
07699

782
793
1963

INLEIDING TOT DE MEDISCHE STATISTIEK I

Uit de afdeling Statistiek van het Nederlands Instituut voor
Praeventieve Geneeskunde te Leiden
Hoofd: C. A. G. Nass

dep. ex.

VERHANDELING VAN HET
NEDERLANDS INSTITUUT VOOR PRAEVENTIEVE GENEESKUNDE
XLI

INLEIDING TOT DE
MEDISCHE STATISTIEK

DEEL 1

FUNDAMENTELE BEGRIPPEN EN TECHNIEKEN
VERDELINGSVRIJE METHODEN

DOOR

H. DE JONGE

2e DRUK

1963

BIBLIOTHEEK NEDERLANDS INSTITUUT
VOOR PRAEVENTIEVE GENEESKUNDE TNO
WASSENAARSEWEG 56 - LEIDEN

VOORWOORD

Dit boek is bestemd voor artsen en andere werkers op biometrisch gebied, die kennis willen nemen van de statistische technieken voor het analyseren van waarnemingsuitkomsten. Het is ontstaan uit de syllabus, die gedurende enige jaren bij verschillende bio-statistische cursussen door mij werd gebruikt. Ik heb getracht de statistische gedachtengang er zodanig in weer te geven, dat de beschrijving van de statistische technieken niet alleen behoeft te bestaan uit het geven van reeksen 'recepten', die slechts blindelings kunnen worden gevolgd.

Bij deze opzet kan (en mag) de wiskunde niet achterwege blijven. De dosis wiskunde is echter bewust beperkt, zodat zij ook voor de niet-mathematisch geïnteresseerde lezer verteerbaar is. Wèl is een ruim gebruik gemaakt van symbolen en formules, doch deze worden geleidelijk ingevoerd en verduidelijkt. De moeite die men zich aanvankelijk moet getroosten om met deze symbolen en formules vertrouwd te raken wordt echter ruimschoots vergoed door de voordelen, die het gebruik ervan oplevert. Het verdient daarom aanbeveling de onderdelen IA (fundamentele begrippen en technieken) en IB (beginselen van de mathematische statistiek) grondig en in een kalm tempo door te nemen, de voorbeelden nauwkeurig te bestuderen en de opgaven zoveel mogelijk te maken. Men zal dan de technieken, die in IC en in deel II behandeld worden, vlot kunnen volgen.

Deel IC is gewijd aan de *verdelingsvrije* methoden. Zoals bekend mag worden verondersteld behoeft men hierbij geen veronderstellingen te maken omtrent de vorm van de bestudeerde frequentieverdelingen, terwijl uit didactisch oogpunt van belang is dat zij merendeels betrekkelijk eenvoudig kunnen worden afgeleid en gewoonlijk weinig rekenwerk vergen. Het erkennen van deze klaarblijkelijke voordelen van de verdelingsvrije methoden behoeft echter niet in te houden, dat men de *klassieke* methoden – die in deel II aan de orde komen – terzijde schuift. Er bestaan immers ook op medisch-biologisch gebied vele problemen, waarbij men voldoende omtrent de bestudeerde verdelingen weet of te weten kan komen, om tot een verantwoorde toepassing van een klassieke methode over te gaan. Men verkrijgt dan niet alleen meer informatie uit het waarnemingsmateriaal dan met de overeenkomstige verdelingsvrije methode, doch men kan bovendien gebruik maken van verschillende complexe, maar bijzonder doeltreffende proefopzetten, die zich (nog) niet voor een verdelingsvrije bewerking lenen. In de praktijk zal men bij elk probleem tot een bewuste keuze tussen beide soorten methoden moeten komen en in deel II zal mede worden nagegaan, welke factoren daarbij een rol spelen.

Voor de medewerking die ik bij het tot stand komen van dit boek heb ondervonden, ben ik velen dank verschuldigd.

Deze dank gaat in de eerste plaats uit naar het Nederlands Instituut voor Praeventieve Geneeskunde, dat mij in staat stelde deze arbeid ter hand te nemen. Voorts naar de Heer C. A. G. Nass, Hoofd van de Afdeling Statistiek van dit Instituut, die mij wegwijs maakte op het biometrische toepassingsgebied van de statistiek. Zijn lessen en zijn visie op de statistica zijn bepalend geweest voor de door mij gevolgde gedachtengang. De samenwerking binnen deze Afdeling met Dr. M. J. W. de Groot was mede van grote betekenis, daar ik veelvuldig gebruik kon maken van zijn inzichten op het terrein van de medische statistiek; verschillende voorbeelden en opgaven zijn van hem afkomstig.

Ook aan de andere leden van de Afdeling Statistiek ben ik dank verschuldigd en wel in het bijzonder aan de dames J. S. de Graaff en V. F. Ooms en aan de Heer P. van Leeuwen, die vele berekeningen uitvoerden en mij een grote hulp zijn geweest bij de correctie. Vele medewerkers aan het Instituut en een aantal cursisten kwamen met suggesties, die tot wezenlijke verbeteringen hebben geleid.

In vele opzichten is dit boek een compromis. Van opmerkingen, die tot aanvullingen en verbeteringen leiden zal ik dan ook gaarne kennis nemen. Intussen hoop ik, dat het werk reeds in de huidige vorm van nut zal kunnen zijn voor degenen, die de statistiek willen gebruiken bij de oplossing van problemen op medisch-biologisch gebied.

Leiden, October 1958. H. DE JONGE

BIJ DE TWEEDE DRUK

Het verschijnen van de tweede druk van dit boek heeft een aantal verbeteringen en aanvullingen mogelijk gemaakt. Het belangrijkst zijn een aanzienlijke uitbreiding van § 4 van hoofdstuk 1 (de bespreking van onderzoekingen op medisch gebied), een herziening van hoofdstuk 3 (kansrekening) en een aanpassing van hoofdstuk 12 (verdelingsvrije toetsen tegen verloop) aan de in deel II behandelde stof. Verder is de tabellenverzameling in haar geheel in deel I ondergebracht; de antwoorden bij de opgaven en het register behouden hun plaats in deel II.

Veel dank ben ik verschuldigd aan allen die door hun opmerkingen en adviezen aan het tot stand komen van deze herdruk hebben medegewerkt. Mej. J. S. DE GRAAFF ben ik bijzonder erkentelijk voor het geduld waarmee zij, naast haar vele andere werkzaamheden, vrijwel de gehele correctie heeft verzorgd.

Leiden, januari 1963. H. DE JONGE

INHOUDSOPGAVE

A: FUNDAMENTELE BEGRIPPEN EN TECHNIEKEN

1. INLEIDING

1.1. POPULATIE EN STEEKPROEVEN 3
1.1.1. Populaties. 3
1.1.2. Attributen en waarden 5
1.1.3. Frequentieverdelingen 6
1.1.4. Steekproeven 9
1.2. STATISTIEK. 10
1.3. STEEKPROEFTREKKEN 14
1.3.1. Aselecte getallen (toevalcijfers). 14
1.3.2. Aselecte permutaties (lotingstabellen). 16
1.3.3. Laagsgewijs aselecte steekproeven 17
1.3.4. Groepssteekproeven 18
1.3.5. Andere steekproeftechnieken. 19
1.3.6. Onzuivere steekproeven. 20
1.3.7. Opgaven 25
1.4. ONDERZOEKINGEN OP MEDISCH GEBIED 25
1.4.1. Inleiding 25
1.4.2. Onderzoeken. 26
1.4.3. Algemene gang van zaken bij een onderzoek. . 37
1.5. ANALYSE VAN BESTAANDE GEGEVENS 50
1.5.1. Inleiding 50
1.5.2. Statistieken inzake sterfte, ongevalsletsel en ziekte. 53
1.5.3. Andere medische gegevens. 56
1.5.4. Gegevens omtrent medische voorzieningen . . 60
1.6. SAMENVATTING 61

2. REDUCTIE VAN WAARNEMINGSUITKOMSTEN

2.1. FREQUENTIETABELLEN. 62
2.1.1. Attributen (kwalitatieve kenmerken) 62
2.1.2. Waarden 63
2.1.3. Verlies bij reductie tot een frequentieverdeling. 67
2.1.4. De correlatietabel 68
2.1.5. Sorteermethoden. 70
2.2. GRAFIEKEN. 73
2.2.1. Het kolommendiagram 73
2.2.2. De frequentiepolygoon 76
2.2.3. Het spreidingsdiagram 78
2.2.4. Het verloopdiagram 78
2.2.5. Andere grafiekvormen 80
2.2.6. Wenken en waarschuwingen. 81
2.2.7. Opgaven 86

2.3. GEMIDDELDEN 87
2.3.1. De modus en de modale klasse 87
2.3.2. De mediaan 87
2.3.3. Het rekenkundig gemiddelde 88
2.4. SPREIDINGSMATEN 91
2.4.1. De spreidingsbreedte 91
2.4.2. De gemiddelde afwijking 91
2.4.3. Fractielen en het kwartieleninterval 91
2.4.4. De variantie en de standaardafwijking 92
2.4.5. Opgaven 95
2.5. VEREENVOUDIGDE BEREKENING VAN HET GEMIDDELDE EN DE STANDAARDAFWIJKING 95
2.5.1. Enkele eigenschappen 95
2.5.2. Berekeningsmethoden 96
2.5.3. Opgaven 98
2.6. KARAKTERISERING VAN EEN FREQUENTIEVERDELING . 98
2.6.1. Enkele verdelingstypen 98
2.6.2. Karakterisering van pseudo-normale verdelingen 103
2.6.3. Karakterisering van andere verdelingstypen . . 104
2.6.4. Populaties en steekproeven 105
2.7. KENGETALLEN VOOR STERFTE EN ZIEKTE 106
2.7.1. Kengetallen voor sterfte 106
2.7.2. Kengetallen voor ziekte 111
2.7.3. Literatuur 111

3. KANSREKENING

3.1. HET BEGRIP KANS IN HET DAGELIJKS LEVEN 113
3.2. KANSDEFINITIE 113
3.3. KANSVERDELINGEN 114
3.3.1. Inleiding 114
3.3.2. Gemiddelde en variantie van een kansverdeling 116
3.4. KANSREGELS 117
3.4.1. De somregel 117
3.4.2. De productregel 117
3.4.3. Afleiding van kansverdelingen 118
3.4.4. Gemiddelde en variantie van een afgeleide kansverdeling 121
3.4.5. Opgaven 122
3.5. KANSVERDELINGEN VAN STEEKPROEFFUNCTIES 123
3.5.1. Steekproeffuncties 123
3.5.2. Steekproeven met teruglegging 124
3.5.3. Steekproeven zonder teruglegging 127
3.5.4. Toepassing 130
3.5.5. Opgaven 133
3.6. OVERSCHRIJDINGSKANSEN 134
3.6.1. Definities 134
3.6.2. Opgaven 134

B: BEGINSELEN VAN DE MATHEMATISCHE STATISTIEK

4. ENKELE KANSVERDELINGEN

4.1. INLEIDING: STEEKPROEVEN UIT EEN DICHOTOMIE . . . 136
4.2. HYPERGEOMETRISCHE VERDELINGEN 137
4.2.1. Afleiding, gemiddelde en variantie 137
4.2.2. Toepassing 138
4.2.3. Opgaven 139
4.3. BINOMIALE VERDELINGEN 140
4.3.1. Afleiding, gemiddelde en variantie 140
4.3.2. Tabellen van binomiale verdelingen. 143
4.3.3. Aanpassing van een binomiale verdeling. . . . 145
4.3.4. Binomiale benadering van een hypergeometrische verdeling 146
4.3.5. Opgaven 146
4.4. POISSON-VERDELINGEN 148
4.4.1. Een POISSON-verdeling als benadering van een binomiale verdeling. 148
4.4.2. Berekening 149
4.4.3. POISSON-verdelingen als exacte kansverdelingen 150
4.4.4. Aanpassing van een POISSON-verdeling 151
4.4.5. Opgaven 152
4.5. NORMALE VERDELINGEN 152
4.6. SAMENVATTING 153

5. NORMALE VERDELINGEN

5.1. EEN NORMALE VERDELING ALS LIMIET VAN EEN BINOMIALE 154
5.2. NORMALE VERDELINGEN EN DE STANDAARDNORMALE VERDELING. 156
5.3. DE STANDAARDNORMALE VERDELING 159
5.3.1. Continue kansverdelingen 159
5.3.2. Linkse, rechtse en tweezijdige overschrijdingskansen 160
5.3.3. Tabellen van de standaardnormale verdeling . 161
5.3.4. Opgaven 164
5.4. BENADERING VAN EEN BINOMIALE VERDELING DOOR DE BEST-PASSENDE NORMALE VERDELING 165
5.4.1. Binomiale verdelingen met $P = Q = \frac{1}{2}$ 165
5.4.2. Binomiale verdelingen met $P \neq \frac{1}{2}$ 168
5.4.3. Opgaven 168
5.5. AANPASSING VAN EEN NORMALE VERDELING 169
5.5.1. Berekening van de best-passende normale verd. 169
5.5.2. Normaal waarschijnlijkheidspapier 169
5.5.3. Opgaven 170

6. TOETSINGSTHEORIE

6.1. GRONDSLAGEN VAN DE TOETSINGSTHEORIE 171
6.1.1. Probleemstelling 171
6.1.2. Termen en begrippen 172
6.1.3. De verdelingen van de toetsingsgrootheid onder H_0 en H_1 172
6.1.4. Het opstellen van een beslissingsschema. . . . 173
6.1.5. De invloed van de steekproefomvang op de fouten-kansen 178
6.1.6. Het vaststellen van de steekproefomvang bij gegeven waarden van P_0, P_1, α en β 181
6.1.7. Opgaven 184
6.2. TOETSING VAN DE HYPOTHESE $P = 0{,}5$ 184
6.2.1. Algemene gang van zaken. 184
6.2.2. Eénzijdige of tweezijdige toetsing. 186
6.2.3. De werkingskromme 187
6.2.4. De uitvoering van de toetsing via overschrijdingskansen 190
6.2.5. Bepaling van overschrijdingskansen. 190
6.2.6. Opgaven 192
6.3. TOETSING VAN DE HYPOTHESE $P = P_0 \neq 0{,}5$. . . . 193
6.3.1. Exacte toetsing 193
6.3.2. Normale benadering 193
6.3.3. Opgaven 194
6.4. SEQUENTE TOETSING. 194
6.4.1. Algemene gang van zaken. 194
6.4.2. Rechts éénzijdige sequente binomiale toetsing . 195
6.4.3. Sequente binomiale toetsing met drie beslissingsmogelijkheden 201
6.4.4. Opgaven 206
6.5. SAMENVATTING 207

7. SCHATTINGSTHEORIE

7.1. GRONDSLAGEN VAN DE SCHATTINGSTHEORIE 209
7.1.1. Termen en begrippen 209
7.1.2. Puntschattingen 210
7.1.3. Intervalschattingen. 210
7.2. BETROUWBAARHEIDSGRENZEN VAN EEN FRACTIE . . . 211
7.2.1. Exacte grenzen 211
7.2.2. Grafieken met betrouwbaarheidsgrenzen van P. 214
7.2.3. Tabellen met betrouwbaarheidsgrenzen van P . 215
7.2.4. Berekening van betrouwbaarheidsgrenzen van P 218
7.2.5. Berekening van de steekproefomvang, nodig voor het verkrijgen van een betrouwbaarheidsinterval van gegeven breedte 219
7.2.6. Opgaven 220
7.3. SCHATTEN EN TOETSEN 221

C: METHODEN DER MATHEMATISCHE STATISTIEK

I. Verdelingsvrije methoden

8. OVERZICHT VAN DE VERDELINGSVRIJE METHODEN

8.1. MATHEMATISCH-STATISTISCHE METHODEN 225
8.2. VERDELINGSVRIJE METHODEN. 226
8.2.1. De verdelingsvrije methoden in de hoofdstukken 9 en 10 226
8.2.2. De verdelingsvrije methoden in hoofdstuk 11 . 229
8.2.3. De verdelingsvrije toetsen in hoofdstuk 12. . . 229
8.2.4. Andere verdelingsvrije methoden. 230
8.2.5. De ontwikkeling van verdelingsvrije methoden. 230

9. KWALITATIEVE VERDELINGSVRIJE TOETSEN

9.1. POPULATIEVERDELINGEN MET r CATEGORIEËN. 232
9.1.1. Steekproeven zonder teruglegging 232
9.1.2. Steekproeven met teruglegging. 232
9.2. TOETSING VAN DE HYPOTHESE $P_1 = P_2 = P_3$ 234
9.2.1. Exacte toetsing 234
9.2.2. G als maat voor de discrepantie tussen waargenomen en verwachte frequenties 238
9.2.3. χ^2-verdelingen 241
9.2.4. Opgaven 244
9.3. DE χ^2-TOETS BETREFFENDE EEN POPULATIE MET r CATEGORIEËN 245
9.3.1. Uitvoering van de toetsing 245
9.3.2. Populatieverdelingen met 2 categorieën 247
9.3.3. Toepassing op kwantitatieve waarnemingen . . 248
9.3.4. Combinatie van categorieën bij χ^2-toetsen . . . 248
9.3.5. Opgaven 249
9.4. DE TEKENTOETS 250
9.4.1. Toepassing 250
9.4.2. De toetsingsgrootheid. 250
9.4.3. De verdeling van de toetsingsgrootheid onder H_0 250
9.4.4. Uitvoering van de toetsing 251
9.4.5. Toepassing op kwantitatieve waarnemingen . . 252
9.4.6. Opgaven 252
9.5. DE TOETS VAN COCHRAN. 252
9.5.1. Toepassing 252
9.5.2. De toetsingsgrootheid en haar verdeling onder H_0 253
9.5.3. Uitvoering van de toetsing 254
9.6. DE TOETS VAN FISHER (methode der 2×2 tabel) . . . 254
9.6.1. Toepassing 254
9.6.2. De toetsingsgrootheid. 255
9.6.3. De verdeling van de toetsingsgrootheid onder H_0 255
9.6.4. Exacte toetsing 256

9.6.5. De normale benadering 258
9.6.6. De χ^2-benadering. 259
9.6.7. Betrouwbaarheidsgrenzen van $P_1 - P_2$ 262
9.6.8. Toetsing voor kleine kansen bij even grote steekproeven. 262
9.6.9. Opgaven 263
9.7. DE χ^2-TOETS OP DE $2 \times r$ TABEL. 264
9.7.1. Toepassing 264
9.7.2. De toetsingsgrootheid. 265
9.7.3. De verdeling van de toetsingsgrootheid onder H_0 266
9.7.4. Uitvoering van de toetsing 267
9.7.5. Opgaven 267
9.8. DE χ^2-TOETS OP DE $k \times 2$ TABEL. 267
9.9. DE χ^2-TOETS OP DE $k \times r$ TABEL. 268
9.10. GEBRUIK VAN DE χ^2-TOETSEN BIJ TOETSING VAN DE HYPOTHESE, DAT TWEE CATEGORISCHE SYSTEMEN STOCHASTISCH ONAFHANKELIJK ZIJN 269
9.10.1. Toepassing 269
9.10.2. De toetsingsgrootheid. 270
9.10.3. Verdeling van de toetsingsgrootheid onder H_0 en uitvoering van de toetsing 270
9.10.4. Opgaven 271

10. KWANTITATIEVE VERDELINGSVRIJE TOETSEN

10.1. DE χ^2-TOETS VOOR AANPASSING. 273
10.1.1. Toepassing 273
10.1.2. Uitvoering van de toetsing 273
10.1.3. Opgaven 275
10.2. DE TOETS VAN KOLMOGOROV-SMIRNOV (*K-S* toets) . . . 275
10.2.1. Toepassing 275
10.2.2. Uitvoering van de toetsing 275
10.2.3. Schatting van de populatieverdeling 277
10.2.4. Toepassing op discreet verdeelde waarnemingen 277
10.2.5. Opgaven 278
10.3. DE RANG-TEKENTOETS 279
10.3.1. Toepassing 279
10.3.2. De toetsingsgrootheid. 280
10.3.3. De verdeling van de toetsingsgrootheid onder H_0 281
10.3.4. Uitvoering van de toetsing zonder gelijken. . . 282
10.3.5. Uitvoering van de toetsing met gelijken. . . . 284
10.3.6. Opgaven 286
10.4. DE TOETS VAN FRIEDMAN 288
10.4.1. Toepassing 288
10.4.2. De toetsingsgrootheid. 288
10.4.3. De verdeling van de toetsingsgrootheid onder H_0 289
10.4.4. Uitvoering van de toetsing zonder gelijken. . . 290
10.4.5. Uitvoering van de toetsing met gelijken. . . . 291
10.4.6. Opgaven 292

10.5. DE TOETS VAN WILCOXON 293
10.5.1. Toepassing 293
10.5.2. De toetsingsgrootheid. 294
10.5.3. De verdeling van de toetsingsgrootheid onder H_0 295
10.5.4. Uitvoering van de toetsing zonder gelijken . . 295
10.5.5. Uitvoering van de toetsing met gelijken . . . 297
10.5.6. Rekenschema 298
10.5.7. Opgaven 300
10.6. DE TOETS VAN KRUSKAL & WALLIS 301
10.6.1. Toepassing 301
10.6.2. De toetsingsgrootheid. 301
10.6.3. De verdeling van de toetsingsgrootheid onder H_0 302
10.6.4. Uitvoering van de toetsing zonder gelijken. . . 302
10.6.5. Uitvoering van de toetsing met gelijken. . . . 303
10.6.6. $k = 2$: overgang in de toets van WILCOXON . . 304
10.6.7. Opgaven 304
10.7. DE RANGCORRELATIETOETS VAN SPEARMAN. 305
10.7.1. Toepassing 305
10.7.2. De toetsingsgrootheid. 306
10.7.3. De verdeling van de toetsingsgrootheid onder H_0 306
10.7.4. Uitvoering van de toetsing zonder gelijken. . . 306
10.7.5. Uitvoering van de toetsing met gelijken. . . . 308
10.7.6. De rangcorrelatiecoëfficiënt van SPEARMAN . . 309
10.7.7. Opgaven 310
10.8. DE RANGCORRELATIETOETS VAN KENDALL 310
10.8.1. Toepassing 310
10.8.2. De toetsingsgrootheid. 310
10.8.3. De verdeling van de toetsingsgrootheid onder H_0 312
10.8.4. Uitvoering van de toetsing zonder gelijken. . . 313
10.8.5. Uitvoering van de toetsing met gelijken . . . 313
10.8.6. De rangcorrelatiecoëfficiënt van KENDALL . . . 314
10.8.7. Opgaven 314
10.9. KEUZE VAN EEN RANGCORRELATIETOETS. 315

11. VERDELINGSVRIJE TOETSEN TEGEN VERLOOP IN EEN WAARNEMINGSREEKS

11.1. INLEIDING 317
11.2. DE RANGCORRELATIETOETSEN ALS TOETSEN TEGEN VERLOOP 320
11.3. DE SERIETOETS OP DE WAARNEMINGEN 321
11.3.1. Toepassing 321
11.3.2. De toetsingsgrootheid. 321
11.3.3. De verdeling van de toetsingsgrootheid onder H_0 322
11.3.4. Uitvoering van de toetsing 323
11.4. TOETS TEGEN VERLOOP VAN EEN KANS 325
11.5. DE SERIETOETS OP DE VERSCHILLEN TUSSEN OPEENVOLGENDE WAARNEMINGEN. 326
11.5.1. Toepassing 326

11.5.2. De toetsingsgrootheid. 326
11.5.3. De verdeling van de toetsingsgrootheid onder H_0 327
11.5.4. Uitvoering van de toetsing 327
11.6. DE TOETS OP DE TEKENS VAN DE OPEENVOLGENDE WAARNEMINGEN 328
11.7. TOETSEN GEBASEERD OP DE SERIELENGTE 329
11.8. KEUZE VAN EEN TOETS TEGEN VERLOOP. 329
11.9. OPGAVEN. 332

12. VERDELINGSVRIJE TOETSEN TEGEN VERLOOP IN MEERDERE WAARNEMINGSREEKSEN
12.1. INLEIDING 333
12.2. TOETS TEGEN VERLOOP IN k VERWANTE WAARNEMINGSREEKSEN. 334
12.2.1. Toepassing 334
12.2.2. De toetsingsgrootheid. 335
12.2.3. Uitvoering van de toetsing 336
12.3. TOETS TEGEN VERLOOP IN k VERWANTE WAARNEMINGSREEKSEN MET DICHOTOME WAARNEMINGEN. 337
12.3.1. Toepassing 337
12.3.2. De toetsingsgrootheid. 338
12.4. TOETS TEGEN VERLOOP IN k ASELECTE STEEKPROEVEN 339
12.4.1. Toepassing 339
12.4.2. De toetsingsgrootheid. 340
12.4.3. Uitschakeling van de invloed van verschillen in steekproefomvang 342
12.5. TOETS TEGEN VERLOOP IN EEN $k \times 2$ TABEL 343
12.6. OPGAVEN. 347

BIJLAGEN

I. TABELLENVERZAMELING 349

II. ELEMENTAIRE ALGEBRA 395
1. Het gebruik van symbolen 395
2. Permutaties en combinaties. 397
3. Exponenten en wortels 399
4. Logarithmen 400
5. Benaderende getallen. 402
6. Coördinaatassen 405

III. LITERATUUROPGAVE 408

A

FUNDAMENTELE BEGRIPPEN EN TECHNIEKEN

'The statistical method is essentially a *technic*, which finds its justification in its usefulness in helping to solve the problems of the basic sciences. Statistics, in any proper sense, has no, or at best few, problems of its own. Its technical problems are really problems of mathematics. The statistical method is, or should be, a working tool of science, just as is the microscope or the kymograph. But it is probably of wider utility than any other single technical method which science has discovered or devised. For it has an applicability and a usefulness, direct or indirect, in virtually every problem. It is, in short, a fundamental element of scientific methodology.'

R. PEARL

HOOFDSTUK 1

INLEIDING

De statistiek is een hulpwetenschap die methoden verschaft voor het bewerken, presenteren, analyseren en interpreteren van waarnemingsuitkomsten.

De voorgaande omschrijving is noodgedwongen vaag. Zij kan pas worden verscherpt door de specifieke 'taal' te gebruiken, die de statistiek voor allerlei termen en begrippen heeft ontwikkeld. Een ieder die zich in de statistiek wil verdiepen dient zich eerst deze taal eigen te maken en wij bespreken dus in de volgende paragraaf een aantal termen en begrippen die in de statistiek een centrale plaats innemen.[1]

Ten aanzien van de indeling van dit boek en de voor allerlei onderdelen gebruikte nummering geldt het volgende:

1. Evenals de paragrafen zijn de tabellen, figuren, formules, voorbeelden en opgaven per hoofdstuk doorlopend genummerd. 'Zie 3.2' betekent dus 'zie hoofdstuk 3, paragraaf 2'; figuur 2.1 is dus de eerste figuur in hoofdstuk 2. Bij verwijzing naar formules is het nummer van de formule steeds tussen haakjes geplaatst: 'volgens (3.4)' betekent dus 'volgens formule 3.4'.
2. De paragrafen die met een sterretje (*) zijn aangegeven kunnen bij eerste lezing zonder bezwaar worden overgeslagen.
3. Alle met *hoofdletters* aangeduide tabellen staan in de tabellenverzameling (bijlage 1).
4. Een verwijzing naar de alfabetische literatuuropgave in bijlage 3 vindt plaats door vermelding van de naam van de auteur met tussen haakjes het jaar van publikatie. Wanneer een auteur in één jaar meerdere publikaties op zijn naam heeft, zijn deze met letters (a, b, . . .) onderscheiden.
5. In bijlage 2 zijn enkele onderwerpen uit de algebra opgenomen, die men zo nodig kan raadplegen. Bijlage 4 (in deel twee) bevat de antwoorden bij de opgaven.

1.1. Populatie en steekproeven

1.1.1. POPULATIES

Een wetenschappelijk onderzoek komt in principe neer op het bestuderen van eigenschappen van populaties. In dit verband verstaat men onder een *populatie* (Eng. population) de verzameling van operationeel gedefinieerde eenheden ('elementen'), waarop men de conclusies

[1] Wij volgen hierbij — behoudens een enkele uitzondering — het normblad NEN 3117, Statistische termen.

van het onderzoek wil betrekken. Een operationele definitie is een voorschrift met behulp waarvan kan worden vastgesteld of een element al dan niet tot de populatie behoort.

De elementen van een populatie kunnen in de eerste plaats materiële objecten zijn (zoals mensen, dieren of voorwerpen). Zij kunnen echter ook gebeurtenissen zijn (ongelukken, huwelijken, sterfgevallen) of abstracte of denkbeeldige elementen (waarnemingsuitkomsten, getallen, ogenblikkelijke toestanden van materiële objecten).

Het aantal elementen van een populatie noemt men de *populatieomvang* en als symbool voor deze omvang gebruiken wij steeds de (hoofd)letter N. Sommige populaties zijn *onbegrensd* (zie voorbeeld 1.5).

Voorbeelden [1]

Om het populatiebegrip te verduidelijken geven wij een reeks voorbeelden van verschillende soorten populaties:

1.1. De dienstplichtigen van lichting 1962 in Nederland (de operationele definitie is vervat in de Dienstplichtwet).

1.2. De mannen van 20 t/m 64 jaar die op 1 januari 1960 in een bepaald bedrijf werkzaam waren.

1.3. De mannen van 20 t/m 24 jaar die op 1 januari 1960 in dat bedrijf werkzaam waren. Dit is een *deelpopulatie* van de populatie in 1.2.

1.4. De gehele getallen van 1 t/m 1000.

1.5. Wegingen van een bepaald voorwerp, die een bepaalde waarnemer met een bepaalde balans op voorgeschreven wijze heeft verricht of zou kunnen verrichten.

1.6. De kinderen van 6 en 7 jaar die op 30 september 1957 op 3 bepaalde lagere scholen in de stad R. aanwezig waren.

1.7. Het aantal huishoudens in Nederland op 30 juni 1956, volgens de Algemene Woningtelling 1956 (incl. de huishoudens in gestichtswoningen en op woonschepen, doch excl. de huishoudens in militaire inrichtingen, in gestichten en onder de varende en rijdende bevolking).

1.8. Het aantal geregistreerde verkeersongevallen op de openbare weg in Nederland in 1959.

Vervolgens enkele voorbeelden waarbij de operationele definitie ontbreekt of onvoldoende is:

1.9. De bomen in een bepaald bos (zonder dat is aangegeven wanneer een jonge plant een boom wordt genoemd). Het is overigens wel duidelijk dat men bij inventarisatie van het bos met de bedoeling tot het kappen en verkopen van de bomen over te gaan een andere definitie van 'boom' zal gebruiken dan wanneer men dit onderzoek uitvoert om uit te maken hoe het bos zich in de eerstkomende 25 jaar zal ontwikkelen.

1.10. De Indianen in een bepaald land (zonder dat ook voor personen van gemengd bloed wordt aangegeven wanneer zij als Indianen worden beschouwd).

Bij die onderzoeken op medisch gebied, waarbij de populatie bestaat uit patiënten met een bepaalde aandoening, is een nauwkeurige definitie daarvan bijzonder belangrijk. Gewoonlijk kan men in zo'n geval niet volstaan met het geven van een 'etiket' (in de zin van: de diabetici in Nederland), maar zal men

[1] Hoewel bij deze en vele volgende voorbeelden een kleine druk is gebruikt om ruimte te sparen, vormen zij een zo wezenlijk deel van de tekst dat zij *niet* kunnen worden overgeslagen.

voor nauwkeurig omschreven diagnostische criteria moeten zorgen. WITTS c.s. (1959) bespreken uitvoerig de voorwaarden waaraan deze criteria moeten voldoen.

1.1.2. ATTRIBUTEN EN WAARDEN

Een kenmerk is een eigenschap, die de elementen van een populatie kunnen bezitten. Men spreekt van:

a. Een *attribuut* (kwalitatief kenmerk, categorie) als de eigenschap kwalitatief is (bijv. de kleur blauw, indien het over kleuren gaat).

b. Een *waarde* als het kenmerk kwantitatief is (bijv. de leeftijd van 20 jaar als het over leeftijden gaat).

Een combinatie van kenmerken vormt zelf ook een kenmerk (bijv. de leeftijd van 20 t/m 24 jaar).

Een *categorisch systeem* wordt gedefinieerd met betrekking tot een bepaalde populatie en is een zodanige verzameling van attributen (kwalitatieve kenmerken) dat ieder element van deze populatie geacht wordt er juist één van te bezitten. Een categorisch systeem bestaande uit twee kenmerken heet een *dichotomie* of tweedeling. Categorische systemen met r (meer dan 2) kenmerken kunnen worden onderscheiden in:

1. *Niet-geordende categorische systemen.* De attributen bezitten geen natuurlijke of logische volgorde.
2. *Geordende categorische systemen.* De attributen bezitten een natuurlijke of logische volgorde.

Voorbeelden

1.11. De toevoeging 'met betrekking tot een bepaalde populatie' in de definitie van een categorisch systeem is noodzakelijk. Een categorisch systeem (van enkelvoudige kleuren) kan categorisch zijn met betrekking tot één populatie (van effen stoffen), maar niet met betrekking tot een andere (van bonte stoffen).

1.12. Een categorisch systeem met drie kenmerken met betrekking tot de populatie van dienstplichtigen in voorbeeld 1.1 is: 'goedgekeurd — afgekeurd — niet gekeurd'.

1.13. Een niet-geordend categorisch systeem met betrekking tot de populaties in de voorbeelden 1.1, 1.2, 1.3, 1.6 en 1.10 vormen de bloedgroepen: O, A, B, AB.

1.14. Met betrekking tot de populatie in voorbeeld 1.7 vormen de kenmerken 'huishoudingen zonder echtpaar — huishoudingen met echtpaar' een dichotomie.

1.15. Met betrekking tot de populatie in voorbeeld 1.8 vormen de kenmerken 'met dodelijke afloop — met ten hoogste ernstig letsel — met ten hoogste licht letsel — met uitsluitend materiële schade' een categorisch systeem met 4 attributen. Beschouwt men dit systeem ten opzichte van de ernst van het aan mensen toegebrachte letsel, dan is het een geordend categorisch systeem.

1.16. Met betrekking tot de populatie in voorbeeld 1.6 vormen de graden 'goed — voldoende — matig — onvoldoende — slecht', die bij medisch onderzoek worden toegekend op grond van de fysieke conditie van de kinderen, een geordend categorisch systeem.

Een *grootheid* (Eng. variable) wordt gedefinieerd met betrekking tot een bepaalde populatie en is een zodanige verzameling van waarden (kwantitatieve kenmerken), dat ieder element van deze populatie geacht wordt er juist één van te bezitten.

Een grootheid heet *discreet* (of discontinu) als zij uitsluitend een aantal afzonderlijke waarden kan aannemen (voorbeelden: het aantal kinderen per gezin, het aantal ziekmeldingen per arbeider per jaar). Bij een discrete grootheid worden de waarnemingsuitkomsten door tellen verkregen. Elke waarde van de grootheid is in het algemeen een geheel getal en het verschil tussen twee waarden bedraagt dus steeds nul, één of meer eenheden.

Een grootheid heet *continu* als zij binnen zekere grenzen iedere waarde kan aannemen (voorbeelden: de lichaamslengte, de systolische bloeddruk, het tijdsverloop tussen twee gebeurtenissen). Bij een continue grootheid worden de uitkomsten door meten of wegen verkregen. Tengevolge van de begrensde nauwkeurigheid van iedere meettechniek worden deze uitkomsten in werkelijkheid steeds in *afgeronde* waarden uitgedrukt en deze zijn dus eveneens discreet. Men noemt het kleinst mogelijke verschil — nul niet meegerekend — tussen twee van deze afgeronde waarden het *afrondingsinterval.*

Voorbeelden

1.17. Een discrete grootheid met betrekking tot de populatie in voorbeeld 1.8 is het aantal personen dat door het ongeluk is gedood. Deze grootheid kan de waarden 0, 1, 2, ... aannemen.

1.18. Een discrete grootheid met betrekking tot de populatie in voorbeeld 1.6 is het aantal kinderen met struma.

1.19. Een continue grootheid met betrekking tot de populaties in de voorbeelden 1.2 en 1.3 is de leeftijd van de mannen. Deze leeftijd wordt gewoonlijk in gehele jaren uitgedrukt, waarbij een leeftijd van 35 jaar betekent, dat de 35-ste verjaardag wel, maar de 36-ste nog niet is gevierd. In symbolen uitgedrukt betekent $x = 35$ dus dat $35 < x < 36$.

1.20. Een continue grootheid met betrekking tot de populatie in voorbeeld 1.6 is het hemoglobinegehalte van de kinderen. Gewoonlijk drukt men de waarden van deze grootheid in g% uit met een afrondingsinterval van 0,1 g%.

1.1.3. FREQUENTIEVERDELINGEN

Onder *variabiliteit* verstaat men het verschijnsel dat de elementen van een populatie verschillende kenmerken van een categorisch systeem, resp. verschillende waarden van een grootheid bezitten. *Fluctuatie* is het verschijnsel van variabiliteit bij elementen, die op één of andere wijze (bijv. in volgorde van waarneming in de tijd) zijn gerangschikt.

Het aantal elementen van een populatie dat een bepaald kenmerk (of een bepaalde waarde) bezit noemt men de (absolute) *frequentie* van

dat kenmerk (van die waarde).[1] Voor een frequentie gebruikt men het symbool f en met een index kan worden aangegeven op welk kenmerk of op welke waarde de frequentie betrekking heeft. Zo betekent bijv. f_A: de frequentie van kenmerk A (van een bepaald categorisch systeem), en f_{x_1}: de frequentie van de waarde x_1 (van de grootheid x). Deelt men een frequentie door de omvang van de populatie, dan verkrijgt men de fractie $p = f/N$, die de *relatieve frequentie* of het *frequentiequotiënt* wordt genoemd.

De verzameling van de (relatieve) frequenties van een categorisch systeem, resp. van de waarden van een grootheid, noemt men de (relatieve) *frequentieverdeling* (Eng. frequency distribution) van dat systeem, resp. van die grootheid. Een frequentieverdeling die betrekking heeft op een populatie zullen wij met de term *populatieverdeling* aanduiden.

Een getal waardoor een populatieverdeling kan worden gekarakteriseerd noemt men een *parameter* van die verdeling. Zo wordt bijv. de populatie waarvan elk element kenmerk A òf kenmerk B draagt volledig gekarakteriseerd door de fractie (resp. het percentage) elementen met kenmerk A (òf door de fractie met kenmerk B).

De *somfrequentie* van een waarde van een grootheid kan worden verkregen door de frequentie van deze waarde te vermeerderen met de som van de frequenties van alle lagere waarden. Een somfrequentie (ook wel cumulatieve frequentie genoemd) wordt aangegeven met het symbool F. De relatieve somfrequentie F/N noemt men ook wel de somfunctie. De verzameling van somfrequenties van een grootheid noemt men de *somfrequentieverdeling*.

Voorbeelden

1.21. Frequentieverdeling van het categorisch systeem in voorbeeld 1.12 op de populatie in voorbeeld 1.1 (fictieve gegevens).

Kenmerk	Frequentie (f)
Goedgekeurd	66647
Afgekeurd	18518
Niet gekeurd	2147
Totaal	87312 = N

1.22. Frequentieverdeling en relatieve frequentieverdeling van het categorisch systeem in voorbeeld 1.13 op de populatie in voorbeeld 1.2.

Kenmerk	Frequentie f	Freq. quotiënt f/N
Bloedgroep O	4621	4621/11000 = 0,420
Bloedgroep A	4895	4895/11000 = 0,445
Bloedgroep B	1213	1213/11000 = 0,110
Bloedgroep AB	271	271/11000 = 0,025
Totaal	11000 = N	1,000

[1] In het vervolg zullen wij onder frequentie steeds de absolute frequentie verstaan.

1.23. Frequentieverdeling en relatieve frequentieverdeling van het categorisch systeem in voorbeeld 1.15 op de populatie in voorbeeld 1.8.

Kenmerk	f	f/N
Dodelijke afloop	1622	0,011
Ten hoogste ernstig letsel	22607	0,150
Ten hoogste licht letsel	14795	0,098
Uitsluitend materiële schade	111746	0,741
Totaal	150770	1,000

1.24. Frequentieverdeling, somfrequentieverdeling en relatieve somfrequentieverdeling van de grootheid in voorbeeld 1.19 op de populatie in voorbeeld 1.3.

Waarde	Frequentie f	Somfrequentie F	Rel. somfreq. F/N
24 (jaar)	485	1625	1,000
23 (jaar)	354	1140	0,700
22 (jaar)	361	786	0,483
21 (jaar)	227	425	0,262
20 (jaar)	198	198	0,122
Totaal	1625	—	—

Een kwantitatief kenmerk dat bestaat uit een interval, d.i. een aantal opeenvolgende waarden van een grootheid, noemt men een *klasse*. De eindpunten van zo'n interval noemt men de *klassegrenzen* en de afstand tussen deze grenzen de *klassebreedte*. Onder de *klassewaarde* verstaat men het midden van het interval, d.i. de waarde die men verkrijgt door te nemen (bovengrens + ondergrens)/2.

Voorbeeld

1.25. Frequentieverdeling, somfrequentieverdeling, relatieve frequentieverdeling en relatieve somfrequentieverdeling van de grootheid in voorbeeld 1.20 op de populatie in voorbeeld 1.6.

De hemoglobinegehalten zijn op 1 en 2 oktober 1957 bepaald door analyste A met Sicca-hemometer nr. . . . . op bloed verkregen uit een vingerprik. Het hemo-

Klassegrenzen $x_o - x_b$	Klassewaarde x	f	F	f/N	F/N
15,75—16,25	16,0	2	190	0,011	1,000
15,25—15,75	15,5	3	188	0,016	0,989
14,75—15,25	15,0	11	185	0,058	0,973
14,25—14,75	14,5	19	174	0,100	0,915
13,75—14,25	14,0	26	155	0,137	0,815
13,25—13,75	13,5	46	129	0,242	0,678
12,75—13,25	13,0	38	83	0,200	0,436
12,25—12,75	12,5	24	45	0,126	0,236
11,75—12,25	12,0	15	21	0,079	0,110
11,25—11,75	11,5	4	6	0,021	0,031
10,75—11,25	11,0	1	2	0,005	0,010
10,25—10,75	10,5	1	1	0,005	0,005
Totaal		190	—	1,000	—

globinegehalte vormt een continue grootheid, die in discrete waarden met een afrondingsinterval van 0,1 g% is gemeten. De waarden zijn in klassen van 5, die dus een klassebreedte van 0,5 g% bezitten, bijeengebracht. De laagste klasse bevat de waarden 10,3, 10,4, 10,5, 10,6 en 10,7 g% en heeft als klassemidden = klassewaarde 10,5 g%. De ondergrens x_o van deze klasse ligt $\frac{1}{2} \times 0{,}5 =$ $= 0{,}25$ g% beneden de klassewaarde en is 10,25 g%; de bovengrens x_b is $10{,}5 + 0{,}25 = 10{,}75$ g% (zie ook 2.1.2).

1.1.4. STEEKPROEVEN

In de praktijk is het gewoonlijk niet mogelijk (en soms ook niet noodzakelijk) om de volledige populatie in een onderzoek te betrekken. In de eerste plaats zijn vele populaties zo groot of bezitten hun elementen een zo grote geografische spreiding, dat een volledig onderzoek te tijdrovend en te kostbaar wordt. Voor andere populaties geldt dat de elementen door het onderzoek geheel of gedeeltelijk verloren gaan. Met deze moeilijkheid heeft men vooral in de industrie te maken: men kan bijv. niet de brandduur van elke gefabriceerde gloeilamp bepalen, of van elke lucifer die de fabriek verlaat onderzoeken of hij wel ontvlamt. Verder zijn er vele populaties waarvan de elementen niet volledig of nog niet bekend zijn, zoals bijv. de populatie bestaande uit alle personen in Nederland die aan een bepaalde aandoening lijden of zullen gaan lijden. Tenslotte noemen wij populaties die slechts denkbeeldig zijn, zoals de populatie bestaande uit de uitkomsten van alle hemoglobine bepalingen die analyst *A* met een bepaald apparaat op voorgeschreven wijze aan één bloedmonster (van flinke omvang) zou kunnen verrichten. Overigens zijn er vele situaties waarin een onderzoek van de volledige populatie onnodig en zodoende economisch niet verantwoord is, omdat men uit een deel van de populatie reeds voldoend betrouwbare en dus praktisch bruikbare informatie kan verkrijgen.

In deze en soortgelijke gevallen gaat men over tot het samenstellen ('trekken') en onderzoeken van een *monster* of *steekproef* (Eng. sample), d.i. een operationeel gedefinieerde verzameling van elementen uit de te bestuderen populatie. De operationele definitie die aangeeft hoe de steekproef moet worden genomen, heet het *steekproefvoorschrift*. In principe bevat dit voorschrift ook de *steekproefomvang* (Eng. sample size), het aantal elementen dat deel zal uitmaken van de steekproef. Wij zullen dit aantal steeds met de (kleine) letter n aangeven. Daar een steekproef uit elementen van een populatie bestaat gelden de termen en begrippen in 1.1.2 en 1.1.3 ook voor steekproeven. Voor de frequentieverdeling van een categorisch systeem of grootheid met betrekking tot een steekproef voegen wij de term *waargenomen frequentieverdeling* toe.

Indien men, gebruik makend van de methoden der mathematische statistiek, uit steekproeven conclusies wil trekken omtrent de popula-

ties waaruit zij afkomstig zijn, moeten deze steekproeven *aselect* (Eng. at random) getrokken zijn. Deze voorwaarde houdt in dat de trekkingsmethode zo moet zijn, dat elke steekproef van n elementen die uit de N elementen van de populatie kan worden samengesteld, dezelfde kans moet hebben om bij de trekking te verschijnen. Men kan dit bereiken door elk element van de steekproef uit de populatie te trekken door middel van een methode, die onafhankelijk is van alle eigenschappen van het element die voor het onderzoek van belang zijn. Hierbij kan onderscheid worden gemaakt tussen:

1. *Trekking met terugleggíng*: elk getrokken element wordt, nadat hieraan waarnemingen zijn verricht, weer aan de populatie toegevoegd (het wordt 'teruggelegd'), zodat hetzelfde element bij een volgende trekking opnieuw kan verschijnen.
2. *Trekking zonder terugleggíng*: elk getrokken element wordt tijdens het steekproeftrekken niet weer aan de populatie toegevoegd, zodat het bij een volgende trekking niet nogmaals kan worden aangewezen. Deze trekkingsmethode wordt gewoonlijk in de praktijk toegepast.

Er zijn verschillende typen van aselecte steekproeven mogelijk. Een *enkelvoudige* of *volledig aselecte steekproef* (die wij kortweg een *aselecte steekproef* zullen noemen; Eng. (unrestricted) random sample) wordt verkregen indien alle elementen aselect worden getrokken. Een *laagsgewijze* of *gelede aselecte steekproef* (Eng. stratified random sample) wordt samengesteld uit enkelvoudige steekproeven, die aselect getrokken worden uit elk der deelpopulaties, die tezamen de gehele populatie vormen. Een *groepssteekproef* (Eng. clustersample) wordt verkregen door uit de populatie aselect een aantal (gewoonlijk kleine) deelpopulaties (of groepen, Eng. clusters) te trekken.

In 1.3 bespreken wij hoe men in de praktijk aselecte steekproeven kan trekken. Daarbij behandelen wij tevens het *aselecteren* (Eng. to randomize), d.i. de procedure waarmee men op aselecte wijze een groep elementen in twee of meer subgroepen kan verdelen.

1.2. Statistiek

Wij kunnen de omschrijving van de statistiek die wij aan het begin van dit hoofdstuk hebben gegeven nu uitbreiden en verscherpen. Daarbij maken wij een onderscheid tussen de beschrijvende en de mathematische statistiek.

De *beschrijvende statistiek* verschaft methoden, die zowel op waarnemingen betreffende populaties als op waarnemingen betreffende steekproeven kunnen worden toegepast. Deze methoden kunnen worden gebruikt voor:

a. De ordening van waarnemingen, bijv. door het samenstellen van

tabellen waaruit de frequentieverdeling van het bestudeerde categorisch systeem of de bestudeerde grootheid blijkt.

b. Het tekenen van grafische voorstellingen (grafieken), bijv. voor de fluctuatie van waarnemingen of van frequentieverdelingen.

c. Het reduceren van waarnemingen of frequentieverdelingen tot kengetallen, zoals bijv. gemiddelden, spreidings- en correlatiematen.

Deze methoden worden in hoofdstuk 2 van dit boek behandeld.

De *mathematische statistiek* (statistica) verschaft methoden, waarmee men uitgaande van een aselecte steekproef tot uitspraken kan komen omtrent de populatie, waaruit deze steekproef afkomstig is. Naar het karakter van deze uitspraken kan men twee soorten methoden onderscheiden:

a. Methoden voor het *toetsen van hypothesen* betreffende één of meer populatieverdelingen. Deze kunnen worden gebruikt wanneer men een onderzoek instelt naar de juistheid van veronderstellingen betreffende populatieverdelingen.

b. Methoden voor het *geven van schattingen* betreffende één of meer populatieverdelingen.

In vele gevallen betreffen de toetsingen of schattingen één of meer parameters van de bestudeerde populatieverdelingen.

Voorbeelden van toetsingsproblemen

1.26. Men wil voor een bepaalde populatie van patiënten nagaan of de veronderstelling houdbaar is, dat zich hierin evenveel mannen als vrouwen bevinden. Men trekt daartoe uit deze populatie aselect zonder teruglegging een steekproef van 100 personen en vindt hierbij 62 mannen. Is deze uitkomst nog te verenigen met de getoetste hypothese? Of is het verantwoord uit dit onderzoek te concluderen dat de populatie meer mannen dan vrouwen bevat?

1.27. Een onderzoeker wil voor twee uit kinderen bestaande populaties, A en B, nagaan welke van de twee volgende veronderstellingen juist is: 1) Het percentage kinderen met struma is in beide populaties gelijk, 2) Het percentage in populatie A is lager dan dat in populatie B. Hij trekt uit elke populatie, volgens een methode die als aselect kan worden beschouwd, een steekproef van 200 kinderen. De steekproef uit A bevat 21 kinderen met struma (10,5%), de steekproef uit B bevat er 32 (16%). Welke conclusie kan uit deze uitkomsten worden getrokken?

1.28. In de farmacologie wordt de reactietijd op de z.g. hot plate gebruikt om na te gaan of een geneesmiddel een analgetische werking heeft. Deze reactietijd is de tijd die verloopt tussen het ogenblik waarop een muis op een zinken plaat met een temperatuur van ongeveer 56° C wordt gebracht en het moment waarop hij, als afweerreactie, zijn voorpootjes optrekt. Normaliter varieert deze tijd om 10 sec., maar wanneer een stof analgetisch werkt wordt deze reactietijd verlengd.

Men wil de hypothese toetsen dat twee stoffen, X en Y, een even sterke analgetische werking bezitten. Men splitst daartoe 40 muizen van een bepaalde stam aselect in twee groepen van 20 en bepaalt de reactietijden, nadat de ene groep

met een bepaalde dosis van stof X en de andere met eenzelfde dosis van stof Y is ingespoten. Hierbij verkrijgt men dus twee reeksen van reactietijden (x en y, die wij bij dit gefingeerde voorbeeld weglaten), waarop een uitspraak over de werking van de stoffen kan worden gebaseerd.

Voorbeelden van schattingsproblemen

1.29. Men wil voor een bepaalde groep uit de Nederlandse bevolking (die enige honderdduizenden personen telt) weten, welk deel daarvan aan diabetes lijdt. Men onderzoekt een groep van 500 personen uit deze populatie en vindt daarbij (uiteraard volgens bepaalde tevoren gestelde criteria) 25, d.i. 5% diabetici. Welke conclusie kan men uit deze bevinding trekken omtrent het percentage in de populatie?

1.30. Men splitst 40 ratten van een bepaalde stam aselect in vier groepen van 10. Men geeft de dieren een inspuiting met een vitamine D_3 preparaat, waarbij aan de groepen resp. 50000, 100000, 150000 en 200000 I.E. worden gegeven. Na drie dagen bepaalt men van alle dieren het calciumgehalte in het serum. Uit de verkregen gegevens wil men een schatting geven van het verloop van het gemiddelde calciumgehalte bij toeneming van het aantal I.E.

De mathematische statistiek maakt gebruik van de kans- of waarschijnlijkheidsrekening, een onderdeel van de wiskunde. Door middel van deze kansrekening kan worden vastgesteld, welke variatie optreedt in steekproeven van n elementen, die aselect worden getrokken uit een populatie met een (bekende) verdeling van een bepaald type. Deze kennis kan worden aangewend voor het oplossen van het omgekeerde probleem, het trekken van conclusies omtrent een bepaalde populatie van dit type op grond van één aselecte steekproef.

Het is duidelijk dat deze weg niet tot een uitspraak omtrent de populatie voert die zonder enige reserve juist is. Steeds loopt men een zeker *risico*, een zekere kans, dat de op een steekproef gebaseerde uitspraak *onjuist* is. In dit opzicht is een 'statistische' uitspraak dus evenmin exact als een intuïtief getrokken conclusie. De exactheid die door de toepassing van de mathematische statistiek wordt ingevoerd bestaat hieruit, dat men kan aangeven hoe groot dit risico (de onbetrouwbaarheid van de uitspraak) is. Hierbij dient men zich goed te realiseren, dat deze exactheid uitsluitend aanwezig is, indien aan een aantal voorwaarden is voldaan. Wij bespreken deze later uitvoeriger, maar herhalen hier met nadruk de twee basisvoorwaarden:

1. De populatie(s) waaruit de steekproef (-proeven) afkomstig zijn moeten nauwkeurig zijn gedefinieerd (daar de uitspraak alleen voor deze populatie(s) van kracht is).

2. De steekproef (-proeven) moeten aselect zijn samengesteld (omdat alleen voor aselecte steekproeven de kansrekening als basis voor de berekening van de onbetrouwbaarheid kan dienen).

Onder deze (en enkele neven-) voorwaarden (waaraan vaak met minder moeite is te voldoen) kan de statistiek als *'bewijsmiddel'* worden gebruikt: men weet voor welke populatie(s) de uitspraak geldt en welke onbetrouwbaarheid zij bezit.[1]

In de praktijk zal men bij het groepsonderzoek op medisch gebied (zie 1.4) aan deze basisvoorwaarden niet altijd kunnen voldoen. Zo is men bijv. vaak gedwongen om een onderzoek uit te voeren met patiënten uit één kliniek, waarvan men niet precies weet uit welke populatie zij afkomstig zijn (of vrijwel zeker weet dat zij een select deel van de feitelijk te onderzoeken populatie vormen). In deze en soortgelijke gevallen missen de uitspraken die langs statistische weg verkregen zijn de bovenbedoelde 'bewijskracht' (men weet niet meer voor welke populatie en/of met welke onbetrouwbaarheid zij gelden). Zij verliezen daarmee natuurlijk niet alle waarde: men kan de statistiek hier als *'detectiemiddel'* gebruiken en door een nauwgezette analyse van de 'onvolmaaktheden' de uitspraken min of meer aannemelijk maken. Deze uitspraken kunnen dan de basis vormen voor een beter opgezet onderzoek, dat wel tot een statistisch bewijs kan leiden.

Voorbeeld 1.29 betreft een steekproef van 500 personen met 25 diabetici, d.i. 5%. Indien *verondersteld* wordt dat de beschouwde steekproef aselect getrokken is, kan men langs statistische weg de parameter, het percentage diabetici in de populatie, schatten op een waarde tussen 3,3 en 7,3%, waarbij de onbetrouwbaarheid van deze uitspraak 5% (dus de betrouwbaarheid 95%) is. Dit wil zeggen, dat wij mogen verwachten dat van de op deze wijze verkregen uitspraken *gemiddeld* 5% onjuist is.

Deze schatting bezit twee bijzonderheden. In de eerste plaats is zij niet 'precies', maar wordt voor de parameterwaarde een interval aangegeven. Dit interval bepaalt de nauwkeurigheid van de schatting: hoe kleiner het is, des te nauwkeuriger is de schatting. Verder bevat de uitspraak een element van onzekerheid; er is een bepaald risico dat zij onjuist is, en dit wordt de onbetrouwbaarheid van de uitspraak genoemd. Deze onbetrouwbaarheid kan slechts worden verkleind ten koste van de nauwkeurigheid van de schatting; stelt men haar bij het voorbeeld op 1%, dan wordt het schattingsinterval 2,8–8,0%. In de praktijk hangt de keuze van de onbetrouwbaarheid gewoonlijk af van de consequenties die een onjuiste uitspraak heeft.

Door het trekken van een grotere steekproef kan men een schatting verkrijgen, die bij dezelfde onbetrouwbaarheid nauwkeuriger is dan de schatting die een kleinere steekproef oplevert. Vindt men bijv. in een steekproef van 1000 personen 5% diabetici, dan zijn de schattingsintervallen met 5% en 1% onbetrouwbaarheid 3,7–6,5 resp. 3,4–7,0.

De kansrekening wordt in hoofdstuk 3 geïntroduceerd. In de hoofdstukken 4 t/m 7 bespreken wij de gedachtengang van de mathematische statistiek uitvoeriger. Wij gaan daarbij uit van een populatie, waarvan de verdeling een dichotomie is, dus één van twee kenmerken bezit. De daaropvolgende hoofdstukken bevatten een groot aantal mathematisch-statistische technieken, die op populaties met andere verdelingstypen en/of situaties met twee of meer steekproeven betrekking hebben.

[1] Zie HEMELRIJK (1958).

1.3. Steekproeftrekken

1.3.1. ASELECTE GETALLEN (toevalcijfers)

Wanneer men uit een populatie van N elementen aselect met teruglegging een steekproef van n elementen wil trekken, kan men als volgt handelen: 1. Nummer de elementen van de populatie van 1 t/m N. 2. Stel een modelpopulatie samen door bijv. kaartjes met de nummers 1 t/m N in identieke holle metalen kogeltjes te plaatsen. 3. Meng deze kogeltjes in een roterende trommel grondig dooreen, trek vervolgens één kogeltje en noteer het daarbij behorende nummer. 4. Voeg het getrokken kogeltje weer aan de populatie toe, meng de kogeltjes opnieuw grondig, trek weer een kogeltje en noteer daarvan het nummer. 5. Zet de trekking voort totdat n nummers verkregen zijn. 6. Neem de elementen uit de oorspronkelijke populatie, die deze n nummers dragen in de steekproef op. Bij trekking zonder teruglegging kan men volstaan met de kogeltjes grondig te mengen, waarna achtereenvolgens n kogeltjes worden getrokken.

Deze lotingsprocedure kan sterk worden vereenvoudigd. Men kan namelijk volstaan met een modelpopulatie van 10 elementen, genummerd 0, 1, 2, . . ., 9, waaruit men aselect met teruglegging telkens één element trekt. Wil men bijv. een steekproef van 5 elementen met teruglegging trekken uit een populatie van 10000 elementen (die genummerd zijn van 1 t/m 10000), dan kan men achtereenvolgens aselect met teruglegging 4 reeksen van 5 cijfers uit de modelpopulatie trekken. Veronderstel dat men daarbij de volgende uitkomsten verkrijgt:

	Reeks 1	Reeks 2	Reeks 3	Reeks 4
5 cijfers	5	2	9	1
	0	2	9	6
	9	7	1	3
	3	5	8	0
	0	0	2	7

In de steekproef kiest men dan de 5 elementen uit de populatie, die de nummers 5291, 296, 9713, 3580 en 27 dragen (komt de uitkomst 0000 voor dan wordt deze gelezen als 10000). Deze methode kan, met modificaties die wij hier buiten beschouwing laten, voor elke populatieomvang worden gebruikt.

Voor het trekken van aselecte steekproeven behoeft men zelfs deze vereenvoudigde lotingsprocedure niet uit te voeren, daar lange reeksen van op deze wijze vervaardigde zg. *aselecte getallen* (Eng. random numbers) ter beschikking staan,[1] zoals de tabellen van TIPPETT (1927),

[1] Zie voor een uitvoeriger beschouwing omtrent de constructie van deze tabellen en hun onderzoek op aselectiviteit HAMAKER (1948) en ENTERS (1951).

KENDALL en BABINGTON SMITH (1946), THE RAND CORPORATION (1955) en FISHER en YATES (1953). In tabel Y hebben wij als voorbeeld een klein deel van de laatstgenoemde tabel opgenomen.

Gebruik van aselecte getallen voor steekproeftrekken. Veronderstel dat men aselect een steekproef van 5 elementen wil trekken uit de populatie in tabel X, waarvan de elementen de nummers 1 t/m 201 dragen en waarden van 10 t/m 30 bezitten. Men leest dan uit een tabel met aselecte getallen, te beginnen op een willekeurige plaats en systematisch in één richting gaande (langs de kolommen, langs de rijen, diagonaalsgewijs) getallen van drie cijfers af. De eerste vijf getallen van 001 (= 1) tot en met 201 die men daarbij tegenkomt bepalen de nummers van de vijf elementen, die in de steekproef worden opgenomen:

a. Bij trekking met teruglegging wordt een element waarvan het nummer op deze wijze meermalen wordt getrokken steeds opnieuw in de steekproef opgenomen.

b. Bij trekking zonder teruglegging kiest men een element slechts de eerste maal dat zijn nummer verschijnt en slaat het bij volgende keren over.

Voert men een trekking zonder teruglegging uit, te beginnen met kolom 4, regel 26 in tabel Y en kolomsgewijs naar beneden gaande dan leest men af

604 767 169 400 005 768 229 593 395 407 595 065 449 321 130 449 070

De steekproef bestaat dus uit de elementen met de nummers 169, 5, 65, 130 en 70, die blijkens tabel X resp. de waarden 21, 20, 16, 19 en 24 bezitten.

Men kan deze trekkingsmethode voor iedere populatie- en steekproefomvang toepassen. Voor populaties van niet te grote omvang kan men soms beter lotingsgetallen gebruiken (zie 1.3.2).

Gebruik van aselecte getallen voor aselecteren. Ook als men N elementen aselect in 2 of meer groepen wil verdelen kan men met aselecte getallen werken.

Veronderstel dat men een groep van 20 proefdieren aselect wil splitsen in drie groepen van resp. 6, 7 en 7 dieren. Men nummert de dieren dan van 1 t/m 20 en leest vervolgens uit tabel Y zonder teruglegging getallen van twee cijfers af. De 6 getallen van 1 t/m 20 die daarbij het eerst verschijnen bepalen de nummers van de dieren die in de eerste groep worden opgenomen, de daaropvolgende 7 getallen van 1 t/m 20 die van de dieren in de tweede groep en de overblijvende 7 getallen vormen de derde groep. Begint men bijv. in tabel Y met

kolom 2 en regel 8, dan vindt men regel voor regel van links naar rechts gaande de onderstaande reeks getallen

16 95 55 67 **19** 98 **10** 50 71 75 **12** 86 73 58 **07** 44 39 52 38 79 33 21 12 34
29 78 64 56 07 82 52 42 07 44 38 **15** 51 00 **13** 42 99 66 **02** 79 54 57 60 86
32 44 **09** 47 27 96 54 49 **17** 46 09 62 90 52 84 77 27 **08** 02 73 43 28 **18** 18
07 92 46 44 17 16 58 09 79 83 86 19 62 **06**

(Wij hebben deze reeks volledig opgenomen en de getallen die achtereenvolgens de elementen in de subgroepen aanwijzen vet gedrukt. In de praktijk zal men alleen de getallen noteren die men nodig heeft, dus alle getallen groter dan 20 en alle getallen tot en met 20 die men reeds heeft opgeschreven weglaten).

De groepen bevatten dus de dieren met de volgende nummers

Groep	Volgorde van trekking	Dieren in volgorde
1	16 19 10 12 7 15	7 10 12 15 16 19
2	13 2 9 17 8 18 6	2 6 8 9 13 17 18
3	— — — — — — —	1 3 4 5 11 14 20

1.3.2. ASELECTE PERMUTATIES (lotingstabellen)

Wanneer de populatie-omvang of het aantal te aselecteren elementen tussen twee veelvouden van 10 ligt moet men bij het gebruik van aselecte getallen een deel der trekkingsuitkomsten verwaarlozen. Zo zal men bij de voorbeelden in 1.3.1 met $N = 201$ en $N = 20$ gemiddeld ongeveer 80% van de getrokken getallen niet nodig hebben. Hierbij komt nog dat bij trekking zonder teruglegging getallen meerdere malen kunnen voorkomen, terwijl men ze maar eenmaal nodig heeft.

Bij het trekken zonder teruglegging en bij het aselecteren kan men voor $N \leqslant 1000$ in vele gevallen beter gebruik maken van *lotingstabellen* (Eng. random permutations), waarin de getallen van 1 t/m N *in aselecte volgorde* zijn opgenomen. In 'Normalisatiecommissie 70' (1951) staan deze tabellen voor $N = 10, 20, 50, 100$ en 200 en in 'Statistische Tabellen en Nomogrammen' voor $N = 100, 400, 600$ en 1000. In tabel Z hebben wij er enkele opgenomen voor $N = 10, 20, 50$ en 100.

Gebruik van lotingstabellen voor steekproeftrekken ($N \leqslant 1000$). Stel dat men uit een populatie van 90 elementen aselect een steekproef van 5 elementen wil trekken. Men zoekt dan in de lotingstabellen een tabel op voor de naasthogere waarde van N, d.i. hier dus een tabel voor $N = 100$. Hieruit leest men (te beginnen van boven en koloms-, regels- of diagonaalsgewijs gaande) 5 getallen af. De elementen waarvan de nummers met deze getallen corresponderen kiest men in de steekproef.

Neemt men bijv. in tabel Z voor $N = 100$ serie 2 en leest men links bovenaan beginnend kolomsgewijs af, dan vindt men de getallen 03,

77, (92), 43, 69 en 83 (daar de populatie slechts 90 elementen telt, die van 1 t/m 90 genummerd zijn, wordt 92 overgeslagen). De steekproef bestaat dus uit de elementen met de nummers 3, 43, 69, 77 en 83.

Gebruik van lotingstabellen voor aselecteren. Beschouw nogmaals het geval dat men een groep van 1 t/m 20 genummerde proefdieren aselect in drie groepen van resp. 6, 7 en 7 wil splitsen. Gebruikt men hiertoe in tabel Z voor $N = 20$ serie 1 en leest men regelsgewijs af, dan vindt men voor de nummers van de dieren in de drie groepen:

Groep 1:	6	16	9	7	19	3	
Groep 2:	18	12	11	1	20	8	13
Groep 3:	17	14	5	15	10	2	4

Wil men 16 dieren aselect in twee groepen van 8 verdelen, dan vindt men voor de nummers van de dieren in deze groepen met tabel Z, $N = 20$, serie 5 (de getallen 17 tot en met 20 verwaarlozend en kolomsgewijs aflezend):

Groep 1:	6	10	7	4	13	16	12	14
Groep 2:	1	15	2	5	8	3	9	11

1.3.3. LAAGSGEWIJS ASELECTE STEEKPROEVEN

Men kan gebruik maken van een laagsgewijs aselecte steekproef, indien de onderzoekpopulatie op grond van beschikbare informatie in homogene deelpopulaties kan worden gesplitst, die onderling heterogeen zijn. Betreft het onderzoek bijv. een grootheid, dan heeft de laagsgewijze trekking zin wanneer de deelpopulaties zo zijn samen te stellen dat men weet of vermoedt dat *binnen* elke deelpopulatie een variatie optreedt die klein is met betrekking tot de variatie die *tussen* (de gemiddelden van) de deelpopulaties aanwezig is.

Een laagsgewijze steekproef wordt samengesteld door uit iedere deelpopulatie (laag) aselect één steekproef te trekken (zie t.a.v. de omvang van deze deelsteekproeven 23.3). Dit steekproeftype wordt veel gebruikt bij opinie-, budget- en marktonderzoeken, waarbij men met de gehele bevolking van een land (of groot deel daarvan) te maken heeft, waarvan men de opbouw naar geslacht, leeftijd, grootteklasse van gemeente van inwoning, e.d. (ongeveer) kent. Het gaat dan gewoonlijk om schattingen betreffende eigenschappen van de bevolking en deze bezitten bij een laagsgewijs aselecte steekproef een grotere nauwkeurigheid dan bij een volledig aselecte steekproef van dezelfde omvang, omdat het voorkomen van éénzijdige steekproeven, waarin toevallig één of meer deelpopulaties slecht (of niet) zijn vertegenwoordigd, wordt uitgesloten.

Op medisch gebied komen laagsgewijs aselecte steekproeven minder frequent voor, onder meer omdat men gewoonlijk niet over de vereiste

gegevens betreffende de populatiesamenstelling beschikt, of omdat men juist geïnteresseerd is in de verschillen tussen de deelpopulaties en de daaruit getrokken steekproeven vergelijkt. Men kan echter wel vaak *laagsgewijs aselecteren*, d.w.z. bij het aselect verdelen van patiënten of proefdieren in groepen, die verschillende behandelingen zullen ondergaan, een laagsgewijze opbouw toepassen. Neem als voorbeeld dat men een groep van 102 patiënten in twee even grote subgroepen wil splitsen, die behandeling A, resp. B zullen krijgen en veronderstel dat zij naar geslacht en leeftijd de nevenstaande verdeling bezitten. Indien men mag verwachten dat zowel geslacht als leeftijd van invloed zullen zijn op de uitkomst van de behandelingen, heeft het zin de twee subgroepen laagsgewijs aselect samen te stellen. Men splitst daartoe met lotingsgetallen elk der zes delen van de groep aselect in twee gelijke onderdelen, waarvan men er één aan behandeling A en de tweede aan behandeling B toewijst (bij delen met een oneven aantal laat men (aselect) één patiënt uitvallen, of voegt men alsnog één patiënt toe). Bij het voorbeeld zou men dan laagsgewijs aselecte groepen krijgen van de nevenstaande samenstelling. Daar geen patiënten zijn toegevoegd dienen er in totaal 4 patiënten uit te vallen.

Leeftijd	M	V
20–39	12	7
40–59	31	20
60–79	17	15
Totaal	60	42

Laagsgewijs aselecte groepen

Leeftijd	*A*			*B*		
	M	V	Tot.	M	V	Tot.
20–39	6	3	9	6	3	9
40–59	15	10	25	15	10	25
60–79	8	7	15	8	7	15
Totaal	29	20	49	29	20	49

Het is duidelijk dat men het aantal (relevante) criteria, waarop de laagsgewijze opbouw wordt gebaseerd niet te groot moet kiezen om het optreden van te veel uitvallers te voorkomen.

1.3.4. GROEPSSTEEKPROEVEN

Deze steekproefmethode kan worden toegepast indien de populatie bestaat uit een groot aantal betrekkelijk kleine deelpopulaties (of groepen), die onderling niet te veel in omvang verschillen en waarbij de elementen in elke groep t.a.v. de te onderzoeken eigenschappen heterogeen zijn (een grote variabiliteit te zien geven). De groepssteekproef wordt dan samengesteld uit een aselecte steekproef van k groepen en alle elementen van een groep (of een aselect gekozen deel daarvan) worden onderzocht.

Deze methode is voordelig, wanneer de kosten voor het onderzoeken van één element laag, maar die voor het verkrijgen van dit element betrekkelijk hoog zijn, bijv. door een sterke geografische spreiding van

de tot de populatie behorende elementen. Veronderstel bijv. dat de onderzoekpopulatie bestaat uit alle Nederlandse kinderen, die op 1 oktober 1962 de eerste klasse van een lagere school bezoeken en dat het eigenlijke onderzoek per kind ca. 5 minuten vergt. Trekt men uit de populatie aselect een steekproef van bijv. 1000 kinderen, dan zullen deze door het gehele land verspreid zijn, zodat per kind een veelvoud van de onderzoekkosten voor reis- en verblijfkosten van de onderzoeker(s) moet worden uitgetrokken. Kiest men echter een aselecte steekproef van bijv. 100 eerste klassen dan behoeft men per klasse van gemiddeld ca. 35 kinderen slechts één reis te maken, waarna men in een ochtend of een middag al deze kinderen kan onderzoeken. Dit onderzoek van ca. 3500 kinderen zou wel eens minder tijdrovend en dus minder kostbaar kunnen zijn dan het eerstgenoemde en het zal wellicht meer informatie opleveren.[1]

1.3.5. ANDERE STEEKPROEFTECHNIEKEN

Soms kan men bij het steekproeftrekken geen gebruik maken van aselecte getallen. Het inventariseren en nummeren van de populatie kan op praktische bezwaren stuiten, bijv. door de grote omvang van een populatie (denk in dit verband aan een produkt, waarvan een fabriek dagelijks tienduizenden stuks vervaardigt). Ook kan het onmogelijk zijn, zoals bijv. bij een onderzoek naar de visstand in een groot meer, of de wildstand in een gebied van grote afmetingen.

In deze en soortgelijke gevallen dient men te zoeken naar een trekkingsmethode, waarvan redelijkerwijs mag worden aangenomen dat zij een aselecte steekproef zal opleveren. Wij geven hier enkele voorbeelden van.

Systematisch steekproeftrekken (Eng. systematic sampling).

Hierbij kiest men, bijv. uit de stroom van een bepaald produkt die door een machine wordt afgeleverd, of uit een lijst of kaartsysteem waarin alle elementen van de populatie stuk voor stuk zijn opgenomen, successievelijk elk i-de exemplaar in de steekproef. Men begint hierbij met een element, dat aselect uit de eerste i elementen wordt gekozen. Teneinde een steekproef van n elementen te verkrijgen uit een populatie van de omvang N, neme men voor i het gehele getal dat zo dicht mogelijk bij N/n ligt. Voor $N = 1000$ en $n = 100$, is $i = 10$; wordt uit de getallen van 1 t/m 10 aselect het getal 7 gekozen, dan bestaat de steekproef uit de elementen met de volgnummers 7, 17, 27, . . ., 997.

Deze methode levert een aselecte steekproef op, indien de elementen een volgorde bezitten, die onafhankelijk is van de eigenschap(pen) die men gaat onderzoeken. Zij kan echter tot een niet-aselecte (onzuivere) steekproef leiden, wanneer zich in de volgorde systematische fluctuaties voordoen of wanneer de elementen van een bepaalde soort de neiging vertonen in groepen voor te komen. Het is dan ook beter om binnen elke groep van i elementen het nummer van het te kiezen element aselect aan te wijzen, dus bijv. voor $i = 10$:

[1] De genoemde cijfers zijn vanzelfsprekend gefingeerd. Men dient echter te bedenken dat bij het onderzoek van één kind vrijwel evenveel tijd aan voorbereiding zal worden besteed als bij het onderzoeken van een gehele klas.

aselect
uit 1 t/m 10: het 6e exemplaar, d.i. exemplaar met volgnummer 6,
uit 11 t/m 20: het 3e exemplaar, d.i. exemplaar met volgnummer 13,
uit 21 t/m 30: het 9e exemplaar, d.i. exemplaar met volgnummer 29,
enz.

Aselecte trekking van deelgebieden (Eng. area sampling).

Dit is een soort 'cluster sampling' dat kan worden gebruikt voor het trekken van een steekproef uit populaties, die een grote geografische spreiding vertonen (en waarvan vaak de elementen niet individueel bekend of geregistreerd zijn), zoals bijv. de populatie bestaande uit alle planten van een bepaalde soort in een bepaald gebied, de populatie bestaande uit alle gezinnen in een land. Op grond van een kaart van het gebied, waarin de onderzoekpopulatie voorkomt, splitst men dit in deelgebieden, waaruit men aselect een steekproef trekt. De hierin opgenomen deelgebieden worden dan volledig onderzocht, of uit elk deelgebied wordt opnieuw een steekproef getrokken. De mogelijkheden (en moeilijkheden) van deze methode worden door YATES (1953) uitvoerig besproken.

1.3.6. ONZUIVERE STEEKPROEVEN

De ervaring heeft geleerd dat de mens een bijzonder slecht instrument is voor het maken van een aselecte keuze. Als men geen aselecte trekkingsmethode kan gebruiken dient het surrogaat in ieder geval vrij te zijn van iedere persoonlijke invloed bij de trekking, daar deze invloed tengevolge van (on)bewuste preferenties bij de persoon die de keuze verricht gewoonlijk tot *onzuiverheid* of *selectie* (Eng. bias) leidt. Om deze reden is het zg. *doelbewust steekproeftrekken* (Eng. purposive sampling) dan ook niet aan te bevelen. Deze methode van uitzoeken van wat typerend voor de populatie wordt geacht zal immers in sterke mate gevoelig zijn voor vooroordelen van de onderzoeker. Hierbij komt nog, dat deze in de regel elementen zal kiezen die dicht bij de 'norm' liggen, zodat een onjuiste indruk van de spreiding van de populatie wordt verkregen.

Uit YULE en KENDALL (1946), gegevens van YATES. Uit een proefveld met tarwe trok men in mei (voordat de aren gevormd waren) en in juni (na de vorming van de aren) resp. 58 en 56 steekproeven van 8 planten. Per steekproef werden 6 planten aselect en 2 planten, steeds door dezelfde onderzoeker, doelbewust (op het oog) gekozen. De 8 planten in elke steekproef werden naar grootte gerangschikt; hierbij konden de op het oog gekozen planten dus twee van de plaatsen 1, 2, . . ., 8 innemen. Als zij, evenals de overige zes planten, een aselecte keuze vormen, zullen zij bij beschouwing van alle steekproeven deze plaatsen met ongeveer gelijke frequentie moeten bezetten. Tabel 1.1 geeft de uitkomsten van dit onderzoek.

Tabel 1.1. Afmetingen van tarweplanten. Frequenties van de op het oog gekozen planten voor de rangnummers van 1 t/m 8

Datum	Rangnummers								Totaal
	1	2	3	4	5	6	7	8	
31 mei	9	7	11	8	11	18	21	31	116
28 juni	9	19	27	23	15	10	5	4	112

Deze uitkomsten tonen duidelijk aan dat van een aselecte keuze geen sprake is. In mei treedt een uitgesproken preferentie voor de grotere planten op. In juni, na de vorming van de aren, geeft de onderzoeker de voorkeur aan planten van ongeveer gemiddelde grootte en vermijdt hij vooral de grote planten.

Het gebruik van vrijwilligers, waarbij men in feite de tot een populatie behorende personen zelf laat uitmaken of zij al dan niet in de steekproef of in een groep met een bepaalde behandeling worden opgenomen, kan eveneens tot onzuivere steekproeven leiden.

De moeilijkheden waartoe deze *autoselectie* kan leiden blijken, wanneer men denkt aan een onderzoek naar het effect van vaccinatie tegen influenza in een bedrijf, op basis van vrijwilligheid. Welke personen melden zich voor vaccinatie? Zijn het voornamelijk degenen die vaak influenza-achtige ziekten hebben? Komen zij uit bepaalde afdelingen van het bedrijf (waar a priori de kans op dit soort ziekten, bijv. omdat men buiten werkt, reeds groter is)? Bezit de gevaccineerde groep dezelfde leeftijdsopbouw en dezelfde opbouw t.a.v. gezinsgrootte als de niet-gevaccineerden? Enz.

Een tweede voorbeeld betreft een onderzoek naar het effect van schoolmelkverstrekking op de groei van kinderen (van ca. 6 t/m 13 jaar) op een groot aantal lagere scholen in een grote stad. Hierbij maakten de ouders uit, of hun kinderen de schoolmelk — tegen betaling — zouden nuttigen. Na één schooljaar vergeleek men o.m. de gemiddelde lengte en het gemiddelde gewicht van de schoolmelk- en de niet-schoolmelkgebruikers.

Het is duidelijk dat men de (iets) hogere gemiddelden van de schoolmelkdrinkers, niet zonder meer aan het melkgebruik mag toeschrijven, daar de vrijwillige keuze van de ouders tot selectie kan hebben geleid. Het is immers bekend dat kinderen uit betere welstandsgroepen gemiddeld langer en zwaarder zijn dan kinderen uit ongunstiger welstanden. De ouders in eerstgenoemde groepen kunnen echter het geld voor de schoolmelk gemakkelijker missen dan die uit laatstgenoemde, zodat men rekening moet houden met de mogelijkheid dat de schoolmelkgroep 'gemiddeld' uit betere welstanden komt en daardoor gemiddeld langer en zwaarder is. Bij gezinnen met ongeveer dezelfde welstand kunnen ouders, die verstandig met hun inkomsten omspringen hun kinderen eerder gebruik laten maken van de schoolmelkverstrekking dan ouders, die hun geld onverstandig uitgeven. Overigens bestaat ook de mogelijkheid dat juist ouders die thuis voor een verantwoorde voeding (melk inbegrepen) zorgen, de schoolmelk *niet* nodig zullen achten. Maar de eerstgenoemde ouders zullen wellicht hun kinderen ook thuis reeds beter verzorgen (verstandige, goede voeding!). Ook kan het zijn, dat men voor jongere kinderen (die gemiddeld korter en lichter zijn) de schoolmelk meer (of juist minder?) nodig acht dan voor oudere. Het is duidelijk dat het achteraf vrijwel onmogelijk zal zijn dit complex van factoren, die selectief kunnen hebben gewerkt, te ontwarren.

Selectie kan ook optreden wanneer men de steekproef wel juist heeft samengesteld, maar niet aan alle elementen een waarneming kan verrichten, zodat *uitvallers* optreden. Wanneer een populatie uit personen of gezinnen bestaat kan de reden hiervan zijn, dat men ze niet alle kan bereiken of dat zij onvoldoende medewerken of zelfs medewerking weigeren. Heeft men bijv. te doen met een steekproef van huisvrouwen, die voor een enquête overdag worden bezocht, dan zal men

relatief te veel vrouwen zonder kinderen missen, omdat deze meer uithuizig zijn of werkzaamheden buitenshuis verrichten.

WITTS c.s. (1959) schrijven hierover het volgende: 'The most relentless investigators have usually failed to obtain information from between 5 and 10 per cent of the (sample) studied. It is tempting to assume that these lapses are only a random sub-group, differing in no material way from the subjects succesfully investigated. There is little or no evidence that this assumption is justifiable. In fact COCHRANE [1] found a significantly increasing proportion of cases with tuberculosis in successively seen one-thirds of a coalmine population, in which more pressure was brought to bear on the late comers than on the first to come to be X-rayed. The true prevalence in the entire population was 50 per cent greater than that in the first two-thirds. It would reasonably be argued that in general those who fail to co-operate are either disproportionally normal or disproportionally abnormal; it is unduly optimistic to assume that neither possibility outweighs the other.'

HILL [1] zegt dan ook: 'I would myself infinitely sooner have, say, a one in four sample of the population, of a size thereby which enabled me to pursue relentlessly, and complete the records for all or nearly all the persons in it, than have to interpret figures derived from survey of the 'whole' population from which finally a quarter was missing.' DOLL [1] meent dat een uitvalpercentage *ten hoogste* 5 procent mag bedragen en dat 'anything appreciably higher may materially bias the results.'

In de praktijk komen veel hogere uitvalpercentages regelmatig voor, zonder dat men de mogelijke consequenties daarvan t.a.v. selectie voldoende overweegt. Bij het onderzoek van VAN DER WIELEN [2] trekt men (bijv.) primair een steekproef van 1 op 5 uit de Nederlandse huisartsen (uit het Geneeskundig Jaarboekje 1953). Bij nadere informatie bleken 46 van de 766 getrokken artsen geen huisarts meer te zijn. De overblijvende 720 huisartsen werden volgens een lotingsprocedure genummerd en ook zoveel mogelijk in de door het lot aangewezen volgorde bezocht, teneinde hun medewerking aan het onderzoek te vragen en een datum voor een bezoek ter uitvoering van een enquête vast te stellen. In totaal werden op deze wijze voor het vragen van medewerking 372 huisartsen bezocht, waarvan er 64 vroeger of later medewerking weigerden (waaronder 29 omdat zij de enquête ongewenst achtten en 12 omdat zij 'geen begrip voor het onderzoek' bleken te bezitten). Verder werden 40 artsen niet in het onderzoek opgenomen omdat 'hun omstandigheden zodanig waren, dat medewerking niet kon worden gevraagd'. In het geheel zijn er dus niet minder dan 104 op de 372, d.i. 28% uitvallers. Over de mogelijkheid dat hierdoor de steekproef een vertekend beeld zou kunnen opleveren wordt echter niet gesproken. Uit andere ter beschikking staande gegevens blijkt echter dat de steekproef veel te weinig kleine praktijken (100–1000 ingeschrevenen) bevat en dat hiertegenover een te veel aan grote praktijken staat.

Een algemeen verbreide misvatting is, dat een 'willekeurige', een *'lukrake'* (Eng. haphazard) keuze van elementen een aselecte steekproef oplevert. Dit is vaak niet het geval, omdat allerlei niet-vermoede of qua invloed moeilijk waardeerbare factoren selectief kunnen hebben gewerkt.

[1] In 'The application of scientific method to industrial and service medicine', H.M.S.O., London, 1951.

2 De huisarts en de doeltreffendheid van zijn aandeel in de gezondheidszorg, Van Gorcum, Assen, 1960.

Trekt men bijv. uit een kooi met 25 ratten lukraak een steekproef van 5 ratten, dan bestaat de mogelijkheid dat zich hierin juist die dieren bevinden, welke in minder goede conditie zijn, daardoor trager reageren en zich gemakkelijker laten vangen dan hun levendiger soortgenoten. Bevat de kooi alleen dieren in prima conditie, dan zou men bij zo'n lukrake trekking wel eens overwegend de zwaardere dieren in de steekproef kunnen opnemen, omdat deze indolenter zijn dan de andere. Het is daarom noodzakelijk om ook (zo niet vooral) bij dierproeven een aselecte trekkingsmethode te gebruiken, zelfs als daar oppervlakkig gezien geen redenen voor schijnen te zijn. MAINLAND (1954) merkt dan ook op: 'To omit randomization because one cannot see clearly how bias could occur is like trusting that glassware in chemistry is clean because it does not look dirty.'

Een ander geval van lukrake trekking doet zich vaak voor bij onderzoeken, waarbij de populatie bestaat uit patiënten met een bepaalde aandoening. Zo'n populatie is zelfs binnen één land gewoonlijk niet volledig bekend en men verricht dan noodgedwongen het onderzoek bij een steekproef bestaande uit die patiënten, welke zich ergens betrekkelijk eenvoudig laten 'vangen'.

RÜMKE (1961 d) geeft een voorbeeld van een klinisch-farmacologisch onderzoek over de invloed van een bepaald geneesmiddel op de zuurgraad van het maagsap bij patiënten met een maagzweer. Men wil 'in het algemeen', dus zeg voor de populatie van deze patiënten in Nederland op een bepaald moment, weten welke invloed dit middel uitoefent. Omdat men deze populatie niet in haar geheel kan onderzoeken, beperkt men zich tot de patiënten die op een bepaald tijdstip op de interne afdeling van een ziekenhuis zijn opgenomen. Deze 'lukrake' steekproef zal echter de populatie op nogal onbevredigende wijze vertegenwoordigen, omdat zij bijv.:

1. Voornamelijk bestaat uit patiënten die afkomstig zijn uit de streek of stad waarin het ziekenhuis gevestigd is.

2. Mogelijk overwegend patiënten bevat uit bepaalde sociale lagen.

3. Geen patiënten bevat met een maagzweer die zo weinig klachten oplevert dat zij niet zijn opgenomen.

4. Geen patiënten bevat die niet wensen te worden opgenomen.

5. Geen patiënten bevat bij wie een indicatie tot operatie bestaat; deze vindt men op de chirurgische zaal of op de wachtlijst.

De conclusies die uit het onderzoek worden getrokken hebben dus betrekking op een deel van de populatie — en dit is, zoals gewoonlijk bij dit soort situaties, moeilijk nauwkeurig te definiëren. Het staat zonder meer zeker niet vast dat zij onveranderd voor de gehele populatie gelden.

Voor *proefdieren* is de situatie in wezen niet veel anders. MAINLAND c.s. (1959) merken hierover op: '... if we visualize the circumstances and practices in laboratory animal breeding farms, including the way in which animals are born in litters, and the way in which a breeder collects a batch of animals of a specified type, we see good reason to believe that no batch of animals of a particular species, strain, sex and weight range, can be accepted as a strictly random sample of all animals to which these labels could be applied, even in that particular farm.'

Reeds op deze plaats willen wij wijzen op het feit, dat de verschillende soorten selectie die bij opname in een ziekenhuis kunnen optreden, het vrijwel onmogelijk maken om uit ziekenhuispatiënten algemeen geldige conclusies te trekken omtrent relaties tussen verschillende ziekten.

Een voorbeeld, waarin de werkelijkheid nog sterk vereenvoudigd is, moge dit illustreren.

Een onderzoeker wil nagaan of de veronderstelling juist is dat een chronisch reuma (R) bij lijders aan hartaandoening B meer optreedt dan bij lijders aan hartaandoening A. Hij wil daarbij gebruik maken van het materiaal dat in een groot ziekenhuis beschikbaar is.

Wij veronderstellen dat zijn hypothese onjuist is en dat in de populatie chronisch reuma even vaak voorkomt bij A als bij B en wel met een frequentie van 1 op 4, zoals in nevenstaande tabel is aangegeven. Veronderstel nu verder dat bij de opname in het ziekenhuis, tengevolge van de verschillen in de indicaties die tot deze opname leiden, de volgende opnamepercentages gelden: voor A—60%, voor B—20%, voor R—50%. Men vindt dan de volgende aantallen opgenomen patiënten:

		A	B	
R	+	200	100	300
	—	600	300	900
		800	400	1200

A, R (—): 60% van 600 = 360,

A, R (+): wegens A 60% van 200 = 120; van de resterende 80 50% wegens R = 40, dus totaal 160,

B, R (—): 20% van 300 = 60,

B, R (+): wegens B 20% van 100 = 20; van de resterende 80 50% wegens R = 40, dus totaal 60.

Deze ziekenhuisfrequenties kunnen eveneens in een tabel worden samengevat:

		A	B		
R	+	160	60	220	31% van A lijdt tevens aan R,
	—	360	60	420	50% van B lijdt tevens aan R.
		520	120	640	

De onderzoeker zou op grond van het ziekenhuismateriaal dus ten onrechte de conclusie trekken, dat chronisch reuma meer voorkomt bij hartaandoening B dan bij hartaandoening A.[1]

Wanneer men de groep patiënten bestudeert die op een *bepaald tijdstip* in een ziekenhuis aanwezig is, zal deze t.o.v. de *opgenomen* patiënten te veel personen met langere opnameduur bevatten. Naarmate men langer in een ziekenhuis verblijft is nl. de kans om op een bepaalde moment aanwezig te zijn, zoals men eenvoudig kan inzien, groter. Dit verschijnsel kan tot een relevante selectie leiden, daar de opnameduur mede afhangt van de ernst van de ziekte (of van de in het ziekenhuis verrichte ingreep), de burgerlijke staat van de patiënten (personen die ongehuwd, gescheiden of weduwnaar (weduwe) zijn en dus niet op een eigen gezin kunnen terugvallen verblijven in het algemeen langer in het ziekenhuis), milieufactoren (personen uit ongunstige sociale of hygiënische milieu's blijven gemiddeld langer), e.d. SCHWARTZ en ANGUERA (1957) geven hier voorbeelden van.

[1] Op dit verschijnsel is het eerst gewezen door BERKSON (1946). Het staat dan ook in de literatuur bekend als 'BERKSON's fallacy'.

Uit het voorgaande blijkt dat het in de praktijk niet altijd eenvoudig is om een steekproeftechniek toe te passen, die aselect is of waarvan men verwachten mag dat zij tot een aselecte steekproef leidt. Men dient echter aan het steekproeftrekken bijzonder veel zorg te besteden, omdat onzuivere steekproeven tot misleidende conclusies kunnen leiden.

1.3.7. OPGAVEN

1.1. Trek met tabel Y tien steekproeven van 5 elementen (met teruglegging) uit de modelpopulatie in tabel X en noteer hieronder de uitkomsten (resp. het nummer van het getrokken element en zijn waarde):

Volgorde van trekking	Steekproeven																			
	1		2		3		4		5		6		7		8		9		10	
	Nr	x	Nr	x	Nr	x	Nr	x	Nr	x	Nr	x	Nr	x	Nr	x	Nr	x	Nr	x
1																				
2																				
3																				
4																				
5																				

1.2. Verdeel een groep elementen genummerd van 1 t/m 10 aselect in twee groepen van 5 elementen:
a. Met tabel Y (begin in kolom 5, regel 2 en ga kolomsgewijs naar beneden),
b. Met tabel Z.

1.3. Verdeel een groep elementen genummerd van 1 t/m 30 aselect in 3 groepen van 10 elementen:
a. Met tabel Y (begin in kolom 1, regel 20 en ga regelsgewijs naar rechts),
b. Met tabel Z.

1.4. Onderzoekingen op medisch gebied

1.4.1. INLEIDING

Het is nog niet zo lang geleden dat onderzoek op medisch gebied eigenlijk alleen werd verricht via waarnemingen bij individuele patiënten. Pas in de eerste decennia van deze eeuw werd door hygiënisten en tropenartsen en later door voedingsdeskundigen gebruik gemaakt van gegevens betreffende groepen personen. Dit *groepsonderzoek* betrof oorspronkelijk gehele populaties of althans grote delen daarvan, zodat een uitgebreid waarnemingsmateriaal moest worden geordend, beschreven en geanalyseerd. Men kon hierbij vrijwel steeds volstaan met gebruik te maken van de methoden die de beschrijvende statistiek daartoe ter beschikking stelde en deze gingen geleidelijk een grotere rol op het 'epidemiologische' gebied spelen.

Spoedig bleek dat het principe van het groepsonderzoek niet alleen voor het bestuderen van de frequentie, de verspreiding en de epidemiologie van ziekten kon worden gebruikt, maar ook voor het bestuderen

van de ziekte zelf, bijv. wat betreft haar beloop, de frequentie van het voorkomen van bepaalde symptomen, de reactie op medicamenten, enz. Voor de studie van klinische verschijnselen was men dus niet meer uitsluitend aangewezen op de uitkomsten van casuïstisch onderzoek bij individuele patiënten. Bij het 'klinische' groepsonderzoek dient men zich echter gewoonlijk te beperken tot groepen patiënten (of proefdieren), die slechts een — vaak relatief klein — deel van de totale te bestuderen populatie(s) vormen. Om op basis van zo'n deel, de steekproef, tot een uitspraak over de populatie te komen kan men gebruik gaan maken van de methoden der mathematische statistiek, waarvan de ontwikkeling juist in de eerste decennia van deze eeuw grote vooruitgang heeft gemaakt.

In figuur 1.1 zijn de verschillende soorten van groepsonderzoek in een schema opgenomen. In de eerste plaats is hier een onderscheid gemaakt tussen een *onderzoek* en een *analyse van reeds bestaande gegevens*.

Bij een *onderzoek* worden de waarnemingen door de onderzoeker zelf (gewoonlijk in samenwerking met een aantal medewerkers) verricht. Wij bespreken de verschillende vormen die een onderzoek kan aannemen in 1.4.2. Een onderzoek wordt voorafgegaan door de *planning*, dat zijn de werkzaamheden die moeten plaatsvinden om te komen tot het vaststellen van de proefopzet. De *proefopzet* bestaat uit de verzameling van voorschriften waarin de gang van zaken bij een onderzoek is vastgelegd. De verschillende overwegingen waarmee men hierbij rekening moet houden worden in 1.4.3. besproken.

Bij een *analyse van bestaande gegevens* gaat men uit van waarnemingsmateriaal dat reeds ter beschikking staat (bijv. sterfte- of ziekteverzuimcijfers, gegevens van patiënten die in een ziekenhuis opgenomen zijn geweest). Van een proefopzet is dus geen sprake, maar wel moet zorgvuldig worden nagegaan welke waarde het materiaal bezit en hoe het zal worden bewerkt en geanalyseerd. Wij bespreken dit in 1.5.

1.4.2. ONDERZOEKEN

In figuur 1.1 is reeds onderscheid gemaakt tussen experimenteel en observationeel onderzoek en deze twee onderzoekvormen worden nu achtereenvolgens besproken.

Bij een *experimenteel onderzoek* (Eng. experiment, trial) gaat men uit van een steekproef van elementen uit de te onderzoeken populatie. Deze elementen worden aan één (of aan een aantal opeenvolgende) behandelingen onderworpen, of zij worden in twee of meer groepen verdeeld die verschillende behandelingen ondergaan. Hierbij worden vóór, gedurende of na de behandeling(en) waarnemingen verricht. In dit verband verstaan wij onder *behandeling* bijvoorbeeld het toedienen van een geneesmiddel, het geven van een dieet, het verrichten

van een chirurgische ingreep, enz. Wij hebben het experimentele onderzoek onderverdeeld in:

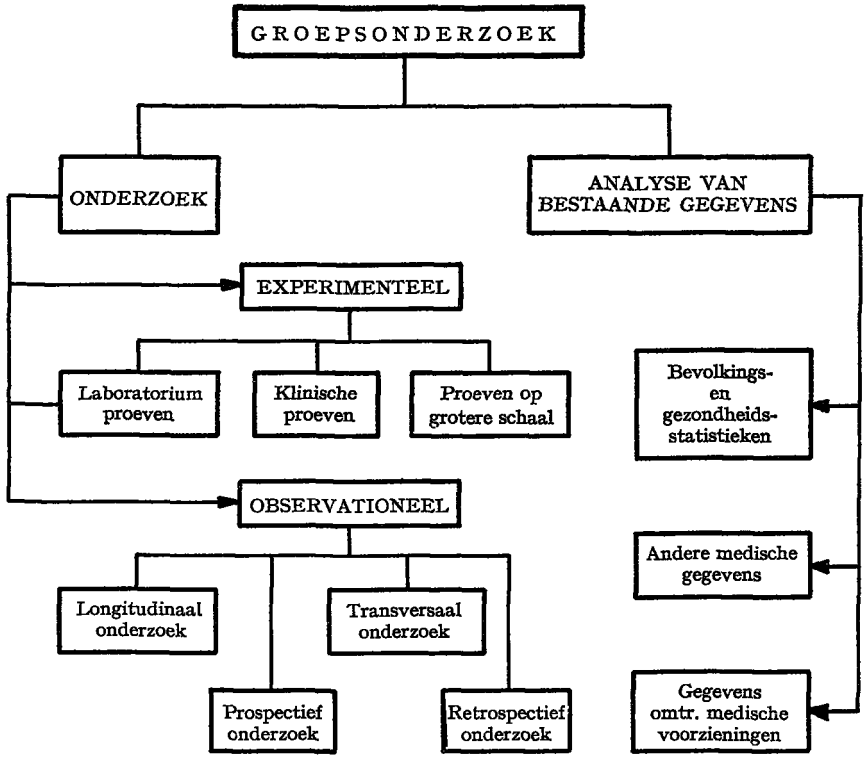
1. *Laboratoriumproeven* (Eng. laboratory experiments). Hieronder vallen de in het laboratorium uitgevoerde proeven, zoals: experimenteel farmacologisch onderzoek; experimenteel epidemiologisch onderzoek (waarbij men bij grote groepen proefdieren het beloop van een infectie bestudeert); onderzoek over de werkzaamheid of toxiciteit van geneesmiddelen via dierproeven. Ook proeven ter bepaling van de betrouwbaarheid (zuiverheid en nauwkeurigheid) van waarnemingsmethoden willen wij hieronder rekenen.

2. *Klinische proeven* (Eng. controlled clinical trials). Hieronder vallen wetenschappelijk verantwoorde therapietoetsingen, waarbij men door middel van een vergelijkend onderzoek informatie verzamelt over het effect van verschillende behandelingen. Voor een inleiding over deze methoden verwijzen wij naar HILL (1951), PANNEKOEK (1956), DE JONGH (1958) en GOSLINGS (1961). Zij worden uitvoerig, met vele voorbeelden en met een uitgebreide literatuuropgave behandeld in WITTS (1959) en speciaal in 'Controlled Clinical Trials' (1960).

3. *Proeven op grotere schaal* (Eng. controlled field trials). Dit zijn gewoonlijk onderzoekingen betreffende de profylactische werking van een bepaalde maatregel of behandeling. Voorbeelden zijn: de fluoridering van drinkwater ter voorkoming van tandcariës en vaccinatie tegen poliomyelitis [zie VAN STRIK (1958a)].

Bij het experimentele onderzoek kan soms worden volstaan met één proefgroep die wordt behandeld, bijv. bij aandoeningen die 100% dodelijk verlopen of bij ziekten die steeds een vast patroon blijken te volgen. Ook is er de mogelijkheid dat de *patiënt zijn eigen controle* is, waarbij iedere patiënt de verschillende behandelingen successievelijk ontvangt; deze methode kan echter alleen worden toegepast voor ziekteprocessen, die gedurende de onderzoekperiode vrijwel constant zijn en bij middelen die een, in korte tijd te beoordelen, snel effect hebben. Gewoonlijk zal men echter gebruik moeten maken van *controles*, d.i. van een groep patiënten of dieren die niet, of op andere wijze wordt behandeld. Bij therapeutische proeven met mensen zal men veelal een controlegroep moeten invoeren, die een *schijnbehandeling* ondergaat, doordat een placebo of een dummy wordt verstrekt. Verder zal men in dit geval zo mogelijk een *dubbel-blinde* proef moeten uitvoeren, d.i. een proef waarbij noch aan de patiënt, noch aan de waarnemers (verplegend personeel, behandelende arts) bekend is aan wie het nieuwe middel, dan wel het oude (als dit ter vergelijking dient), resp. de placebo of dummy wordt gegeven.

Figuur 1.1. Schematisch overzicht van de verschillende vormen van groepsonderzoek op medisch gebied.



Een *placebo* dient om de patiënt 'een genoegen te doen' en kan stoffen bevatten met een zeker nuttig effect, zoals vitaminen.

Een *dummy* is een farmacologisch geheel indifferent preparaat, dat meestal zoveel mogelijk dezelfde kleur, smaak, enz. bezit als het te onderzoeken middel. Gewoonlijk maakt men geen onderscheid tussen deze twee soorten schijnmiddelen en worden zij beide placebo genoemd. Onder een placebo of schijnmiddel verstaat men dan elke stof, die een invloed van niet farmacologische aard uitoefent op het beloop of de verschijnselen van een ziekte.

De Jongh heeft aan de betekenis van het placebo (1954) en de opzet van klinische proeven (1958) zeer lezenswaardige artikelen gewijd. Hierin zegt hij over het gebruik van controles en een dubbelblinde proefopzet o.a. het volgende (1958):

'De vraag of men altijd een controle-waarneming nodig heeft, moet in beginsel bevestigend beantwoord worden. Uitspraken a priori gelden immers niet

in de ervaringswetenschappen. De enige wijze om te weten te komen of een factor al of niet een bepaalde werking heeft is het experiment en altijd moet men aannemelijk maken, dat datgene wat tijdens de inwerking dezer factoren gebeurt niet ook zonder deze zou zijn ingetreden. Voor het laatste is echter een blanco waarneming nodig. De enige schijnbare uitzondering op deze regel is het geval, waarin reeds zoveel controle-waarnemingen zijn verricht met een zodanig ondubbelzinnige uitkomst, dat omtrent die uitkomst praktische zekerheid bestaat. Om een voorbeeld te noemen: men behoeft aan de uiteindelijke afloop van een leukaemie-geval niet te twijfelen, aangezien nog nooit iemand van deze ziekte is genezen.

In de meeste gevallen is het onmogelijk betrouwbare gegevens omtrent de therapeutische werking van een geneesmiddel in handen te krijgen zonder blanco waarnemingen te doen. Dit blijkt ten duidelijkste uit het onafgebroken geharrewar in de medische literatuur over talloze medicamenten. Zo weet men bv. nog steeds niet zeker, of stoffen als khelline, papaverine, polynitraten en purines invloed hebben op angina pectoris. En dit is maar een willekeurig voorbeeld. De Amerikaanse auteur O. B. Ross (J.A.M.A., 145 (1951), 72) laat door een analyse van recente literatuur zien, hoe vaak het gebeurt, dat conclusies de therapie betreffende ongeldig zijn door het ontbreken van controles. Hij verzamelde een 100-tal artikelen uit de 5 voornaamste Amerikaanse geneeskundige tijdschriften, die verschenen na 1 januari 1950. Zijn conclusie luidde nu, dat slechts in ruim een vierde deel van de onderzoekingen adequate controleproeven waren gedaan. Dit betekent dus, dat in de rest der gevallen de conclusie ongeldig is. Ongetwijfeld zijn er onder deze ongeldige tevens onjuiste conclusies die dan toch plegen voort te blijven bestaan in de medische literatuur. Deze krijgt dan wat ik de kookboekvorm noem. In stede van de neerslag te vormen van betrouwbare gegevens wordt daardoor de therapeutische literatuur tot een soort medische folklore:

Wil men geldige conclusies bereiken, dan moet men dus in het algemeen op therapeutisch gebied in iedere proefneming ... blanco waarnemingen opnemen. De betrokken zieken moeten dezelfde behandeling ondergaan als de patiënten die het medicament ontvangen, echter op dit medicament na. Men kan er evenwel niet mee volstaan (alle) controlepatiënten niets te geven, aangezien vast is komen te staan, dat het geven van een voor een geneesmiddel doorgaande stof als zodanig een therapeutisch effect kan hebben, het zg. placebo- of schijnmiddeleffect. Vele ziekelijke verschijnselen verdwijnen wanneer de patiënt ten onrechte gelooft, dat hij een werkzaam geneesmiddel ontvangt. Bovendien is echter gebleken, dat het invloed kan hebben op de uitslag van de proef wanneer de arts van te voren de uitkomst reeds meent te weten. Wanneer de arts welbewust aan sommige patiënten een geneesmiddel geeft en aan anderen een schijnmiddel, dan kan hij dat blijkbaar op zodanige wijze doen, dat het geneesmiddel beter werkt dan wanneer ook de behandelende geneesheer niet weet welk middel de patiënt krijgt, het echte of de 'dummy'. Bovendien moet men aannemen dat soms het resultaat anders wordt beoordeeld wanneer het geneesmiddel dan wanneer het schijnmiddel wordt gegeven. Het is onvermijdelijk, dat de behandelende arts eerder geneigd is verbetering van de toestand te constateren wanneer hij een geneesmiddel dan wanneer hij een placebo heeft gegeven.'

In het artikel van 1954 geeft De Jongh een groot aantal sprekende voorbeelden van het placebo-effect, waarvan wij een drietal aanhalen (voor de literatuurverwijzingen raadplege men het oorspronkelijke artikel):

'Een voorbeeld waarbij inderdaad later werd aangetoond dat het niet gebruiken van een placebogroep tot foutieve conclusies had geleid, is het geruchtmakende onderzoek van Brewster over de invloed van antihistaminica op het beloop van verkoudheid. In een met behoorlijke controles uitgevoerde proefreeks

van HOAGLAND bleef van dit effect niets over. Ook COWAN en DIEHL vonden geen verschil in de duur van verkoudheid wanneer deze met antihistaminica (al of niet samen met vitamine C) of met placebo's werd behandeld.' (Een later onderzoek van de Medical Research Council bevestigde deze resultaten.)

'Placebo's kunnen bij de mens niet alleen hetzelfde therapeutische effect hebben als 'echte' geneesmiddelen, zij kunnen ook dezelfde bijverschijnselen teweegbrengen. WOLF en PINSKY vergeleken, speciaal met dit doel, tolserol met een placebo. Het bleek dat beide middelen bij de patiënten even vaak hoofdpijn, sufheid en anorexie veroorzaakten. Drie patiënten kregen door het schijnmiddel zelfs ernstige verschijnselen.'

'Het placebo-effect kan niet alleen bij zieken, maar onder bepaalde omstandigheden ook bij gezonde personen worden waargenomen. De proefpersonen van JONGBLOED en VAN GOOR leverden betere werkprestaties wanneer zij tijdens rustpozen lucht in te ademen kregen met de mededeling dat het zuurstof was, dan wanneer hun zuurstof onder de naam van lucht werd toegediend.'

Het gebruik van controles kan soms in strijd komen met de medische ethiek, nl. wanneer op grond van theoretische overwegingen of een voorlopig onderzoek een nieuwe behandeling te prefereren lijkt boven de vroeger gebruikte. DRION (1953) zegt over dit ethische probleem het volgende: 'De onderzoeker zal als arts in een dergelijk geval zijn patiënten niet gaarne de door hem verwachte voordelen van de nieuwe methode onthouden. Hij moet dan echter wel beseffen, dat hij ook nimmer de waarde van de nieuwe methode zal kunnen bewijzen, zodat latere teleurstellingen niet zijn uitgesloten. Het is mogelijk, dat dus het heil van *talloos vele latere* patiënten wordt opgeofferd aan het heil van *enkele huidige* patiënten, die de arts, om hun het *veronderstelde* nut van de nieuwe methode niet te onthouden, niet in een controlegroep plaatst. Welke keuze in een dergelijk ethisch dilemma moet worden gedaan, zal de onderzoeker telkens opnieuw voor zijn eigen geweten moeten uitmaken.' [1]

Bij een *observationeel onderzoek* (Eng. survey) beperkt men zich tot het verrichten van waarnemingen aan de elementen in de te onderzoeken populatie(s) of in daaruit getrokken steekproeven. Wij zullen vier vormen van observationeel onderzoek onderscheiden en achtereenvolgens bespreken, nl. longitudinaal, prospectief, retrospectief en transversaal onderzoek [zij worden uitvoeriger en met vele voorbeelden behandeld in WITTS (1959)].[2]

Bij een *longitudinaal onderzoek* (Eng. follow-up study) gaat men uit van een steekproef uit een bepaalde populatie. Alle daarin opgenomen elementen worden in de tijd vervolgd, waarbij men — gewoonlijk periodiek — zekere waarnemingen verricht, bijv. om vast te stellen

[1] Zie ook GOSLINGS (1961) en RÜMKE en VAN EEDEN (1961).

[2] Vanzelfsprekend zijn er verschillende combinaties van deze onderzoeksoorten mogelijk die wij in dit globale overzicht buiten beschouwing laten.

welk verloop in de tijd een grootheid vertoont, of om na te gaan welke bijzonderheden (zoals recidieven, of complicaties na een operatie) in de tijd optreden.

Enkele voorbeelden van longitudinaal onderzoek: (1) periodieke metingen van bepaalde lichaamsmaten bij één groep kinderen, teneinde het algemene groeipatroon en afwijkingen daarvan te leren kennen, (2) halfjaarlijks onderzoek van een aantal personen met een chronische ziekte, om een beeld te krijgen van het 'normale' beloop (bijv. de frequentie van aanvallen bij lijders aan epilepsie).

Van groot belang is, dat het aantal uitvallers tijdens het onderzoek uiterst klein is (zie 1.3.6), dat de waarnemingen bij iedere persoon volledig worden verricht en dat de waarnemingsmethoden in de loop van de tijd niet veranderen. Bij een lange duur van het onderzoek is een zeker percentage uitvallers onvermijdelijk; men dient dan in ieder geval de oorzaken vast te stellen en deze nauwkeurig te registreren.

Bij het prospectieve en retrospectieve onderzoek houdt men zich bezig met *één kenmerk* van de personen van een populatie (een eigenschap, een gewoonte, een gebeurtenis, een reeks gebeurtenissen), waarvan men vermoedt of verwacht dat het (mede) het optreden van een *bepaalde ziekte* veroorzaakt. De term ziekte moet hier (en in het vervolg) ruim worden opgevat, omdat wij er niet alleen een ziekelijke aandoening of ziekteproces, maar ook aangeboren afwijkingen, later ontstane misvormingen, e.d. onder begrijpen.

Bij een *prospectief onderzoek* (Eng. prospective survey) gaat men — evenals bij een longitudinaal onderzoek — uit van een steekproef uit de te onderzoeken populatie. Men stelt van alle in deze steekproef opgenomen personen vast, welke het 'onder verdenking staande' kenmerk bezitten (de 'proefgroep') en welke niet (de 'controles'). Als het kenmerk gegradueerd kan worden (bijv. naar de mate waarin het aanwezig is) zal men meerdere proefgroepen vormen. Alle elementen worden dan in de tijd vervolgd teneinde na te gaan of (en soms wanneer) de ziekte optreedt en of dit optreden in de proefgroep(en) en controlegroep verschilt of overeenstemt. Bij deze vorm van onderzoek heeft men dus, evenals bij een experiment, één of meer proefgroepen en een controlegroep, maar de verschillende 'behandeling' van deze groepen is niet afkomstig van de onderzoeker, maar van de natuur of maatschappij.

Enkele voorbeelden van prospectief onderzoek: (1) Registratie van rookgewoonten, waarna wordt afgewacht of zich bij verschillende soorten rokers (met verschillende rookgewoonten) en bij niet-rokers een verschillende sterfte voordoet; wij noemen hier de onderzoeken van DOLL en HILL (1956) en HAMMOND en HORN (1958), (2) Registratie van bijzonderheden gedurende de zwangerschap, waarna later wordt vastgesteld welke congenitale afwijkingen zich bij de hieruit geboren kinderen voordoen; zie MCDONALD (1958).

Voor het slagen van een prospectief onderzoek moet aan drie voorwaarden worden voldaan. In de eerste plaats moet men er redelijk zeker van zijn, dat het onder verdenking staande kenmerk en de ziekte niet *beide* 'veroorzaakt' worden door dezelfde factor.

Als voorbeeld kan dienen een prospectief onderzoek bij kinderen, waarbij het onder verdenking staande kenmerk het al dan niet verrichten van een tonsillectomie is en waarbij men later nagaat, hoe frequent een neurose voorkomt bij de geopereerde en niet-geopereerde kinderen. Vindt men hierbij meer neurosen bij de geopereerden, dan mag men dit verschil o.i. niet onmiddellijk aan de operatie (i.c. aan psychische traumata door de operatieve ingreep) toeschrijven. Hier bestaat onder meer de mogelijkheid, dat het gezinsmilieu voor operatie én ziekte van belang is, doordat overdreven zorgzame, min of meer neurotische ouders eerder neurotische kinderen hebben, maar ook eerder een tonsillectomie zullen laten verrichten.

In de tweede plaats moet men van alle personen in het onderzoek de andere kenmerken, waarvan men weet of vermoedt dat zij het ontstaan van de ziekte (hoe dan ook) beïnvloeden (zg. *relevante* kenmerken), nauwkeurig registreren. Men kan dan bij de statistische analyse de invloed van deze factoren uitschakelen (of deze invloed mede in het onderzoek betrekken). Bijna altijd zijn hier geslacht en leeftijd van belang en gewoonlijk ook sociale klasse, beroep, woonplaats (en bij kinderen gezinsgrootte en rangnummer in het gezin).

Ook hier een eenvoudig gefingeerd voorbeeld ter illustratie. Veronderstel dat men bij een prospectief onderzoek voor een bepaalde populatie de nevenstaande uitkomsten vindt betreffende een kenmerk A en een ziekte Z. Onder de personen met kenmerk A treft men dus meer lijders aan de ziekte aan, dan onder de personen die kenmerk A niet bezitten.

		Kenmerk A +	Kenmerk A —	
Ziekte	+	100 (25%)	360 (20%)	460
	—	300	1440	1740
		400	1800	2200

Veronderstel nu verder dat men vermoedt, dat de ziekte meer zal voorkomen bij vrouwen dan bij mannen. Splitst men op grond hiervan de gegevens naar geslacht, dan zijn de volgende uitkomsten mogelijk:

		Mannen A(+)	Mannen A(—)	Vrouwen A(+)	Vrouwen A(—)	Totaal
Ziekte	+	10	90	90	270	460
	—	90	810	210	630	1740
Totaal		100	900	300	900	2200

Nu vindt men wel een duidelijk verschil in de ziektepercentages bij mannen en vrouwen (10% tegen 30%), maar 'binnen' elk geslacht is geen invloed van kenmerk A te vinden.

In de derde plaats moet men er bij een prospectief onderzoek (evenals bij een longitudinaal onderzoek) voor (kunnen) zorgen, dat men bij het vervolgen van de personen in de tijd (vrijwel) geen uitvallers krijgt.

Aan de prospectieve methode zijn twee belangrijke bezwaren verbonden:

(1) het kan lang duren voordat de bestudeerde ziekte zich manifesteert,

(2) bij betrekkelijk zeldzame ziekten moet men een groot aantal personen vervolgen om tot een 'proef'-groep van voldoend grote omvang te komen. Tengevolge hiervan is de methode vaak zeer kostbaar en/of praktisch onuitvoerbaar.[1]

Bij een *retrospectief onderzoek* (Eng. retrospective survey) gaat men uit van een groep personen, waarbij de te bestuderen ziekte reeds is opgetreden. Men onderzoekt in de eerste plaats hoe frequent (en/of in welke mate) het kenmerk, dat als (mede)oorzaak wordt verdacht, bij deze groep voorkomt. Ter vergelijking onderzoekt men het voorkomen van het kenmerk bij een controlegroep, welke in principe dient te bestaan uit personen uit *dezelfde populatie* die niet aan de ziekte lijden.

Hier stuiten wij op het fundamentele probleem bij elk retrospectief onderzoek: het omschrijven van de populatie waaruit de patiëntengroep afkomstig is. Gewoonlijk bestaat deze groep namelijk uit een verzameling patiënten (bijv. uit één of meer ziekenhuizen) waarvan men in eerste instantie slechts kan zeggen dat zij zeker geen aselecte steekproef vormt uit de populatie die men feitelijk zou willen onderzoeken (zoals de bevolking uit een land of streek daarvan). De aangelegenheid is echter van vitaal belang en verdient bijzondere aandacht van de onderzoeker, omdat kennis van de (selecte) populatie waaruit de patiëntengroep stamt zowel noodzakelijk is voor het kiezen van de juiste controles als voor het aangeven van de beperkingen, die aan de conclusies van het onderzoek verbonden zijn.

Bij het retrospectieve onderzoek maakt men vaak gebruik van de methode der *gepaarde controles* (Eng. matched controls). Hierbij wordt naast elke patiënt een controle gekozen, die er in een aantal relevant geachte kenmerken (bijv. geslacht, leeftijd(sklasse), beroep) mee overeenstemt.

Wanneer dit mogelijk blijkt vormt men meerdere controlegroepen; de daarin gesignaleerde overeenkomsten en/of verschillen in het voorkomen van allerlei kenmerken (waaronder het onder verdenking staande) kunnen, in relatie tot de wijze waarop zij zijn gekozen, aanwijzingen

[1] Soms kan men de duur beperken door het begin van het onderzoek naar een tijdstip in het verleden te verplaatsen ('prospectief onderzoek in retrospect').

verschaffen omtrent hun waarde voor het onderzoek. Een mogelijkheid tot controle binnen de patiëntengroep levert een onderzoek naar de 'internal consistency', het 'op zichzelf consistent zijn' van de uitkomsten binnen deze groep (zoals overeenstemming in de uitkomsten van overeenkomstige groepen patiënten van verschillende oorsprong, bijv. uit verschillende ziekenhuizen).

Ook hier moeten wij met enkele voorbeelden volstaan. Bij het eerste, retrospectieve onderzoek van DOLL en HILL (1952) betreffende de relatie tussen roken en longkanker, ging men uit van een groep patiënten met longkanker in verschillende ziekenhuizen, waarbij de rookgewoonten in het verleden werden nagegaan. Er werden twee controlegroepen gevormd, (a) bestaande uit patiënten met andere ziekten, exclusief kanker, van hetzelfde geslacht, binnen dezelfde leeftijdsgroep (van 5 jaar) en in hetzelfde ziekenhuis in ongeveer dezelfde periode, (b) bestaande uit patiënten met één van drie andere vormen van carcinoom, die op dezelfde wijze werden gerapporteerd als de longkankerpatiënten. Het onderzoek leverde uiteindelijk nog een derde controlegroep op, bestaande uit patiënten waarbij de diagnose longcarcinoom ten onrechte bleek te zijn gesteld. FABER c.s. (1957) gingen uit van patiënten met myeloïde en andere vormen van leukemie en onderzochten hun vroegere expositie aan ioniserende straling. Als controles kozen zij patiënten met chronisch lymfatische leukemie, er van uitgaande dat deze vorm van leukemie zich onafhankelijk van de blootstelling aan ioniserende straling ontwikkelt.

Een tweede bezwaar van de retrospectieve methode is dat men pas achteraf gegevens over de patiënten en controles verzamelt. In het bijzonder wanneer het hierbij gaat om gewoonten (die in het verleden niet constant geweest behoeven te zijn) of over een reeks gebeurtenissen (bijv. ondergane bestralingen en de aard daarvan) trekt men hierbij een 'wissel op het verleden', waarbij meestal een beroep moet worden gedaan op het geheugen van de in het onderzoek betrokken personen, waarvan men zich steeds moet afvragen in hoeverre hij kan worden gehonoreerd. Hierbij komt nog dat de bekendheid met de diagnose of de ziektegeschiedenis van de patiënt de anamnese kan vertekenen, terwijl ook onzuiverheid kan optreden doordat de patiënt anders op het onderzoek reageert dan de controle.

EKKER (1961) geeft hiervan het volgende voorbeeld: 'In twee recente studies over het verband tussen leukemie op de kinderleeftijd en bestraling gedurende de zwangerschap (bestudeerde) STEWART het vraagstuk door retrospectief onderzoek en vond een duidelijk verband; COURT BROWN kon in een prospectieve studie geen invloed aantonen, en sprak het vermoeden uit dat moeders van een kind met een kwaadaardige aandoening zich gemakkelijker — misschien zelfs te gemakkelijk — bijzonderheden uit de voorgeschiedenis voor de geest halen, dan moeders van gezonde kinderen.'

In dit verband was bij het retrospectieve onderzoek van DOLL en HILL de derde controlegroep (patiënten waarbij oorspronkelijk ten onrechte de diagnose longcarcinoom gesteld was) belangrijk. Een onzuiverheid zou hier uitkomsten in de richting van de kankergroep doen verwachten; zij bleken echter geheel op de andere controlegroepen te gelijken.

Wij hebben eerst de nadelen van de retrospectieve methode opgesomd, omdat het voordeel t.o.v. de prospectieve methode evident is: men krijgt sneller resultaten met een minder omvangrijk onderzoek. Verder is in sommige gevallen wegens de zeldzaamheid van de ziekte een prospectief onderzoek praktisch onmogelijk. Ook als men omtrent de factoren die mogelijk tot de ziekte leiden nog niet over duidelijke aanwijzingen beschikt, kan men deze trachten op te sporen door middel van een retrospectief onderzoek. De bevindingen kunnen dan later door een prospectief onderzoek en — zo mogelijk — door een experiment worden getoetst.

Een ander verschil tussen prospectief en retrospectief onderzoek bestaat in de schattingen omtrent de populatie die zij in principe kunnen verschaffen. Dit kan worden geadstrueerd met nevenstaande tabel waarin de uitkomsten van een onderzoek zijn opgenomen.

	Kenmerk A		
	+	—	
Ziek	a	c	$a+c$
Niet ziek	b	d	$b+d$
	$a+b$	$c+d$	n

Betreffen zij een prospectief onderzoek dan kan men de volgende prospectieve schattingen maken:

$a/(a+b)$ = relatieve morbiditeit bij personen met kenmerk A,

$c/(c+d)$ = relatieve morbiditeit bij personen zonder kenmerk A. De verhouding van deze fracties is een maat voor de 'relatieve kans' op het krijgen van de ziekte.

Betreffen de uitkomsten een retrospectief onderzoek, dan verkrijgt men de retrospectieve schattingen: $a/(a+c)$ = fractie van de zieken met kenmerk A, en $b/(b+d)$ = fractie van de controles met kenmerk A. In de regel levert een prospectief onderzoek geen retrospectieve, een retrospectief onderzoek geen prospectieve schatting op. Soms bestaat deze mogelijkheid wel: zie CORNFIELD en HAENSZEL (1960) en EKKER (1961).

De laatste observationele methode is het *transversale onderzoek* (Eng. prevalence survey, cross-sectional survey). Hierbij is het de bedoeling, van één of meer populaties op *een bepaald moment* de frequentieverdeling van een kenmerk (zoals ziekte), een categorisch systeem (bijv. dat van de bloedgroepen) of variabele (bijv. bloeddruk) vast te stellen of om na te gaan of tussen bepaalde kenmerken en/of variabelen een samenhang aanwezig is. Dit onderzoek wordt in de regel uitgevoerd door middel van steekproeven uit de te bestuderen populaties, waarbij vaak gebruik wordt gemaakt van laagsgewijze trekking of 'clustersampling'.

Een transversaal onderzoek wordt vaak uitgevoerd wanneer men een beeld wil krijgen omtrent het voorkomen van een ziekte of afwijking bij de gehele bevolking van een bepaald gebied, die de (selecte) groep van geregistreerde patiënten niet kan verstrekken ('correcting the clinical picture'). De onderzoekmethode wordt ook gebruikt voor het opsporen van verschillen tussen (bijv.) bepaalde bevolkingsgroepen of beroepsgroepen, die aanwijzingen kunnen opleveren omtrent de etiologie van een ziekte. Zie voor vele voorbeelden FLETCHER en OLDHAM in WITTS (1959).

In 1.2 hebben wij twee basisvoorwaarden genoemd waaraan voldaan moet zijn, indien men de statistiek als 'bewijsmiddel' wil gebruiken. Gaan wij na in hoeverre bij de voorgaande onderzoekvormen aan deze voorwaarden kan worden voldaan, dan blijkt het volgende:

A. *Experimenteel onderzoek.* In principe kan men de statistiek als bewijsmiddel gebruiken, daar de onderzoekpopulatie a priori kan worden gedefinieerd, de steekproef hieruit aselect kan worden getrokken en de verschillende proef- en of controlegroepen aselect kunnen worden samengesteld.

B. *Observationeel onderzoek.* Bij het longitudinaal en transversaal onderzoek kan de statistiek als bewijsmiddel worden gebruikt, indien het mogelijk is de onderzoekpopulatie nauwkeurig te omschrijven, de steekproef aselect te trekken en het percentage uitvallers te beperken. Gaat men uit van een min of meer 'lukraak' samengestelde groep patiënten of personen en/of treedt een te hoog percentage uitvallers op, dan zal de statistiek slechts als detectiemiddel kunnen dienen.

Een prospectief onderzoek bezit gewoonlijk geen 'volledige' bewijskracht omdat het risico bestaat dat niet-vermoede factoren in het spel zijn die zowel met het onder verdenking staande kenmerk als met de ziekte samenhangen (zoals erfelijkheid, voeding, sociale omstandigheden), of omdat mogelijk een systematisch verschil tussen de proef- en controlegroep aanwezig is, dat niet of onvoldoende kan worden uitgeschakeld. Overigens geldt ook hierbij als primaire voorwaarde dat slechts een gering percentage uitvallers toelaatbaar is.

Bij een retrospectief onderzoek is de bewijskracht afhankelijk van de juiste samenstelling van de controlegroep en de mate waarin men er in slaagt (achteraf!) betrouwbare, objectieve gegevens te verzamelen. Wij hebben reeds gezien dat vooral het eerstgenoemde niet eenvoudig is.

De waarde van de uitspraken die men verkrijgt wanneer niet geheel aan de basisvoorwaarden kan worden voldaan hangt af van de aard en de ernst van de onvolmaaktheden van het materiaal en van de mate waarin men erin slaagt deze te onderkennen en hun eventuele invloed te evalueren. De onderzoeker zal hiertoe gespeend moeten zijn van elk vooroordeel, bijzonder kritisch te werk moeten gaan en niet moeten schromen om waar nodig reserves te maken, ook als deze ertoe leiden dat minder besliste uitspraken worden gedaan.

MAINLAND c.s. (1959) geven de raad bij het beoordelen van een observationeel onderzoek de volgende vraag te stellen: 'If the investigation is a survey, what is the simple outline of an imaginary experiment which, if it were possible, we would perform instead of the survey?' Hij zegt verder: 'We should use this imaginary experiment in answering subsequent questions (on the survey). This is the best way of bringing to light the difficulties and uncertainties inherent in survey research, and of showing what we should try to achieve as far as possible, in our survey methods.'

In het voorgaande is reeds opgemerkt dat de bij een onderzoek verkregen uitspraken in principe uitsluitend geldig zijn voor de populatie(s), waarop het onderzoek betrekking heeft. Wanneer men bijv. door een goed opgezette klinische proef heeft vastgesteld, dat

een geneesmiddel een duidelijk therapeutisch effect bezit bij Nederlandse mannen van 20 t/m 50 jaar, dan mag men strikt genomen deze conclusie niet betrekken op Nederlandse vrouwen van deze leeftijden, Nederlandse jongens, of Engelse mannen van 20 t/m 50 jaar. Er kunnen *medische* overwegingen zijn, die de geldigheid van de uitspraak ook voor één of meer van deze populaties *plausibel* maken, maar een (statistisch) bewijs voor deze geldigheid is slechts te leveren als men het onderzoek ook voor deze populaties uitvoert.

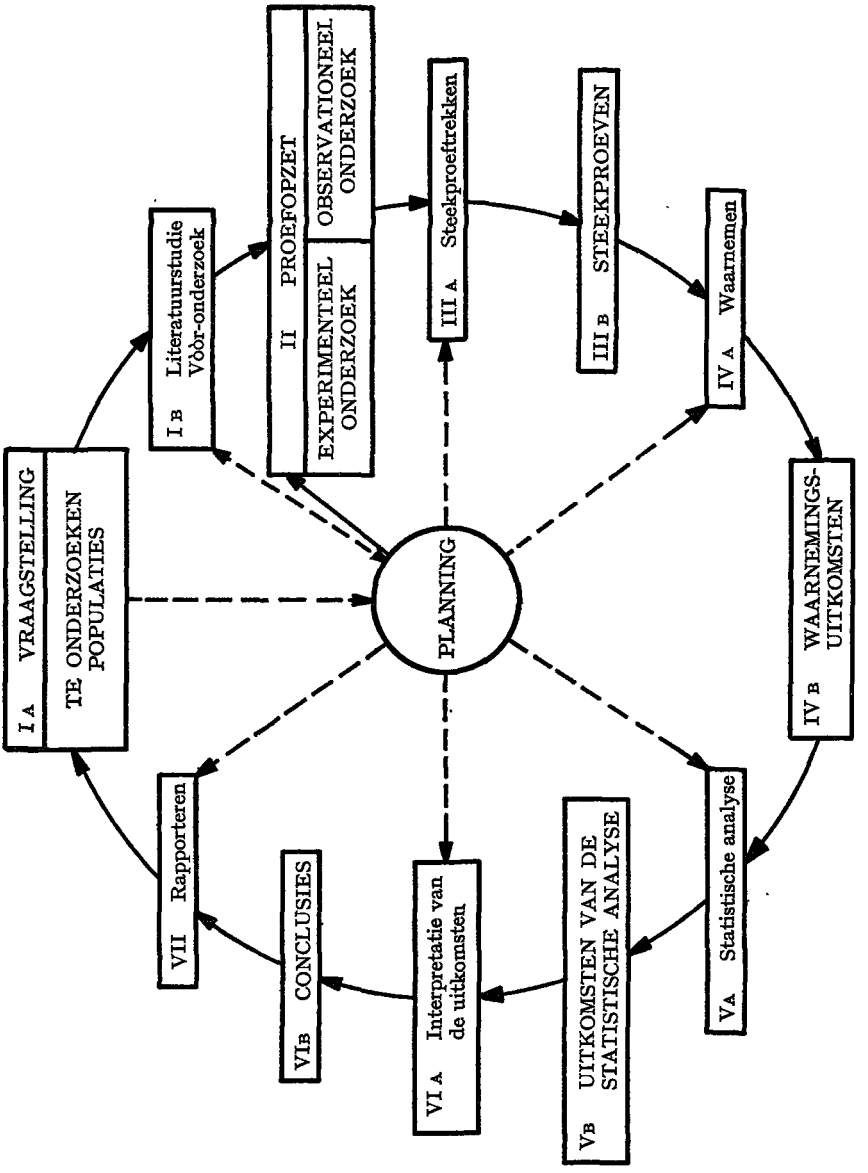
Een overeenkomstig probleem levert de extrapolatie in de tijd. Een conclusie die alleen geldt voor het tijdstip of tijdperk van onderzoek is vanzelfsprekend van weinig waarde. Daar echter bij uit mensen of dieren bestaande populaties in de tijd vele min of meer geleidelijke veranderingen optreden, zullen de conclusies van een nu verricht onderzoek over 10 jaar wellicht niet meer gelden. Zo kan een insekticide nu (en over een jaar) bij een bepaalde insektensoort een uitgesproken effect hebben, maar na verloop van langere tijd kan zich een immuniteit voor de stof ontwikkelen, waardoor dit effect vrijwel geheel verloren gaat. Ook bij extrapolatie in de tijd geldt, dat men vaak op grond van niet-statistische overwegingen aannemelijk kan maken, dat zich in de tijd geen relevante wijzigingen hebben voltrokken, maar het bewijs hiervan kan alleen door het uitvoeren van een nieuw onderzoek worden geleverd. Mede om deze reden is het herhalen van reeds eerder door anderen verrichten onderzoekingen vaak verre van zinloos.

1.4.3. ALGEMENE GANG VAN ZAKEN BIJ EEN ONDERZOEK

Wij besluiten deze paragraaf met een aantal opmerkingen over de algemene *gang van zaken* bij een onderzoek, die gebaseerd zijn op het schema in figuur 1.2, waarin verschillende opeenvolgende fasen zijn onderscheiden.

Fase I. Een onderzoek dient aan te vangen met een duidelijke formulering van de *vraagstelling* die men wil beantwoorden en het definieren van de *populatie(s)* waarop men deze vraagstelling wil betrekken. Het is noodzakelijk dat men hierbij de literatuur die betrekking heeft op het onderzoekprobleem grondig bestudeert. Verder kan het voorkomen dat zich detailproblemen voordoen, die eerst — door middel van een voor-onderzoek (Eng. pilot study) — moeten worden opgelost.

Door het literatuuronderzoek kan men zich van belangrijke informatie voorzien en veel moeilijkheden voorkomen. Het dient echter uitermate *kritisch* te zijn. Men late zich niet (mis)leiden door een met autoriteit uitgesproken conclusie of door een rookgordijn van tabellen, figuren en mathematisch statistische tecbnieken, die soms slechts camoufleren dat van een adequate proefopzet geen sprake is. Steeds zal men de gevolgde proefopzet grondig onder de loupe moeten nemen. Vertoont deze onvolmaaktheden, waarop de onderzoeker niet ingaat, dan zal men zelf moeten trachten uit te maken, welke invloed deze op de onderzoekuitkomsten kunnen hebben.



Figuur 1.2. De gang van zaken bij een onderzoek.

Het voor-onderzoek is *vrijwel altijd* noodzakelijk om de uitvoerbaarheid van de proefopzet aan de praktijk te toetsen (bijv. t.a.v. het tijdschema). Verder zal men er goed aan doen de meettechnieken (bijv. ter bepaling van het gehalte van stoffen in het bloedserum) en waarderingsschalen (bijv. die ter bepaling van het klinisch stadium waarin zich een patiënt bevindt, of die voor het aangeven van de mate van optredende verbeteringen in de toestand van patiënten) tevoren op hun betrouwbaarheid te onderzoeken. Ook verricht men soms waarnemingen in de tijd, zonder dat onderzocht is of wellicht seizoensinvloeden tot een systematisch verloop kunnen leiden. Wij komen hier bij fase IV op terug.

Fase II. Bij het vaststellen van de *proefopzet* dient men in de eerste plaats, op grond van de vraagstelling en de praktische mogelijkheden die ter beschikking staan, uit te maken welke vorm van onderzoek zal worden gekozen en of aan de daarbij noodzakelijke voorwaarden kan worden voldaan. Wanneer de onderzoekvorm vaststaat, dient men alle voorschriften over de gang van zaken bij het verrichten van het onderzoek nauwkeurig vast te leggen. De overwegingen die daarbij een rol spelen komen bij het bespreken van de volgende fasen aan de orde.

Men dient hierbij te bedenken dat tijdens het onderzoek in deze voorschriften geen veranderingen moeten worden aangebracht. Deze kunnen nl. tot een selectie leiden, waarvan het effect achteraf gewoonlijk niet meer te constateren is en die de uitkomsten kan vertekenen. Bij klinisch onderzoek is het soms wenselijk (of zelfs noodzakelijk) dat men bepaalde patiënten uitsluit, bijv. wegens de aanwezigheid van complicaties of te hoge leeftijd. Beslist men hiertoe a priori, dan wordt deze beperking in de voorschriften opgenomen. Zij kan dan uniform worden doorgevoerd en men weet, dat de uitkomsten van het onderzoek voor deze categorie patiënten niet van kracht (behoeven te) zijn. Gaat men tijdens het onderzoek incidenteel tot deze uitsluiting over, dan bestaat de mogelijkheid dat zij soms wel, soms niet worden geëxcludeerd en dat hierbij niet-uniforme criteria worden gehanteerd. Ook moet men tussentijdse veranderingen in een meettechniek vermijden: het invoeren van een nieuwe — en mogelijk op zichzelf betere — methode kan tot uitkomsten leiden, die systematisch verschillen van die, verkregen met de oude methode.

In deze fase kan men — zo nodig — ook het best overleg plegen met een statisticus. Deze kan kennisnemen van de vorderingen die de planning heeft gemaakt en de voorlopige proefopzet onbevooroordeeld en kritisch doornemen, in het bijzonder t.a.v. de statistische aspecten. Het behoeft nauwelijks betoog dat het voor een goed samenspel nodig is dat de onderzoeker vertrouwd is met de terminologie en gedachtengang van de statistiek en dat de statisticus zich met het terrein van onderzoek bekend maakt.

Fase III. Bij het *steekproeftrekken* (waaronder wij hier ook het samenstellen van proef- en of controlegroepen bij een experimenteel onderzoek begrijpen) dient men de grootst mogelijke aandacht te geven aan de aselectiviteit van de te gebruiken methode. Wij verwijzen hier naar 1.3.

Fase IV. Ten aanzien van het *waarnemen* dient men zich af te vragen:
1. *Wat* men gaat waarnemen, 2. *Welke* methoden daarbij worden gebruikt, 3. *Wanneer* de waarnemingen zullen worden verricht, 4. *Wie* de waarnemingen gaan verrichten, 5. *Hoe* zij zullen worden geregistreerd.

Ad. 1. Men moet voorkomen dat zoveel verschillende gegevens worden waargenomen en geregistreerd, dat men door de bomen het bos niet meer ziet; ook hier toont zich in de beperking de meester. Zo'n 'te veel' kan ook leiden tot de veel voorkomende, maar ernstige fout, dat vele waarnemingen niet of niet verantwoord worden uitgevoerd. Men baseert nl. herhaaldelijk conclusies op onvolledige gegevens, zonder dat men zich realiseert hoe misleidend deze kunnen zijn.

Onderstaande tabel bevat de bloedingstijd (in sec.) bij 5 personen, bepaald op verschillende tijdstippen na de toediening van 2 mg adrenoxyl. Door omstandigheden ontbreken de waarnemingen op de plaatsen, waar een streepje staat. Per tijdstip is het rekenkundig gemiddelde van de beschikbare waarnemingen berekend en deze gemiddelden zijn eveneens in de tabel opgenomen:

Patiënt	Tijdsverloop in uren na adrenoxyl					
	¼	½	1	3	6	9
A	120	—	100	85	90	110
B	80	75	—	—	70	75
C	—	140	120	95	100	115
D	100	95	95	80	90	100
E	40	35	30	—	—	40
Gem.	85	86	86	87	87	88

Men concludeert uit deze resultaten, dat de toediening van adrenoxyl geen invloed uitoefent op de bloedingstijd (de gemiddelden vertonen immers slechts geringe verschillen). Deze conclusie behoeft echter niet juist te zijn, daar de volledige uitkomsten heel goed het volgende, van patiënt tot patiënt overeenstemmende, beeld zouden kunnen vertonen:

Patiënt	Tijdsverloop in uren na adrenoxyl					
	¼	½	1	3	6	9
A	120	105	100	85	90	110
B	80	75	65	60	70	75
C	160	140	120	95	100	115
D	100	95	95	80	90	100
E	40	35	30	20	35	40
Gem.	100	90	82	68	77	88

Het voorbeeld lijkt misschien triviaal, maar deze fout blijkt in werkelijkheid regelmatig voor te komen.

Ad. 2. Men dient van elke waarnemingstechniek vast te stellen hoe betrouwbaar zij is.

Bij *metingen* en *tellingen* betekent dit dat men moet weten of de methode *zuiver* is, d.w.z. geen systematische fout bezit, resp. de grootte van een aanwezige systematische fout moet kennen en moet weten hoe *nauwkeurig* de

methode is. Wanneer binnen een onderzoek een bepaalde waarnemingsmethode steeds door één waarnemer wordt gehanteerd, dan kan men uit de uitkomsten wel conclusies trekken betreffende verschillen tussen de proef- en of controlegroepen, omdat daarbij de eventueel aanwezige systematische fout wegvalt. Men dient echter voorzichtig te zijn met een vergelijking t.o.v. de uitkomsten van een ander onderzoek, waaraan ook een systematische fout verbonden kan zijn. Wordt binnen een onderzoek een bepaalde waarnemingsmethode door verschillende waarnemers gehanteerd, dan dient men — zelfs als zij dezelfde apparatuur onder dezelfde omstandigheden gebruiken — op duidelijke systematische verschillen tussen de waarnemers verdacht te zijn. Zo bleek bijv. tussen hemoglobinebepalingen met dezelfde Sicca-hemometer van hetzelfde bloedmonster, onmiddellijk na, maar onafhankelijk van elkaar verricht, tussen *geoefende* analisten een systematische fout van 0,9 tot 1,2 g% (!) voor te komen. Bezitten twee groepen patiënten gemiddeld ongeveer hetzelfde hemoglobinegehalte, maar worden (toevallig — maar het toeval speelt ons deze parten) de bloedmonsters van de ene groep overwegend door de ene, dus die van de andere groep overwegend door de andere analist onderzocht, dan vindt men achteraf een systematisch verschil *tussen de groepen*.

Tabel 1.2. Hemoglobinegehalten, A = Sicca-hemometer, B = colorimetrisch-spectrofotometrisch
Standaardmonster: 16,3 g%

Uitkomsten laboratoria	A	B
12,3		1
13,1—13,5	1	
13,6—14,0	5	5
14,1—14,5	6	7
14,6—15,0	5	22
15,1—15,5	11	28
15,6—16,0	6	30
16,1—16,5		23*
16,6—17,0	1	3
Totaal	35	119
Mediaan	15,1	15,5

* waarvan 5 gelijk aan 16,3

Onderzoekingen van het Rijks Instituut voor de Volksgezondheid hebben aangetoond, hoe groot bij allerlei bepalingen de (overwegend) systematische verschillen tussen laboratoria kunnen zijn [zie bijv. Holtz (1960)]. Wij ontlenen hieraan één voorbeeld, dat nogmaals hemoglobinebepalingen betreft. Bij één der onderzoeken werd aan een groot aantal klinisch-chemische laboratoria een standaard-bloedmonster verstrekt met een hemoglobinegehalte van 16,3 g%. De uitkomsten van 35 laboratoria die gebruik maakten van een Sicca-hemometer en die van 119 laboratoria die colorimetrisch/spectrofotometrisch werkten zijn in tabel 1.2 samengevat en behoeven nauwelijks commentaar (merk echter op, dat het merendeel te laag is).

Wanneer men met waarnemingen van kwalitatieve kenmerken, dus bijv. met een dichotomie (+ of —, kenmerk wel of niet aanwezig) of met een waarderingsschaal te doen heeft, kan bij het waarnemen een subjectief element naar

voren treden, dat extra de aandacht verdient. Het is dringend geboden te onderzoeken of elke waarnemer, die aan het onderzoek deelneemt, de schaal steeds op dezelfde wijze en identiek aan andere waarnemers hanteert. Dit kan alleen als *van te voren* vastgestelde beoordelingscriteria worden gebruikt.

Uit verschillende onderzoeken is gebleken welk een grote subjectieve factor in vele, gewoonlijk als objectief gedachte criteria aanwezig is. In 'Controlled clinical trials' (1960) geeft FLETCHER hiervan een sprekend voorbeeld. Door 8 longspecialisten (A t/m H) werd (onafhankelijk van elkaar) bij 20 patiënten met chronische dyspnoe vastgesteld, hoe vaak zij verschillende op emphyseem wijzende criteria aanwezig achtten te zijn, alsmede bij hoeveel van deze patiënten door hen tenslotte de diagnose emphyseem werd gesteld. Enkele van de uitkomsten van deze proef zijn in figuur 1.3 weergegeven. De (soms) enorme verschillen (bij hyperresonantie gaande van 0–17) zijn uit deze figuur af te lezen.

In het jaarverslag 1959 van de Schoolartsendienst 'Noord-West-Brabant' schrijft men: 'Het is betreurenswaardig te moeten erkennen, dat een groot gedeelte van de door ons verzamelde gegevens van afwijkingen, geconstateerd tijdens het periodiek systematisch onderzoek, niet betrouwbaar zijn. Bij vergelijking van de gegevens van de beide schoolartsen onderling blijkt, dat de subjectieve interpretatie van grote invloed is geweest bij de notering van de afwijkingen, zodat een objectieve beoordeling van de cijfers niet mogelijk is. De geconstateerde verschillen . . . zijn dermate groot, dat deze niet zijn toe te schrijven aan werkelijk bestaande verschillen tussen de twee onderzochte bevolkingsgroepen. Enkele voorbeelden (in onderstaande tabel) mogen dit toelichten.'

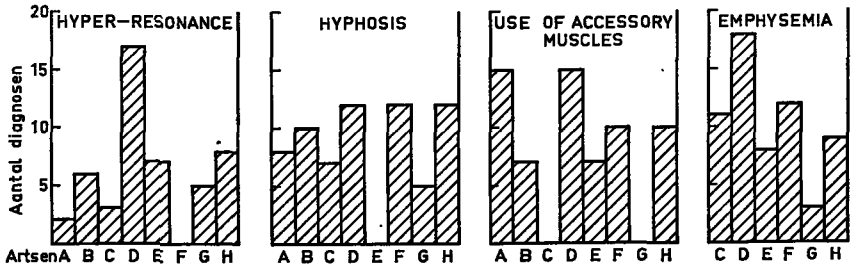
	Schoolarts	
	I	II
matige voedingstoestand	3 %	33 %
slechte voedingstoestand	0,1%	1,1%
hypertrofische tonsillen	5 %	10 %
goed gebit	53 %	37 %
struma +	75 %	53 %
hartafwijkingen	0,5%	0,1%

Wij moeten ons tot deze voorbeelden beperken, maar wijzen de lezer op het belangwekkende onderzoek van YERUSHALMY c.s. (1950) betreffende de subjectieve factor bij de beoordeling van longafwijkingen op kleinbeeld röntgenfoto's.

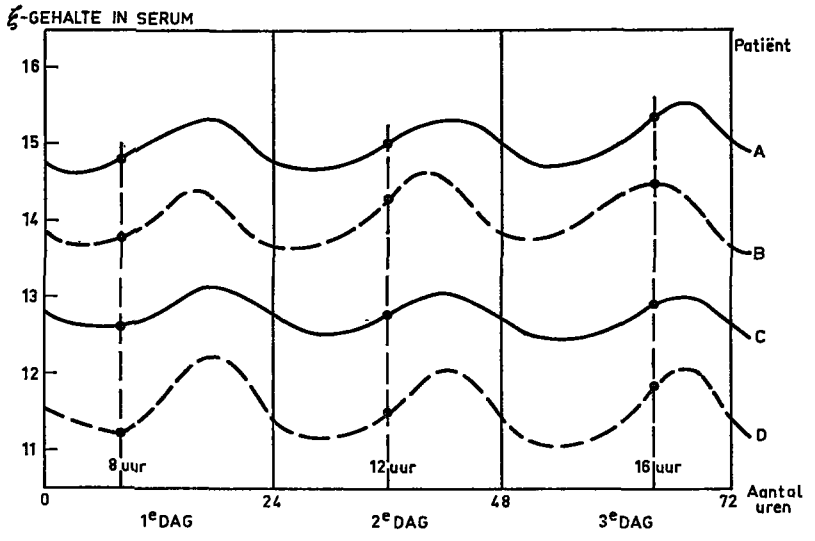
Ad. 3. Bij het vaststellen van het *tijdstip van waarneming* moet men steeds bedacht zijn op het feit dat vele verschijnselen een — min of meer systematisch — verloop in de tijd bezitten (dagritme, seizoensverloop) en dat ook het beloop van een ziekte reeds zonder behandeling niet constant zal zijn.

Figuur 1.4 toont het beeld van een grootheid die, gemeten op drie opeenvolgende data bij vier patiënten, een vrij regelmatig dagritme vertoont (het 'dal' ligt steeds tussen 6 en 8, de 'top' tussen 16 en 19 uur). Verder zien wij — en dit is regel — dat de variatie 'binnen elke patiënt (de fluctuatie om zijn gemiddelde niveau) veel kleiner is dan de variatie 'tussen' de patiënten (de variatie in de verschillende niveau's).

Meet men nu alle patiënten elke dag op hetzelfde, maar van dag tot dag op een verschillend tijdstip, dan verkrijgt men een foutief beeld van de verschillen



Figuur 1.3. De aanwezigheid van enkele op emphyseem wijzende criteria en emphyseem bij 20 patiënten met chronisch dyspnoe. Onafhankelijke beoordeling door 8 (resp. 6) longspecialisten. (Voor 'HYPHOSIS' leze men 'KYPHOSIS'). Ontleend aan FLETCHER in 'Controlled Clinical Trials' (1960).



Figuur 1.4. Verloop in de tijd van een gehalte met een duidelijk dagritme. Gefingeerde gegevens van vier patiënten, A t/m D.

tussen de dagen. Zou men bijv. alle patiënten op de eerste, tweede en derde dag resp. omstreeks 8 uur, 12 uur en 18 uur meten, dan concludeert men ten onrechte tot een systematisch stijgend verloop in de tijd (zie figuur 1.4). Vergelijkt men twee groepen patiënten R en S, die gemiddeld hetzelfde niveau hebben en ongeveer hetzelfde dagritme te zien geven, en meet men groep R om

8 uur en groep S om 12 uur van een bepaalde dag, dan is — door het dagritme — groep S gemiddeld hoger dan groep R. Ligt echter groep S steeds systematisch lager dan groep R, dan wordt dit bij deze wijze van meten min of meer genivelleerd.

Over de fluctuaties bij ziekten zeggen MAINLAND c.s. (1959) onder meer: 'In most diseases the sufferer's condition fluctuates in severity, and he is more likely to seek medical aid when his symptoms are more severe than when they are less severe. Therefore, at any particular time after a group of sufferers has come into a physician's hands he should expect to find some of them in better condition than when they first saw him, even if he has done nothing for them. Although very obvious, this phenomenon has been taken as evidence of benefit from treatment, especially in diseases ... that are characterized by remissions.' Over de volgorde van waarnemingen bij klinische proeven, waarbij twee behandelingen A en B successievelijk aan dezelfde patiënt worden gegeven zegt hij: '... we must not use the sequence AB (or BA) in all patients. Not only is there risk of carry-over effects ... but the trend of the disease, up or down, might entirely mislead us in evaluating the AB difference. We have *confounded*[1] (confused or mixed) time and treatment ...'.

Ad. 5. Bij een onderzoek op grote schaal is het noodzakelijk de waarnemingen zo te registreren, dat hun bewerking efficiënt kan plaatsvinden. Ligt het in de bedoeling gebruik te maken van mechanische bewerking door middel van ponskaarten (zie 2.1.5), dan verdient het aanbeveling het registratieformulier zo in te richten, dat het tevens als basis voor het verponsen kan dienen en zo veel mogelijk van te voren vast te stellen, hoe men de uitkomsten zal coderen.

Enquêtes. Als men een enquête uitvoert en dus de gegevens door een (schriftelijke of mondelinge) ondervraging van de proefpersonen verkrijgt, zijn er enkele punten waarop men in het bijzonder moet letten. Wij citeren DE WOLFF:

'Iedere vraag moet beknopt en vooral duidelijk zijn. Beginners zijn maar al te vaak geneigd te denken, dat een vraag, die voor hen volkomen duidelijk is, ook voor niet-ingewijden begrijpelijk zal zijn en, wat even belangrijk is, slechts één interpretatie zal toelaten. Het stellen van dergelijke vragen is echter lastig en vereist veel ervaring.

De vragen moeten liefst zo gesteld worden, dat er slechts een beperkt aantal antwoorden op gegeven kan worden, bv. ja of neen, de keuze uit een drie- of viertal alternatieven, een getal, enz ... Als men de vragen zo stelt, dat er met een omstandig verhaal op geantwoord kan worden, ondervindt men bij de rubricering veel ... moeite en dikwijls zal blijken dat de indeling in dat geval niet zonder willekeur kan geschieden. Indien een dergelijke vraagstelling slechts bereikt kan worden door één of meer vragen te splitsen en dus door vergroting van het totale aantal vragen, dan nog moet aan deze methode de voorkeur gegeven worden, al is het verder ook een gulden regel op het gebied der enquêtes, dat het totale aantal vragen zo klein mogelijk zij.

Het grootste probleem bij een enquête is echter de vragen zo te stellen dat ze zonder vooringenomenheid beantwoord kunnen worden en verder dat ze de ondervraagde niet een bepaald antwoord suggereren.'

Wij dienen met deze opmerkingen te volstaan en verwijzen voor een zeer volledige beschouwing, waarin ook alle mogelijke fouten bij enquêtes worden besproken, naar DEMING (1950) en STEPHAN en MC CARTHY (1958).

Fase V. Een belangrijk deel van de *statistische methoden* die voor de bewerking en analyse van de bij een onderzoek verkregen waarne-

[1] Ned. *verstrengeld.*

mingen ter beschikking staan wordt in dit boek behandeld of ingeleid. Reeds bij de proefopzet dient men de toekomstige statistische analyse in het oog te houden en in vele gevallen kan men zelfs de proefopzet op een specifieke statistische techniek instellen.

Dit laatste geldt in het bijzonder voor experimenteel onderzoek. Zo kan men bij het klinisch groepsonderzoek soms gebruik maken van de sequente analyse, waarbij verschillende proefopzetten volledig zijn uitgewerkt (zie 6.4.3). Bij laboratoriumproeven kan men vaak steunen op de zg. statistische proefschema's (zie hoofdstuk 24). Ook voor de biologische ijking staan verschillende 'standaard'-opzetten ter beschikking (zie hoofdstuk 25).

Fase VI. De *interpretatie* van de uitkomsten van de statistische analyse vereist inzicht in de betekenis en de draagwijdte van de via statistische methoden verkregen uitspraken. Men dient hierbij steeds de betrekkelijkheid van de statistiek als bewijsmiddel in het oog te houden en moet, zoals HAMAKER (1958) opmerkt '. . . nooit uit het oog verliezen dat de statistische bewijsvoering slechts één schakel is in de reeks van argumenten die samen uiteindelijk tot een juiste conclusie voeren.'

Fase VII. In het *rapport* (publikatie, dissertatie) dat de sluitsteen vormt van het onderzoek behoren (*a*) de proefopzet in zijn verschillende geledingen, (*b*) de onderzoekuitkomsten, (*c*) de analyse daarvan en (*d*) de eventuele tekortkomingen van het onderzoek en hun consequenties, zodanig te worden gepresenteerd dat de lezer in staat is de verschillende uitspraken te verifiëren en ze op hun merites te beoordelen.

Deze regels worden herhaaldelijk overschreden en vele publikaties bevátten dan ook te weinig informatie, die op onbevredigende wijze wordt gepresenteerd. Wij geven hiervan enkele sprekende voorbeelden [1].

(1) Uit N.T.v.G. 98 (1954), 2761–2774.

Wij ontlenen aan dit artikel de volgende gegevens betreffende het resultaat van de behandeling van 272 vrouwelijke epilepsie-patiënten. Op blz. 2765 staat: 'Het resultaat der behandeling werd uitgedrukt in de procentsgewijze vermindering van het aantal aanvallen. Men dient zich wel te realiseren, dat men zichzelf hierdoor als therapeut te kort doet, want veel patiënten gingen psychosomatisch aanzienlijk verbeterd naar huis, ofschoon zij nog evenveel aanvallen hadden. Ook werd vaak niet naar vrijheid van aanvallen gestreefd, omdat nogal wat lij-

[1] Deze en verschillende later volgende voorbeelden zijn ontleend aan de Nederlandse medische literatuur van de afgelopen 10 jaar. De bedoeling hiervan is de lezer te confronteren met praktijkproblemen, die vanzelfsprekend interessanter en overtuigender zijn dan gefingeerde. Voorzover deze opgaven en voorbeelden betrekking hebben op onvolledigheden of onjuistheden bij de planning of bij de presentatie, analyse of interpretatie van de uitkomsten van een onderzoek bezit deze procedure het nadeel, dat hiermee indirect kritiek wordt geleverd op bepaalde onderzoeken. Wij zijn ons van dit nadeel bewust, maar kunnen geen andere weg kiezen als wij aan concrete voorbeelden willen demonstreren dat deze tekortkomingen in de praktijk inderdaad voorkomen.

ders aan epilepsie veel beter zijn mèt, dan zonder aanvallen. Bovendien komt bij de beschrijving van het resultaat der behandeling niet tot uitdrukking hoeveel aanvallen de patiënten hadden. Want het maakt toch wel verschil of men het aantal gevallen tot de helft vermindert bij iemand die 2 aanvallen per jaar heeft of 5 per dag. Het is echter schier onmogelijk de duur van de vrijheid van aanvallen, het aantal aanvallen, enz. in een resultatencontrole statistisch te verwerken. Hiervoor zou men bovendien over een veel groter aantal patiënten moeten beschikken.' Op blz. 2772–3: '*Resultaat van de behandeling*. Van 272 patiënten waren er bij ontslag 234 verbeterd, waarvan 216 voor 50 pCt of meer. 139 (51 pCt) werden geheel aanvalsvrij (75 pCt 48, 50 pCt 29, 25 pCt 18 en 0 pCt 38). Van 206 patiënten was bekend hoe zij het een half tot drie jaar na het ontslag maakten. Van haar waren er 151 nog steeds beter dan voor de opneming en 139 van haar waren voor 50 pCt of meer verbeterd. 59 van 206 patiënten (29 pCt) waren aanvalsvrij (75 pCt 42, 50 pCt 38, 25 pCt 12, 0 pCt 55). Het bleek dat van de 108 patiënten die vrij van aanvallen uit de kliniek waren ontslagen, de helft (55) later weer aanvallen had gehad. Vijf patiënten waren na ontslag overleden.

Van 44 patiënten uit de groep 'petit mal' waren er 38 bij ontslag verbeterd, waarvan 34 voor 50 pCt of meer. 27 (61 pCt) waren vrij van aanvalletjes (75 pCt 3, 50 pCt 4, 25 pCt 4, 0 pCt 6). Van 32 patiënten was bekend hoe zij het een half tot 3 jaar na het ontslag maakten. 25 van haar waren beter dan voor de opneming en 23 waren voor 50 pCt of meer verbeterd. 16 (50 pCt) waren vrij van aanvallen (75 pCt 5, 50 pCt 2, 25 pCt 2, 0 pCt 7). Het bleek dat er van 18 patiënten, die zonder aanvallen waren ontslagen, 4 (22 pCt) weer aanvallen hadden gekregen.

Van 193 patiënten uit de groep van 'grand mal' waren er bij ontslag 165 verbeterd, waarvan 150 voor 50 pCt of meer. 91 (47 pCt) werden vrij van aanvallen (75 pCt 37, 50 pCt 22, 25 pCt 14, 0 pCt 28). Van 149 patiënten was bekend hoe zij het een half tot 3 jaar na het ontslag maakten. 105 van haar waren beter dan voor de opneming en 96 waren voor 50 pCt of meer verbeterd. 33 (23 pCt) waren vrij van aanvallen (75 pCt 30, 50 pCt 33, 25 pCt 9, 0 pCt 44). Van 72 patiënten die bij ontslag geen aanvallen meer hadden, kregen er 41 (57 pCt) deze later weer terug.

Uit de vergelijking van deze 2 groepen blijkt dat de prognose wat de frequentie der aanvallen betreft, bij lijders aan 'petit mal' 2 maal zo gunstig is als die bij lijders aan 'grand mal'. De kans, dat lijders aan 'grand mal' weer aanvallen krijgen na een vermeend therapeutisch succes, was 2,5 maal zo groot als die van lijders aan 'petit mal'.'

[Toelichting: Totaal (272) = Petit mal (44) + Grand mal (193) + Patiënten met één van de andere 7 typen aanvallen (35)].

In de eerste plaats is het duidelijk dat deze resultaten (waarin twee soorten percentages naast elkaar worden gebruikt) in tabelvorm moeten worden gepresenteerd. Hierbij dient men de situatie bij ontslag en bij de latere controle *per patiënte* op te geven (en te vermelden waarom sommige patiënten bij de controle zijn uitgevallen; tevens verdient het aanbeveling op te geven hoe lang na het ontslag elke patiënte is gecontroleerd — waarom?). De navolgende tabel (met gefingeerde gegevens) laat zien, dat bijv. zeer opmerkelijke veranderingen per patiënte kunnen optreden, zonder dat deze uit de los van elkaar staande totaalverdelingen per tijdstip blijken.

Ook hier zijn enkele belangrijke vragen te stellen, waarop het artikel geen antwoord geeft, bijv.: Hoe is de procentsgewijze vermindering der aanvallen bepaald? Op welke periode(n) zijn de aanvalsfrequenties bij opname, ontslag en controle gebaseerd? Hoe zijn deze frequenties bij opname en controle verkregen (van de patiënten, van een arts?). Welke fluctuaties vertonen dergelijke aanvalsfrequenties (zonder behandeling) in de tijd? Enz.

Patiënte	Proc. gewijze verbetering t.o.v. opname		Opmerkingen
	Bij ontslag	Bij controle	
1	50%	0%	
2	100%	50%	Alle patiënten met bij ontslag 50%
3	25%	25%	of meer verbetering zijn bij controle
4	0%	25%	verslechterd: nrs. 1, 2, 5, 7 en 9.
5	100%	0%	
6	0%	100%	Van de 5 patiënten met 0 of 25%
7	50%	25%	verbetering bij ontslag is er één on-
8	25%	75%	veranderd (nr. 3). De overige vier
9	75%	50%	zijn bij controle verbeterd: nrs. 4, 6,
10	25%	100%	8 en 10.
Totaalverdeling, zowel bij ontslag als controle: 100%–2, 75%–1, 50%–2, 25%–3, 0%–2			

(2) Uit: N.T.v.G. 98 (1954), 1296–1298. In een artikel betreffende het proconvertinegehalte in het bloed van vroeggeborenen wordt o.m. gezegd: 'Het is in de couveuse-afdeling . . . de gewoonte aan vroeggeboren kinderen gedurende de eerste drie levensdagen 10 mg vitamine K per dag toe te dienen, indien het geboortegewicht meer dan 2000 gram bedraagt, terwijl kinderen met een lager geboortegewicht dezelfde dosis gedurende vijf dagen krijgen. Wij achtten ons a priori niet gerechtigd bij ons onderzoek bij een deel dezer kinderen de toediening van vitamine K geheel weg te laten, gelijk dit bij voldragen kinderen wel is geschied. Daarom hebben wij de vroeggeborenen in twee groepen verdeeld: *groep A*, waarvan de kinderen de eerste levensdag een intramusculaire injectie van 10 mg vitamine K kregen . . . en vervolgens alleen de tweede en de derde levensdag telkens per os 10 mg vitamine K . . . en *groep B*, waarvan de kinderen ook na de eerste drie levensdagen dagelijks 10 mg vitamine K per os kregen.' Na een beschrijving van de onderzoekmethode wordt vermeld: Het onderzoek omvatte 39 vroeggeborenen met geboortegewichten tussen de 1500 en 2500 gram. In de eerste levensweek werden drie bepalingen in duplo gedaan, daarna werd het gehalte eenmaal per week bepaald.

Uitkomsten. Tabel I en II geven de gemiddelde rekenkundige waarden voor het proconvertinegehalte, die bij het opgegeven aantal vroeggeborenen op de verschillende levensdagen werden gevonden, en tevens de laagste en hoogste voor dit gehalte gevonden waarden. Fig. 1 (onze figuur 1.5) geeft een overzicht van de uitkomsten van alle gevallen.'

Tabel I. Proconvertinegehalte bij vroeggeborenen, die alleen gedurende de eerste drie levensdagen vitamine K toegediend kregen

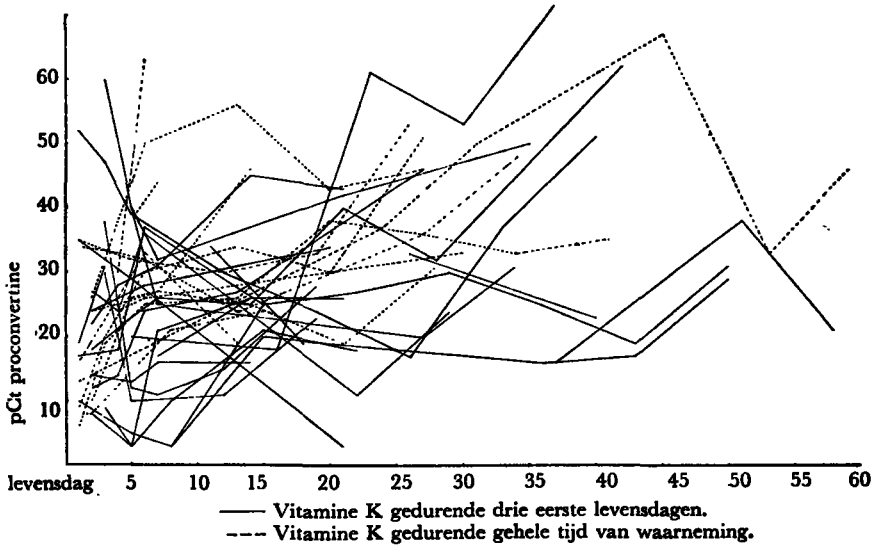
	Levensdag				Levensweek							
	4e dag	5e	6e	7e	1e	2e	3e	4e	5e	6e	7e	8e
Gem. waarde %	21	18½	31	22½	24½	21	26	30	37½	36	18	30
Laagste en hoogste waarde	14–28	3–40	24–37	11–32	3–60	3–45	3–43	11–61	24–53	16–72	17–19	21–38
Aant. onderzochte kinderen	4	13	4	8	48	14	15	10	6	8	2	4

Tabel II. Proconvertinegehalte bij vroeggeborenen, die gedurende het gehele onderzoek dagelijks vitamine K toegediend kregen

	Levensdag				Levensweek							
	4e dag	5e	6e	7e	1e	2e	3e	4e	5e	6e	7e	8e
Gem. waarde %	30	32	37	33	26	31	31	41½	41	35	67	33
Laagste en hoogste waarde	24–38	26–38	25–63	27–44	6–50	19–56	19–48	31–51	33–50			
Aant. onderzochte kinderen	5	3	7	3	27	9	6	7	4	1	1	1

De conclusies, die uit deze gegevens worden getrokken luiden o.m.:

'Deze tabellen en de figuur wijzen op een sterke spreiding in het proconvertinegehalte van het plasma van vroeggeborenen op verschillende levensdagen. . . . Bij de groep vroeggeborenen, die gedurende de gehele periode van waarneming dagelijks vitamine K kreeg, is het proconvertinegehalte in het algemeen een weinig hoger dan bij de vroeggeborenen, die alleen gedurende de eerste drie levensdagen vitamine K kregen.'



Figuur 1.5.

Ook hier volstaan wij met enkele opmerkingen: (a) Er is niet vermeld hoe de groepen A en B zijn samengesteld. Waarschijnlijk niet aselect (groep A bevat ongeveer tweemaal zoveel kinderen als groep B); hoe dan wel? Wat zijn de geboortegewichten in de twee groepen? (b) Is de (zeer grote fluctuatie van de proconvertinegehalten *per kind* wellicht ten dele een gevolg van een geringe nauwkeurigheid van de bepalingsmethode? De duplo's kunnen hierover informatie verschaffen. Vertoont het proconvertinegehalte een dagritme? (c) Als men uit de (onvolledige) tabellen en de (onduidelijke) grafiek de basisgegevens zo goed mogelijk reconstrueert (!) blijkt dat *niet* van ieder kind in de eerste levensweek drie bepalingen ter beschikking staan en dat daarna niet ieder kind éénmaal per week werd onderzocht. Er zijn kinderen met 0, 1 of 2 bepa-

lingen in de eerste levensweek. Er zijn ook kinderen waarbij in de eerste levensweek wel, later geen bepalingen zijn verricht (en omgekeerd). Nemen wij als voorbeeld *enkele* kinderen uit groep A waarvan de gegevens ongeveer konden worden gereconstrueerd, dan ziet men het volgende beeld:

Kind	Eerste levensweek							Latere levensweken							
	1	2	3	4	5	6	7	2	3	4	5	6	7	8	9
A	19	—	31	—	—	—	—	—	—	—	—	—	—	—	—
B	—	—	—	—	—	—	—	3	20	20	—	16 } 16 }	17	29	—
C	35	—	—	—	29	—	—	—	3	—	—	—	—	—	—
D	—	—	—	—	20	—	—	—	18	61	53	72	—	—	—
E	—	18	—	—	22	—	25	24	40	33	—	63	—	—	—

Refererend aan een reeds eerder gegeven voorbeeld (fase IV, ad 2) laten wij het aan de lezer over de consequentie uit deze onvolledigheid te trekken.

(3) Uit N.T.v.G. 99 (1955), 2498–2499. In dit artikel wordt o.m. een onderzoek in een kliniek beschreven, over het effect van hormonale behandeling van lijders aan dystrophia musculorum progressiva. In de aanvang van het artikel wordt gesproken over 20 lijders en in de samenvatting is vermeld dat een overzicht is gegeven van de uitkomsten 'verkregen bij een langdurige behandeling van 20 mannelijke lijders'. De grafieken waarop de verschillende uitspraken berusten betreffen echter slechts 8 lijders. Waarom zijn zij voor de overige 12 lijders niet gegeven? Zijn deze wel in de conclusies verwerkt? Zo ja, dan zijn deze niet voldoende te controleren; zo neen, waarom niet?

De grafieken bevatten voor elk van de 8 lijders, bij 3, 4 of 5 diëten, al dan niet gegeven in combinatie met verschillende hormonen, één uitkomst voor de creatine- en creatinine-uitscheiding per dieet[1]. De patiënten vormen hierbij hun eigen controle (is dit bij een langdurig onderzoek in de tijd mogelijk?). Niet is vermeld: (a) hoe lang elk dieet werd verstrekt, (b) of de verschillende diëten elkaar onmiddellijk opvolgden, (c) of per dieet één of meerdere waarnemingen zijn verricht. Indien meerdere waarnemingen beschikbaar zijn is van belang hoe hun verloop in de tijd is en hoe zij in één maat zijn samengevat, (d) hoe betrouwbaar de bepalingsmethoden zijn, enz.

Bij (c) willen wij opmerken dat twee diëten bijv. gemiddeld hetzelfde resultaat kunnen geven terwijl zij een verschillend effect op de uitscheiding hebben:

	Dieet A						Dieet B					
Weken	1	2	3	4	5	6	7	8	9	10	11	12
Uitscheiding mg/24 uur	350	410	620	790	850	940	980	810	720	530	490	440
Gemiddeld	660						660					

Bij dit gefingeerde voorbeeld geeft dieet A een geleidelijke *stijging* van de uitscheiding. Deze werkt nog even door bij dieet B, dat een *daling* van de uitscheiding veroorzaakt. De lezer vrage zich af wat er met deze gemiddelden gebeurt indien A langer of korter wordt toegepast dan B, of wanneer er waarnemingen ontbreken.

Het is duidelijk dat men zonder de bovengenoemde informatie omtrent de 'proefopzet' in ieder geval geen verantwoorde conclusies kan trekken. Maar

[1] Raadpleeg bij het volgende zo mogelijk deze grafieken.

zelfs als de gegevens volledig bekend zijn en de proefopzet juist blijkt zal men zich in de eerste plaats moeten afvragen of allerlei waargenomen verschillen mogelijk nog toevallig kunnen zijn en verder of een steekproef van één (of ten hoogste enkele) patiënten die op dezelfde wijze (?) behandeld zijn (één van de acht patiënten kreeg bijv. TPP, hoeveel van de resterende 12?) het trekken van meer algemeen geldende conclusies mogelijk maakt.

In het voorgaande overzicht hebben wij een aantal belangrijke overwegingen bij de verschillende onderzoekfasen belicht. Zij berusten grotendeels op een vakkundig en logisch doordenken van de problemen die zich bij de opeenvolgende stadia van het onderzoek voordoen. Als men tot een goede proefopzet en daarmee tot verantwoorde uitspraken wil komen, dienen zij bij de planning afzonderlijk, in hun onderlinge samenhang en in hun betrekking tot de statistische bewerking en analyse te worden bezien.

1.5. Analyse van bestaande gegevens

1.5.1. INLEIDING

Door vele instanties en instellingen worden gegevens verzameld die informatie kunnen verschaffen omtrent de gezondheidstoestand van de bevolking van een land.[1] Men kan hierbij een onderscheid maken tussen:

A. Gegevens die door bepaalde instanties op systematische en zo uniform mogelijke wijze worden verzameld en die na bewerking periodiek worden gepubliceerd in de vorm van zg. (medische) *statistieken*.[2] Deze statistieken hebben betrekking op de gehele bevolking van een land of op een — gewoonlijk duidelijk omschreven — deel daarvan.

B. Gegevens die door allerlei op medisch terrein werkende instellingen en personen bij hun werkzaamheden worden vastgelegd. Zij worden niet altijd (volledig) bewerkt en gepubliceerd en betreffen gewoonlijk delen van de bevolking van een land, die niet nauwkeurig gedefinieerd zijn.

Van beide categorieën geven wij enkele belangrijke voorbeelden. Onder *A* vallen o.a.:

1. Statistieken omtrent huwelijk, geboorte en sterfte ('vital statistics'). Voor de interpretatie hiervan heeft men gegevens nodig omtrent omvang, samenstelling en groei van de bevolking ('population statistics').

In Nederland worden deze statistieken verzorgd door het Centraal Bureau voor de Statistiek. Maandelijkse gegevens vindt men in de *Maandstatistiek van*

[1] Een volledig overzicht van de Nederlandse gegevens verschaft het 'Overzicht van de bestaande Nederlandse gezondheidsstatistieken', Verslagen en mededelingen betreffende de volksgezondheid 10 (1961), 989–1021.

[2] Onder een *statistiek* verstaat men in dit verband een geordende weergave (in de vorm van tabellen, grafieken, frequentieverdelingen, e.d.) van de uitkomsten van (grote aantallen) waarnemingen.

de Bevolking en de Volksgezondheid.[1] Deze bevat o.m. gegevens over huwelijk, geboorte en sterfte; de geboortecijfers van eerstgeborenen zijn bovendien afzonderlijk berekend. Verder zijn er cijfers over doodgeborenen, onwettig geborenen en zuigelingensterfte opgenomen (gebaseerd op de leeftijdsopbouw in 1936), alsmede gegevens uit enkele andere landen. In het *Statistisch Zakboek* worden jaarlijkse gegevens o.a. over de opbouw en loop van de bevolking en de sterfte naar doodsoorzaken samengevat. Verder noemen wij hier de *Sterftetafels*, afgeleid uit de waarnemingen over de perioden 1951–1955 en 1953–1955 en een publikatie over de *Sterfte in Nederland naar geslacht, leeftijd en doodsoorzaken*, 1921–1955. Bovendien stelt het C.B.S. voor sommige gegevens (o.m. voor de sterfte naar doodsoorzaken) uitvoeriger tabellen samen, die afzonderlijk tegen een kleine vergoeding worden verstrekt[2].

Van de internationale publikaties op dit gebied noemen wij het '*Annual of Epidemiological and Vital Statistics*' van de Wereld Gezondheids Organisatie en het '*Demographic Yearbook*' van de Verenigde Naties.

2. Statistieken betreffende de aangegeven gevallen van besmettelijke ziekten.

Deze gegevens, verzameld door de Hoofdinspectie voor de Volksgezondheid, worden o.m. per week opgenomen in het Ned. Tijdschrift voor Geneeskunde. Zij zijn ook te vinden in de Maandstatistiek en (verkort) in het Statistisch Zakboek.

3. Statistieken betreffende de keuring van dienstplichtigen.

Deze worden per lichting verzameld door de Inspectie van de Militair Geneeskundige Dienst. Een deel van de keuringsuitkomsten (redenen van afkeuring, gegevens over lengte en gewicht) zijn te vinden in het Statistisch Zakboek. Zie voor verdere bijzonderheden HOEFLAKE (1959).

4. Statistieken betreffende ziekte- en ongevallenverzuim.

Deze statistieken worden samengesteld door (a) Bedrijfsgeneeskundige diensten, (b) Uitvoeringsorganen van de ziekte- en ongevallenwet en van vrijwillige verzekering, (c) het C.B.S., dat sinds 1 januari 1960 de ziektestatistiek van de bedrijfsverenigingen verzorgt, (d) het Ned. Instituut voor Praeventieve Geneeskunde, dat op basis van vrijwillige deelname voor het bedrijfsleven een ziekte- en ongevallenverzuimstatistiek en een diagnosestatistiek voert.

5. Gegevens die door speciaal daartoe in het leven geroepen organisaties worden verzameld, zoals bijv. door de Centrale Kankerregistratie en het Centraal Bureau voor Keuringen op Medisch-Hygiënisch Gebied.

Wanneer de onder *A* genoemde of andere tot deze categorie behorende statistieken een deel van de bevolking betreft, dient men zich bij het trekken van conclusies vanzelfsprekend rekenschap te geven van de beperkingen, welke dit met zich mee brengt. Zo hebben bijv.

[1] Zie voor een overzicht van de inhoud van de jaargangen 1–9 (1953–1961)T. v. Soc. Gen. 40 (1962), 139–141.

[2] Zie hiervoor het 'Systematisch overzicht van de publicaties van het C.B.S.'

de ziekteverzuimstatistieken betrekking op groepen personen, die bij de overheid en in het bedrijfsleven werkzaam zijn, zodat zij overwegend mannen betreffen. Overigens worden zij meestal niet verzameld met het doel om er een 'gezondheidsstatistiek' uit op te stellen, maar dienen zij primair voor financiële of administratieve verantwoording of om bedrijven of diensten een (vergelijkend) beeld van het bij hen optredende verzuim te verschaffen (dit laatste geldt bijv. voor de statistieken van het N.I.P.G.). In 1.5.2 bespreken wij enkele algemene aspecten inzake de analyse van sterfte- en ziektestatistieken.

Uit het voorgaande blijkt, dat er zelfs in een zo volledig 'geadministreerd' land als het onze, grote bevolkingsgroepen zijn waarvan geen systematisch en op uniforme wijze verzamelde gegevens omtrent morbiditeit beschikbaar zijn. Zolang er op dit gebied geen volledige registratie aanwezig is, zal men *algemene* gegevens over de morbiditeit slechts kunnen verkrijgen door middel van speciale enquêtes onder de bevolking ('health and morbidity surveys'). De opzet en uitvoering hiervan wordt beschreven in het 7de rapport van de 'Expert Committee on Health Statistics' (1961) van de W.G.O.

De onder *B* genoemde gegevens worden door vele instellingen, die op medisch gebied werkzaam zijn, verzameld. Als voorbeeld noemen wij:

1. Ziekenhuizen en poliklinieken.
2. Speciale klinieken en centra (voor geestesziekten, voor geslachtsziekten).
3. Consultatiebureau's (voor a.s. moeders, zuigelingen en kleuters, t.b.c.).
4. Schoolartsendiensten, Bedrijfsgeneeskundige diensten (o.m. medische keuringen).
5. Huisartsen en specialisten.

De bij (5) verkregen gegevens zijn in het algemeen fragmentarisch en van arts tot arts verschillend. Zij worden in het algemeen niet bewerkt of gepubliceerd.[1] Een gedeelte van de onder (1) t/m (4) vallende gegevens wordt wel gepubliceerd, hetzij in de vorm van periodieke verslagen, hetzij in incidentele publikaties (die meestal op een deel van de gegevens of op bepaalde aspecten ervan betrekking hebben). In 1.5.3 bespreken wij enkele moeilijkheden die zich bij de analyse van dit soort waarnemingsmateriaal en de interpretatie van de daarbij verkregen uitkomsten kunnen voordoen. Tenslotte volgen in 1.5.4 enkele opmerkingen over gegevens omtrent medische voorzieningen.

[1] Praktijkanalyses, waarbij gedurende een bepaalde periode een nauwkeurige, doelgerichte registratie plaatsvindt vallen onder het observationele onderzoek. Hetzelfde geldt voor onderzoekingen waaraan een groep huisartsen deelneemt, zoals deze tegenwoordig regelmatig worden ingesteld door het Nederlands Huisartsen Genootschap.

1.5.2. STATISTIEKEN INZAKE STERFTE, ONGEVALSLETSEL EN ZIEKTE

Statistieken betreffende sterfte, ongevalsletsel en ziekte(verzuim) zijn pas dan bruikbaar voor het maken van vergelijkingen tussen verschillende deelpopulaties binnen één land, voor een onderzoek naar het verloop in de tijd, of voor het vergelijken van de situatie in verschillende landen, indien zij op uniforme wijze door middel van eenzelfde *klassificatiesysteem* zijn verzameld.

NEURDENBURG (1960) geeft een overzicht van alles wat (sinds 1853!) is voorafgegaan aan het in 1948 onder auspiciën van de Wereld Gezondheids Organisatie tot stand gekomen (en inmiddels in 1955 gereviseerde) *'Manual of the International Statistical Classification of Diseases, Injuries and Causes of Death'*, bestaande uit twee delen: een systematische indeling en een alfabetisch register.

Deel I van dit handboek geeft in de hoofdstukken I t/m XIII de ziekten, ingedeeld in 13 rubrieken. Hoofdstuk XIV betreft aangeboren afwijkingen, hoofdstuk XV bepaalde ziekten van de vroege jeugd. In hoofdstuk XVI zijn onduidelijke en onvoldoende omschreven ziektebeelden samengebracht. Hoofdstuk XVII betreft ongevallen, resp. gesplitst naar ongevalsoorzaken (N-lijst) en naar de aard van het letsel (E-lijst). Tenslotte geeft het boek een reeks aanvullende klassificaties (o.a. voor de oorzaken van doodgeboorte, voor inentingen en voor lichamelijke gebreken en invaliditeit), die tezamen de Y-lijst vormen. Tevens zijn drie speciale lijsten voor publikatie van gegevens over ziekte en sterfte gegeven: lijst A, een tussenlijst (intermediate list) van 150 oorzaken voor het groeperen van gegevens over ziekte en sterfte; lijst B, een verkorte lijst (abbreviated list) van 50 oorzaken voor het groeperen van sterftegegevens; lijst C, een speciale lijst van 50 oorzaken ten behoeve van ziekte- en ongevalsverzekeringsorganen.

De lijsten voor de klassificatie van doodsoorzaken hebben algemene ingang gevonden. De Nederlandse bewerking ervan, de *Nomenclatuur ten behoeve van de codering van ziekten en ongevallen* (C.B.S., 1954), stemt geheel overeen met de internationale klassificatie. Ten behoeve van ziekenhuizen en specialisten is echter nog een verdere differentiatie aangebracht.

Voor ziekten zijn in verschillende landen min of meer gewijzigde klassificatiesystemen in gebruik, die echter merendeels op de internationale lijst zijn gebaseerd. Voor Nederland is in 1955 door de Nationale Commissie voor Gezondheidsstatistieken een samenvattende lijst van ziektegroepen samengesteld die bestemd is voor publikaties over ziekte en ziekteverzuim, de *Nationale lijst van Ziektegroepen.* Deze bevat 16 hoofdgroepen met een onderverdeling in 59 subgroepen en zij wordt thans door de onder *A*.4 genoemde instanties gebruikt.

Sterftestatistieken

De sterftestatistiek is het oudste middel om een inzicht te krijgen omtrent de gezondheidstoestand van een bevolking. Toen de letaliteit

nog zeer hoog was kon men immers de morbiditeit bevredigend afmeten naar de mortaliteit. Ook nu leveren echter de sterftecijfers (en in het bijzonder de sterfte naar doodsoorzaken) belangrijke aanwijzingen op bij het vergelijkend onderzoek naar het gezondheidsniveau van verschillende landen en bij epidemiologisch speurwerk.

Bij het bestuderen van sterftecijfers dient men steeds rekening te houden met verschillen in de bevolkingsopbouw in verschillende jaren of in de te vergelijken landen. In 2.7.1 bespreken wij beknopt enkele methoden die daarbij kunnen worden gevolgd en kengetallen die kunnen worden berekend. Verder treden er enkele moeilijkheden op, waarvan men zich goed bewust moet zijn:

1. In de loop der tijd hebben periodiek veranderingen in het klassificatiesysteem plaatsgevonden. De mogelijkheid bestaat dat hierdoor een bepaalde doodsoorzaak meermalen een andere 'inhoud' heeft gekregen. Verder zijn in 1938 de richtlijnen gewijzigd die aangeven, hoe men bij meervoudige doodsoorzaken een keuze dient te maken tussen de verschillende mogelijkheden. Voordien werd de ziekelijke toestand die men als de 'voornaamste' doodsoorzaak beschouwde als primaire oorzaak opgenomen, maar op de revisie-conferentie in 1938 werd overeengekomen dat men hiervoor voortaan de 'underlying cause' zou nemen. De invoering van dit nieuwe selectiecriterium kan voor verschillende doodsoorzaken belangrijke gevolgen hebben gehad.

In 1938 bestond de 'tussenlijst' uit 87 groepen van doodsoorzaken. Bij de revisie in 1948 werd deze lijst uitgebreid tot 150 groepen, terwijl het nummer waaronder long- en bronchiaal carcinoom viel van 20 in 50 veranderde. Onderstaande tabel geeft het aantal sterfgevallen in CANADA in 1949, geklassificeerd volgens beide lijsten*:

Registratiesysteem	Aantal sterfgevallen
1938, I.L., no. 20, Malignant tumors of the respiratory system	1436
1948, I.L., no. 50, Malignant neoplasms of trachea, and of bronchus and lung, not specified as secundary	1154
verschil	282
Te verklaren als volgt:	
a) Sterfgevallen, die wèl onder 1938, I.L. 20, maar niet meer onder 1948, I.L. 50 werden gerubriceerd	309
b) Idem, wèl onder 1948, I.L. 50, niet onder 1938, I.L. 20	27
	282

* Bulletin of the W.H.O., Supplement 4, 1952, Comparability of Statistics of Causes of Death according to the fifth and sixth revisions of the International List.

Vermoedelijk betreffen de 309 gevallen die in groep 50 (1948) minder werden opgenomen o.m. kanker van de mondholte en secundaire longkankers, die volgens de klassificatie van 1938 onder groep 20 vallen.

2. Verbeteringen van de medische faciliteiten en vermeerdering van de medische kennis kan tot een verbeterde diagnostiek leiden, welke tengevolge heeft dat aan vele sterfgevallen nu een specifieke doodsoorzaak kan worden verbonden, waar men vroeger met een oppervlakkige aanduiding moest volstaan.[1] Verder bestaat de mogelijkheid dat in de tijd bij de artsen een veranderde instelling t.a.v. bepaalde doodsoorzaken optreedt, waardoor deze meer dan tevoren als primaire doodsoorzaak worden opgegeven.

3. Bij de vergelijking van sterftecijfers uit verschillende landen dient men zich in de eerste plaats af te vragen hoe volledig de registratie is, zowel wat de sterfgevallen als de bevolking in haar geheel betreft. Verder kan ook de betrouwbaarheid van de opgegeven cijfers sterk uiteenlopen. Het percentage gevallen in de rubriek 'doodsoorzaak onbekend' (causes of death unstated or ill-defined) kan hierover enige indruk geven. Een hoog percentage 'onbekend' houdt in dat men ook de wel geklassificeerde doodsoorzaken met reserve zal moeten beschouwen. De grote variatie die in dit opzicht internationaal wordt aangetroffen blijkt wel uit tabel 1.3.

Tabel 1.3. Sterfte zonder (duidelijk omschreven) doodsoorzaak bij mannen, 1946 en 1949*

Land	1946			1949		
	Tot. sterfte	Doods-oorz. onbekend	% onbekend v. totaal	Tot. sterfte	Doods-oorz. onbekend	% onbekend v. totaal
Dominicaanse Republiek	11240	3258	29,0	11041	2659	24,1
Salvador	16076	3171	19,7	14771	3421	23,2
Guatemala	28673	5433	18,9	29778	4380	14,7
Frankrijk	272354	34876	12,8	286018	20247	7,1
Noorwegen	14685	1043	7,1	14395	772	5,4
Chili	49145	3285	6,7	54425	2166	4,0
Nederland	41624	1993	4,8	41750	1793	4,3
Zwitserland	24989	362	1,4	25092	242	1,0
Australië (bl. bev.)	41283	150	0,36	42195	154	0,36
Zweden	35184	68	0,19	35043	61	0,17
Nieuw Zeeland (excl. Maori's)	8722	2	0,023	8765	—	—

* Annual Epidemiological and Vital Statistics, I, Vital Statistics & Causes of Death, W.H.O., 1952.

Ziekte- en ongevallen(verzuim)statistieken

Bij de analyse van deze statistieken gelden in de eerste plaats overwegingen analoog aan die bij de sterftestatistiek. In tegenstelling tot een sterfgeval heeft echter een ziektegeval een begin, een zekere

[1] Zie voor een fraai voorbeeld N.T.v.G. 104 (1960), 243.

duur en een einde, zodat men voor het karakteriseren van de morbiditeit in een bepaalde periode gebruik moet maken van kengetallen voor de frequentie en de ziekteduur. Wij komen hier in 2.7.2 op terug.

Ook ten aanzien van de diagnose kunnen zich bij ziekte moeilijkheden voordoen, die bij sterfte niet of in veel mindere mate optreden. Een onduidelijke diagnose is later vaak niet te achterhalen. Wanneer bij de controle op ziekteverzuim niet alle zieken door een arts worden bezocht, krijgt men met een aantal gevallen met 'diagnose onbekend' te maken, die een vertekening in het beeld van de andere diagnosegroepen kunnen opleveren. Verder is de keuze van de juiste diagnose moeilijk bij het optreden van meer dan één ziekelijke toestand en bij alle ziektegevallen met vage klachten. Tenslotte noemen wij de mogelijkheid dat zich tijdens de ziekte een andere ziekte voordoet (denk hierbij bv. aan een post-operatieve infectie), zodat men de totale ziekteduur over beide ziekten moet verdelen; indien men willekeur wil vermijden, moeten hiervoor duidelijke richtlijnen worden gesteld.

1.5.3. ANDERE MEDISCHE GEGEVENS

Door de in 1.5.1 onder *B* genoemde instellingen en personen worden per patiënt of onderzochte persoon vele gegevens vastgelegd. Naast algemene en anamnestische gegevens zijn dit diagnosen, uitkomsten van metingen, aantekeningen betreffende een toegepaste therapie en haar effect, enz. enz.

Deze en soortgelijke gegevens bevatten een schat aan informatie omtrent vele aspecten van de gezondheidstoestand van de bevolking, het verloop van ziekten, het resultaat van behandelingen, e.d. waaruit men betrekkelijk eenvoudig kan putten. Het waarnemingsmateriaal staat immers ter beschikking en het behoeft slechts te worden bewerkt en geanalyseerd. Bij nader inzien stuit men echter op een aantal moeilijkheden, die men zich vaak niet bewust blijkt te zijn en die in drie woorden kunnen worden samengevat: selectie, diversiteit en onvolledigheid. Bij de bespreking van deze moeilijkheden gaan wij, om de zaken zoveel mogelijk te concretiseren, uit van ziekenhuisgegevens.

Selectie. Het in een ziekenhuis ter beschikking staande waarnemingsmateriaal heeft gewoonlijk betrekking op een geselecteerde groep uit een (vaak niet nauwkeurig omschreven) deel van de gehele bevolking (wij verwijzen hier naar wat wij reeds in 1.3.6 opmerkten). Men zal dus steeds moeten overwegen tot welke reserves in de uitspraken deze beperkingen leiden. Dit is vervelend (en niet eenvoudig) en het wordt dan ook vaak achterwege gelaten.

Ook als in één ziekenhuis verschillende patiënten met dezelfde aandoening op verschillende wijze zijn behandeld (met verschillende

middelen, met verschillende doses van hetzelfde middel, e.d.) zal bij de keuze van de behandeling ongetwijfeld selectie zijn opgetreden. De behandeling zal immers per individu zijn gekozen, afhankelijk van factoren als: verloop of stadium van de aandoening, geslacht, leeftijd of gezondheidstoestand van de patiënt, te verwachten complicaties, enz. En als de verschillende behandelingen (overwegend) in verschillende perioden hebben plaatsgevonden, kunnen hun resultaten afwijkend zijn doordat de populatie waaruit de patiënten afkomstig zijn veranderd is, de virulentie van het ziekteverwekkende agens zich in de tijd heeft gewijzigd, de diagnostische criteria vroeger anders lagen, de kans op complicaties in de tijd verminderd is (bv. door de toepassing van antibiotica), enz. Men zal dan moeten trachten het materiaal te verdelen in subgroepen die uitsluitend verschillen in behandeling en die verder t.a.v. alle relevante factoren overeenstemmen. Afgezien van het feit dat men nooit zeker weet, of elke selectie inderdaad wordt uitgeschakeld, wordt zelfs materiaal van grote omvang op deze wijze reeds spoedig in zoveel subgroepjes, elk met een betrekkelijk klein aantal waarnemingen, gesplitst dat een goede statistische analyse moeilijk of vrijwel onuitvoerbaar kan worden.

Voorbeelden

Beschouw de volgende (fictieve) uitkomsten betreffende de resultaten van streptomycine-behandeling in hoge en lage doses bij lijders aan meningitis tuberculosa:

Dosering	Aantal	Overleden	Letaliteit
Laag	150	60	40%
Hoog	100	60	60%

De hoge dosis blijkt een grotere sterfte op te leveren dan de lage. Het ligt echter voor de hand aan te nemen, dat de dosering is aangepast aan de ernst van het ziektegeval, d.i. aan het klinisch stadium. Wordt dit mede in aanmerking genomen, dan vindt men de volgende uitkomsten:

Klinisch stadium	Dosering					
	Laag			Hoog		
	Aantal	Overleden	Let.	Aantal	Overleden	Let.
Vroeg	120	36	30%	20	4	20%
Middelmatig	20	15	75%	20	11	55%
Laat	10	9	90%	60	45	75%
Totaal	150	60	40%	100	60	60%

Nu blijkt in elk stadium de letaliteit bij de hoge dosis beneden die van de lage dosis te liggen.[1] Zieken in het vroege stadium krijgen echter voornamelijk een lage, die in het late stadium een hoge dosis. De misleidende uitkomst van de

[1] Dit voorbeeld is bedoeld om het mogelijke resultaat van de uitschakeling van enkele factoren te demonstreren. Dit blijkt het duidelijkst bij een vergelijking van de procentuele letaliteit in de subgroepen. Men dient zich hierbij echter te realiseren, dat men deze percentages, die soms op zeer kleine steekproeven berusten, in werkelijkheid niet zonder meer kan vergelijken.

totaalcijfers is een gevolg van de *'weging'* die optreedt, omdat de verdeling over het klinisch stadium bij de patiënten met de lage dosis tegengesteld is aan die bij de patiënten met de hoge dosis.

Men kan zich echter afvragen, of op deze wijze nog niet meer factoren moeten worden uitgeschakeld. Onderstel nu, dat men de patiënten in twee leeftijdsgroepen kan verdelen (in werkelijkheid zullen het er wel meer moeten zijn) en dat men dan vindt:

Klinisch stadium	Lage dosering						Hoge dosering					
	Jong			Oud			Jong			Oud		
	Aant.	Overl.	Let.	Aant.	Overl.	Let.	Aant.	Overl.	Let.	Aant.	Overl.	Let.
Vroeg	80	8	10%	40	28	70%	10	1	10%	10	3	30%
Middelmatig	5	1	20%	15	14	93%	5	1	20%	15	10	67%
Laat	2	1	50%	8	8	100%	20	5	25%	40	40	100%
Totaal	87	10	11½%	63	50	80%	35	7	20%	65	53	81½%

Onze uitspraak dient nu opnieuw een verandering te ondergaan. Het ziet er naar uit, dat bij de jongeren de dosering weinig ter zake doet. Bij de ouderen is er in het vroege en middelmatige stadium nog wel een gunstiger effect bij een hoge dosering te onderkennen, maar de behandeling blijkt bij patiënten in het late stadium niet meer te helpen. Het is duidelijk, dat men bij een verdergaande onderverdeling deze conclusies wellicht weer zal moeten amenderen (de jonge mannen in het vroege stadium met lage dosis reageren misschien anders dan de jonge vrouwen). Men houdt dan echter tenslotte subgroepen over die zo klein zijn, dat men er nauwelijks of niet meer mee kan werken.

Uit N.T.v.G. 98 (1954), 3295 e.v. In dit artikel worden de resultaten bewerkt bij patiënten met lues, die op verschillende wijzen zijn behandeld. Drie groepen patiënten (uit drie verschillende klinieken waarvan resp. 150, 100 en 144 patiënten werden bestudeerd) werden behandeld met de methode van Hoffmann (Bonnse kuren), een vierde groep van 154 patiënten (waarvan geen herkomst wordt opgegeven) is deels met penicilline-bismuth, deels met triplokuren behandeld.

Alleen van de 'kliniekgroep' van 394 patiënten wordt een overzicht gegeven van de regelmaat waarmee zij de kuur volgden (dit is niet gesplitst naar klinieken). Na enig puzzelen is te vinden dat 136 patiënten de kuur ongeregeld volgden en dat zich hieronder 27 patiënten bevinden die de kuur uitgesproken slecht volgden of zelfs vroeg of laat geheel wegbleven.

In het artikel zijn de uitkomsten van de behandelingen onvolledig en niet in tabelvorm opgenomen. Zij zijn in onderstaande tabel samengevat.

KLINIEKGROEPEN		Totaal	Niet genezen	% genezen
Vroege lues		?	?	90
Late lues		?	?	55
Totaal		394 (?)	?	
VIERDE GROEP				
Vroege lues	PB	?	1	90
	TR	?	1	bijna 100
Totaal vroeg		31	2	
Late lues	PB	67	42	37
	TR	46	18	63
Totaal laat		113	60	

Bij de kliniekgroepen is niet opgegeven of de resultaten op alle 394 patiënten betrekking hebben. Bij de vierde groep zijn bij de late lues 10 patiënten te weinig, waarover niets wordt medegedeeld.

Wij menen hier te kunnen volstaan met enkele korte opmerkingen:

(a) De resultaten van de kliniekgroepen kunnen meer informatie geven als men ze splitst naar 'trouw' in het volgen van de behandeling. Dit is zelfs noodzakelijk als men ze wil vergelijken met de resultaten van de vierde groep, waarbij onder de vroege lues geen, onder de late lues tien uitvallers (wegblijvers?) voorkomen. (b) Vierde groep — op welke gronden berust de verschillende behandeling (PB of TR)? Het is hoogst onwaarschijnlijk dat aselecte keuze heeft plaatsgevonden, maar welke criteria zijn dan wel gehanteerd en kunnen deze tot selectie hebben geleid? (Zijn bijv. bij de late lues PB en TR gelijkwaardig, maar wijst men — om welke redenen dan ook — ongunstiger gevallen aan PB toe, dan levert deze kuur vanzelfsprekend *daardoor* een slechter resultaat.) (c) Men kan de resultaten van de kliniekgroepen slechts dan met die van de vierde groep vergelijken, indien men aannemelijk maakt dat deze groepen gelijkwaardige patiënten betreffen. Zou bijv. de vierde groep uit particuliere patiënten bestaan, dan bestaat de mogelijkheid dat deze in relevante opzichten van de kliniekpatiënten verschillen en dat de resultaten mede hierdoor worden beïnvloed. (d) Waarom is binnen de kliniekgroepen geen specificatie van de uitkomsten van de drie verschillende klinieken gegeven? Dit levert een controle op de 'internal consistency' van de uitkomsten. (e) Men zou mathematisch-statistisch moeten onderzoeken of de waargenomen verschillen toevallig tot stand gekomen zouden kunnen zijn.

Diversiteit. Bij de registratie van medische gegevens in ziekenhuizen bestaat in vele landen (en ook in Nederland) in allerlei opzichten een gemis aan uniformiteit, waardoor de gegevens van verschillende ziekenhuizen vrijwel onvergelijkbaar zijn.[1] Afgezien van het feit dat niet overal dezelfde gegevens worden geregistreerd, worden vele min of meer subjectieve waarnemingen verricht, zonder dat daarbij gebruik wordt gemaakt van uniforme, duidelijk omlijnde criteria. Verder hebben wij in het voorgaande reeds gezien, dat zogenaamd objectief waarneembare of meetbare gegevens van onderzoeker tot onderzoeker (van laboratorium tot laboratorium) systematische verschillen te zien kunnen geven. Andere gegevens, zoals het beroep van de patiënt, worden vaak zo oppervlakkig genoteerd dat men er later weinig mee kan beginnen. Overigens kunnen ook de gegevens, die in één ziekenhuis binnen een wat langer tijdsverloop verzameld zijn, onderling onvergelijkbaar zijn, bijv. doordat mutaties in de medische staf met veranderingen in bepaalde maatstaven gepaard zijn gegaan.

Onvolledigheid. De gegevens die tijdens het verblijf in een ziekenhuis van een patiënt worden vastgelegd zullen van patiënt tot patiënt verschillen. Zij zijn immers afgestemd op de individuele patiënt en o.a. afhankelijk van de aard en de ernst van het ziektebeeld, de complicaties die zich gedurende de opname voordoen, de extra informatie die men

[1] Zie ook SCHEIJDE (1960).

bij een bepaalde (bijv. bij een nieuwe) therapie noodzakelijk acht, enz. Zo zal men sommige waarnemingen alleen bij ernstige en andere juist uitsluitend bij niet-ernstige zieken verrichten.

Gaat men nu later de beschikbare gegevens van een bepaalde groep patiënten analyseren, dan stuit men op allerlei onvolledigheden, waardoor sommige van deze patiënten geheel of gedeeltelijk buiten beschouwing moeten worden gelaten. Het percentage uitvallers-door-onvolledigheid is soms aanzienlijk; het bedraagt soms 20 tot 40 procent (wij kennen zelfs één geval waarin het ca. 80% bedroeg). Ook hierdoor kan een selectie ontstaan waarvan het karakter in de regel niet te bepalen is.

Een tweede moeilijkheid wordt veroorzaakt door het uitvallen van patiënten bij een langdurige poliklinische behandeling, of bij latere controles op het effect van een behandeling (bijv. van een operatie). Er zijn vele oorzaken die tot deze uitval kunnen leiden (de patiënt kan overleden zijn, verhuisd zijn, zich genezen achten, ontevreden zijn met de resultaten en zich tot een andere arts hebben gewend, enz.). Gewoonlijk zijn deze oorzaken slechts voor een deel van de uitvallers bekend en voor de overigen niet meer te achterhalen. Het staat echter vast dat men hier met een in verschillende opzichten geselecteerde groep te maken heeft en men mag dus niet zonder meer aannemen, dat deze dezelfde resultaten te zien gegeven zou hebben als de overblijvende groep.

Herhaaldelijk wordt de analyse van bestaande gegevens ter hand genomen, zonder dat men daarbij van een bepaalde vraagstelling uitgaat: men wil weten 'wat erin zit'. Vooral als men hierbij niet op de onvolmaaktheden van het materiaal let of deze onvoldoende kent, gaat men zodoende een 'l'art de grouper les chiffres' bedrijven, die gemakkelijk tot onjuiste conclusies kan leiden. Een verder gevaar is dat toevallige eigenaardigheden, die vooral bij een groot materiaal altijd wel voorkomen, tot theoretische bespiegelingen leiden, waarvan de juistheid dan op grond van hetzelfde materiaal wordt 'bewezen' (deze kringloop kan geheel onbewust plaatsvinden). Bij het ontbreken van een vraagstelling, die a priori is opgesteld, kan men de statistiek slechts als detectiemiddel gebruiken en 'detectief' gevonden vraagstellingen kan men slechts door een nieuw onderzoek op hun juistheid onderzoeken.

1.5.4. GEGEVENS OMTRENT MEDISCHE VOORZIENINGEN

Onder deze categorie vallen de gegevens die de bestaande medische voorzieningen betreffen, zoals gegevens over: ziekenhuizen (bedden, verpleegdagen, patiëntenverloop, verpleegduur, bezettingsgraad, enz.); getalsterkte, leeftijd en geografische verdeling van medisch personeel (artsen, tandartsen, verpleegsters); opleidingen (zowel van medici

als van medische hulpkrachten, verpleegsters, e.d.); wachttijden van patiënten in poliklinieken,[1] enz.

Door middel van deze gegevens kan men een beeld krijgen van de bestaande medische voorzieningen en van hun werking en doelmatigheid, dat kan leiden tot verbetering van de huidige situatie of gebruikt kan worden bij de planning van toekomstige voorzieningen.

1.6. Samenvatting

Nadat in 1.1 een aantal fundamentele statistische termen en begrippen is ingevoerd is in 1.2 het terrein van de statistiek beknopt beschreven. Hierbij is reeds de nadruk gelegd op het onderdeel van deze hulpwetenschap waaraan in dit boek de meeste aandacht zal worden besteed, nl. aan de mathematische statistiek, welke ons in staat stelt conclusies te trekken omtrent populaties uit aselect samengestelde steekproeven. De samenstelling van deze steekproeven wordt in 1.3 besproken.

In 1.4 is de problematiek van het groepsonderzoek op medisch gebied ingeleid. Dit onderwerp is bewust reeds in het eerste hoofdstuk geplaatst, omdat de mathematische statistiek haar taak slechts bij een goede proefopzet optimaal, d.i. als bewijsmiddel, kan vervullen. Men kan fundamentele tekortkomingen in de wijze waarop een onderzoek is opgezet of uitgevoerd niet opheffen door het gebruik van statistische methoden — wel kan men ze er soms mee verdoezelen. De statistiek kan een waardevolle tovenaarsleerlinge zijn, maar een tovenares is zij niet!

Tenslotte is in 1.5 de analyse van bestaand waarnemingsmateriaal aan de orde gesteld. In het bijzonder is hierbij gewezen op de zorgvuldigheid waarmee men bij het uitvoeren van zo'n analyse tewerk dient te gaan.

[1] Zie bijv. FORTUIN (1959).

HOOFDSTUK 2

REDUCTIE VAN WAARNEMINGSUITKOMSTEN

Wanneer men een aantal waarnemingen heeft verricht staan de uitkomsten daarvan in één of andere (bijv. chronologische of alfabetische) volgorde, die nog weinig inzicht geeft in de frequentieverdeling van het categorisch systeem of van de variabele grootheid, waarop de waarnemingen betrekking hebben. Men zal dan beginnen de waarnemingen per kenmerk of naar grootte te ordenen door het opstellen van een *frequentietabel*. In 2.1 bespreken wij hoe men daarbij te werk kan gaan.

Vaak bestaat er behoefte aan het geven van een aanschouwelijk beeld van waarnemingsuitkomsten, zoals bijv. van het verloop van de waarden van een grootheid in de tijd, van de in een frequentietabel weergegeven verdeling, of van de onderlinge samenhang van de waarden van twee grootheden. Voor dit doel staan verschillende vormen van *grafieken* ter beschikking, waarvan wij de voornaamste in 2.2 bespreken.

Een frequentietabel geeft een min of meer gedetailleerde beschrijving van de verdeling van de waarnemingen, maar soms wenst men deze verdeling in enkele kengetallen samen te vatten. Men kan voor deze karakterisering onder meer gebruik maken van centrale waarden of *gemiddelden*, die in 2.3 behandeld worden, en van maten voor de variabiliteit of *spreidingsmaten*, die in 2.4 worden besproken.

Tenslotte volgt als 2.5 een korte bespreking van enkele *typen van frequentieverdelingen*, die op medisch-biologisch gebied regelmatig worden waargenomen. Hierbij komt ook de toepassing van de gemiddelden en de spreidingsmaten aan de orde.

2.1. Frequentietabellen

2.1.1. ATTRIBUTEN (KWALITATIEVE KENMERKEN)

Wanneer de waarnemingen een categorisch systeem van kenmerken betreffen, kan men de frequentietabel opstellen door alle attributen van het systeem op te schrijven en door middel van turven of sorteren vast te stellen, hoe frequent elk kenmerk voorkomt.

Voorbeeld

2.1. In tabel 2.1a staan waarnemingen betreffende de bloedgroepen

van een steekproef van 150 arbeiders uit de populatie in voorbeeld 1.22. Deze waarnemingen staan, regel voor regel, van links naar rechts, in de volgorde waarin zij zijn verricht.

Men kan nog wel snel zien dat bloedgroep AB weinig voorkomt en ook, dat B minder frequent is dan A of O. De frequentietabel in 2.1b, die via een turftabel is opgesteld, geeft de volledige informatie omtrent de waargenomen verdeling van het systeem van bloedgroepen. (Wanneer, zoals bij dit voorbeeld, alle waarnemingen reeds bij elkaar staan kan men de turftabel achterwege laten en de frequenties per bloedgroep direct tellen.)

Tabel 2.1. Bloedgroepen van 150 arbeiders

a. Waarnemingen

A	A	A	B	A	A	A	A	A	B	A	O	A	A	O	A	O	O	O	A
A	O	A	O	A	A	O	A	A	O	A	O	O	A	A	O	A	O	O	A
A	A	O	B	B	O	B	A	B	O	A	AB	O	O	A	A	A	A	O	A
O	O	O	O	A	A	A	AB	O	O	O	A	O	A	AB	A	A	A	A	A
AB	O	O	O	B	A	O	A	O	O	O	A	O	A	A	A	B	O	B	A
B	O	O	A	A	A	AB	B	B	B	O	O	O	AB	A	O	O	A	O	A
O	O	B	O	A	A	A	B	B	A	A	B	A	AB	A	B	A	O	A	O
A	A	A	O	O	B	O	B	O	O										

b. Turftabel en frequentietabel

Bloedgroep	Turftabel	f
O	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸	55
A	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 ///	68
B	𝍸 𝍸 𝍸 𝍸	20
AB	𝍸 //	7
Totaal		150

2.1.2. WAARDEN

Bij het verrichten van waarnemingen aan een variabele verkrijgt men een reeks waarden, waaruit de frequentietabel eveneens door turven of sorteren kan worden opgesteld. Hierbij doen zich enkele aspecten voor die wij aan de hand van een tweetal voorbeelden bespreken.

Voorbeeld

2.2. De waarnemingen in tabel 2.2a betreffen het aantal erythrocyten (in mill. per mm^3) bij een steekproef van 50 mannen uit de populatie in voorbeeld 1.2. De discrete variabele is geteld met een afrondingsinterval van 0,1 mill. per mm^3. De hoogste waargenomen waarde is 6,3 en de laagste 4,5 mill. per mm^3. In tabel 2.2b, kolom 1, zijn deze en alle mogelijke tussenliggende waarden opgenomen. Via de

turftabel in kolom 2 verkrijgt men de frequentietabel in kolom 3.

De verkregen verdeling is echter nog onvoldoende overzichtelijk, omdat de waarnemingen over een relatief groot gebied verspreid zijn. Dit bezwaar kan worden opgeheven, als men door het twee aan twee samen nemen van aangrenzende waarden klassen formeert met een breedte van 0,2 (mill. per mm³), zoals in kolom 4 is aangegeven. De hoogste klasse bevat bij de door ons gekozen groepering de waarden 6,2 en 6,3, heeft als klassegrenzen 6,15 en 6,35 (kolom 5) en als klassewaarde 6,25 (kolom 6). De klassegrenzen worden in één decimaal meer gegeven dan de waarnemingen en eindigen op een 5. Hiermee bereikt men, dat de bovengrens van een klasse tevens de benedengrens is

Tabel 2.2. Aantal erythrocyten in mill. per mm³ bij een steekproef van 50 mannen uit de populatie in voorbeeld 1.2

a. Waarnemingen

5,7	5,3	5,6	5,4	5,6	4,9	5,0	5,1	4,9	5,3	5,0	5,4
5,2	5,9	5,4	6,3	5,2	5,5	5,2	5,1	4,6	5,8	5,5	4,5
5,5	5,7	5,3	5,2	5,5	5,2	5,3	4,8	5,0	5,7	4,8	5,3
5,9	5,7	5,3	5,4	5,5	5,0	5,6	4,6	5,1	5,0	4,8	5,4
6,1	5,2										

b. Frequentietabellen

(1) Waarden	(2) Turftabel	(3) f	(4) Klassen	(5) Klassegrenzen	(6) Klassewaarde	(7) f	(8) F
6,3 6,2	/	1 —	6,2 en 6,3	6,15—6,35	6,25	1	50
6,1 6,0	/	1 —	6,0 en 6,1	5,95—6,15	6,05	1	49
5,9 5,8	// /	2 1	5,8 en 5,9	5,75—5,95	5,85	3	48
5,7 5,6	//// ///	4 3	5,6 en 5,7	5,55—5,75	5,65	7	45
5,5 5,4	𝍸 𝍸	5 5	5,4 en 5,5	5,35—5,55	5,45	10	38
5,3 5,2	𝍸 / 𝍸 /	6 6	5,2 en 5,3	5,15—5,35	5,25	12	28
5,1 5,0	/// 𝍸	3 5	5,0 en 5,1	4,95—5,15	5,05	8	16
4,9 4,8	// ///	2 3	4,8 en 4,9	4,75—4,95	4,85	5	8
4,7 4,6	//	— 2	4,6 en 4,7	4,55—4,75	4,65	2	3
4,5 (4,4)	/	1 (—)	4,4 en 4,5	4,35—4,55	4,45	1	1
Totaal		50				50	

van de naasthogere klasse en dat toch elke waarde ondubbelzinnig kan worden geklassificeerd.

De gegroepeerde frequentietabel die op deze wijze wordt verkregen staat in kolom 7. Nu ziet men duidelijk dat een groot deel der waarnemingen in enkele centraal gelegen klassen valt en dat de frequentie afneemt, naarmate een klasse verder van dit centrum verwijderd is. Men zegt dan dat de verdeling een *centrale tendentie* vertoont.

Tenslotte zijn in kolom 8 nog de somfrequenties gegeven. Deze worden berekend door per klasse vast te stellen, hoeveel waarnemingen in totaal beneden de bovengrens van de klasse liggen. De somfrequentie is dus het aantal waarnemingen in een klasse vermeerderd met de somfrequentie van de aangrenzende lagere klasse.

Voorbeeld

2.3. De waarnemingen in tabel 2.3a betreffen de systolische bloeddruk (in mm Hg) gemeten bij een steekproef van 60 kinderen uit de populatie in voorbeeld 1.6. Deze continu verdeelde grootheid is gemeten met een afrondingsinterval van 1 mm Hg. De waarnemingen blijken te variëren van 86 t/m 130 mm Hg. De *spreidingsbreedte*, d.i. het verschil tussen de hoogste en laagste waarneming, bedraagt 130–86 = 44 mm Hg en is zo groot in relatie tot het totale aantal van 60 waarnemingen, dat men beter direct tot een klasse-indeling kan overgaan. In het algemeen gelden hierbij de volgende richtlijnen:

a. De klassebreedte. Kies een constante klassebreedte zodanig, dat ten minste 7 en ten hoogste 25 klassen bezet zijn. Het aantal klassen is gewoonlijk gunstig, als het van de zelfde orde van grootte is als de wortel uit het aantal waarnemingen. De klassebreedte kan globaal worden bepaald door de spreidingsbreedte te delen door de wortel uit het aantal waarnemingen.

Kiest men de klassebreedte te klein, dan ontstaat een te fijne indeling die tot een onvoldoende overzicht leidt. Kiest men te brede klassen, dan heeft dit een onverantwoord verlies aan nauwkeurigheid ten gevolge, daar bij de latere manipulaties met de verkregen verdeling de klassewaarde (= het klassemidden) de waarnemingen in de klasse onvoldoende representeert. In geval van twijfel kan men beter een te kleine klassebreedte aanhouden, omdat men daaruit door verdere groepering altijd nog op een grovere indeling kan overgaan.

Er zijn gevallen, waarin de klassebreedte bezwaarlijk constant kan worden gekozen, bijv. bij verdelingen met een zeer grote variabiliteit. Een bepaalde klassebreedte kan dan bij het centrum van de verdeling goed voldoen, terwijl in de uiteinden (staarten) van de verdeling vele dun bezette of lege klassen (met frequentie 0) optreden. Kiest men echter een grotere breedte, dan gaan in het centrum details verloren. Het zal dus soms noodzakelijk zijn de klassebreedte in het centrum kleiner te kiezen dan in de uiteinden van de verdeling.

b. De ligging der klassen. Kies die groepering waarbij zo weinig mogelijk lege klassen optreden. Als de uitkomsten de neiging vertonen

zich bij bepaalde waarden op te hopen, ligt het voor de hand de klassen zo te plaatsen, dat de klassewaarden zo veel mogelijk met deze concentratiepunten samenvallen. De fout die men maakt door de klassewaarde te beschouwen als representant voor alle waarnemingen in de klasse wordt zodoende zo klein mogelijk.

Het afronden van klassewaarden. De klassewaarde is het gemiddelde van de beneden- en bovengrens van de klasse, bijv.:

Klassegrenzen	Klassewaarde
399,5 —499,5	449,5
399,95 —499,95	449,95
399,995—499,995	449,995

Waarden als 449,95 en 449,995 kan men zonder bezwaar afronden tot 450.

Tabel 2.3. Systolische bloeddruk (mm Hg) bij een steekproef van 60 kinderen van 6 en 7 jaar uit de populatie in voorbeeld 1.6

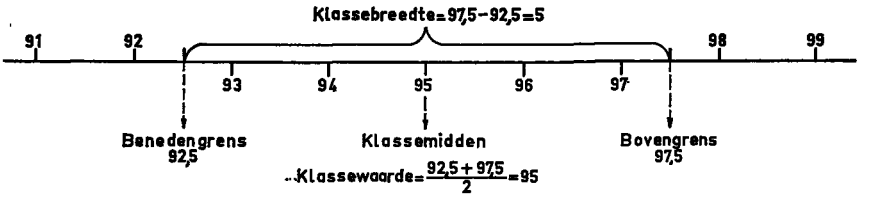
a. Waarnemingen

100	111	105	126	112	110	96	115	114	99
107	94	109	119	104	111	105	98	99	115
110	107	86	110	119	103	116	102	118	125
88	113	106	114	130	104	90	100	116	110
105	89	115	95	109	108	121	109	103	105
105	111	114	101	120	105	109	95	100	106

b. Frequentietabellen

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Klassen van – t/m	Klasse-waarde	Puntendiagram	f	$\frac{f}{n}$	F	$\frac{F}{n}$
128–132	130	.	1	0,0167	60	1,0000
123–127	125	..	2	0,0333	59	0,9833
118–122	120	.....	5	0,0833	57	0,9500
113–117	115	.....	9	0,1500	52	0,8667
108–112	110	.....	13	0,2167	43	0,7167
103–107	105	.....	14	0,2333	30	0,5000
98–102	100	.....	8	0,1333	16	0,2667
93– 97	95	....	4	0,0667	8	0,1334
88– 92	90	...	3	0,0500	4	0,0667
83– 87	85	.	1	0,0167	1	0,0167
		Totaal	$n = 60$	1,0000	—	—

Bij de bloeddrukwaarden in tabel 2.3a bedraagt de spreidingsbreedte 44 mm Hg, zodat er 9 à 10 klassen met een breedte van 5 mm kunnen worden gevormd. Het valt op, dat een belangrijk deel der waarnemingen (22 van de 60) op een 0 of een 5 eindigt: blijkbaar is bij het aflezen (onbewust) afgerond. Wij kiezen daarom de ligging van de klassen zodanig, dat de klassewaarden veelvouden van 5 zijn, dus: 85, 90, . . ., 130. Ter verduidelijking is onderstaand de klasse met klassewaarde 95 getekend:



In tabel 2.3b is de frequentietabel van de waargenomen bloeddrukwaarden met deze klasse-indeling opgesteld. In kolom 1 staan de waarden, die tot elke klasse behoren en in kolom 2 de klassewaarden. In kolom 3 zijn deze keer punten in plaats van turfstreepjes genoteerd; het voordeel hiervan is, dat een *puntendiagram* ontstaat, dat reeds een aanschouwelijk beeld van de frequentieverdeling geeft. In de verdere kolommen zijn de (relatieve) frequentie- en somfrequentieverdelingen geplaatst.

Tenslotte een opmerking omtrent het voorkomen van *uitbijters*, d.w.z. sporadisch voorkomende extreem hoge of lage waarden. Als men, zoals bij voorbeeld 2.3, direct tot het opstellen van een gegroepeerde frequentieverdeling overgaat, dient men de waarnemingen eerst op het voorkomen van uitbijters te inspecteren. Deze vergroten nl. de spreidingsbreedte vaak aanzienlijk en men zou dan tot een te grote klassebreedte komen, waardoor de overige waarnemingen in te weinig klassen bijeen worden gebracht. In principe kan men daarom, als uitbijters optreden, beter de klasse-grootte op de overige waarnemingen baseren en de uitbijters in een extra toegevoegde, zg. open klasse, onderbrengen. Onderstel bijv., dat aan de waarnemingen in tabel 2.3a een uitbijter ad 144 mm Hg wordt toegevoegd. Kiest men nu de klassebreedte op grond van de overige 60 waarnemingen, dan ontstaat de volgende verdeling (vergelijk met tabel 2.3b):

Klasse	83–87	88–92	.....	128–132	133–137	138–142	143–147
f	1	3	.....	1	0	0	1

Men vervangt nu de drie hoogste klassen door de open klasse:

133 en hoger+

en geeft de uitbijter in een voetnoot onder de tabel aan:

+) 144 mm Hg

2.1.3. VERLIES BIJ REDUCTIE TOT EEN FREQUENTIEVERDELING

De reductie van de oorspronkelijke waarnemingen tot een frequentietabel houdt in, dat men de eerste definitief door de laatste vervangt en met de overzichtelijker en beter hanteerbare tabel verder werkt. Hierbij gaat echter informatie verloren inzake:

a. De volgorde der waarnemingen. Meestal heeft deze geen betekenis; de kinderen waar de waarnemingen in tabel 2.3 betrekking op hebben kunnen bijv. in alfabetische volgorde gemeten zijn, of in de willekeurige volgorde waarin zij zich voor het onderzoek hebben opgesteld.

Soms kan de volgorde van belang zijn in verband met een onderzoek naar de aselectiviteit van de steekproef. Een tweede uitzondering treedt op, als men juist de fluctuatie van de waarnemingen (bijv. de wijze waarop zij in de tijd variëren) wil onderzoeken. Wij komen hier in hoofdstuk 11 op terug.

b. De identiteit der afzonderlijke waarnemingen. Bij het samenstellen van een gegroepeerde frequentietabel worden alle waarnemingen binnen een klasse a.h.w. gelijkgeschakeld: men doet, alsof zij gelijkmatig over de klasse verdeeld zijn, zodat zij gerepresenteerd kunnen worden door hun klassemidden (de klassewaarde). Fijnere nuances van de variabiliteit gaan hierdoor verloren, maar als de klasse-indeling niet te grof is wordt het geringe verlies aan nauwkeurigheid gecompenseerd door de grotere overzichtelijkheid en de vereenvoudiging van de verdere bewerkingen.

c. De samenhang tussen twee tegelijk waargenomen verschijnselen. Heeft men bijv. van een aantal kinderen de leeftijd en de lichaamslengte bepaald, dan kan men uit de afzonderlijke verdelingen van leeftijd en lengte niet meer zien, welke leeftijd en welke lengte samengaan. Dit bezwaar kan worden ondervangen, door een *correlatietabel* samen te stellen. Bij een klein aantal waarnemingen kan men beter een *spreidingsdiagram* tekenen; wij bespreken deze grafiek in 2.2.3.

2.1.4. DE CORRELATIETABEL

Bij het vervaardigen van een correlatietabel moet men de waarnemingen van beide categorische systemen of van beide grootheden gelijktijdig verwerken.

Bij kwalitatieve gegevens stelt men daartoe van het ene categorische systeem een frequentietabel op voor iedere categorie van het tweede systeem. Betreft het ene systeem bijv. de bloedgroepen en het tweede de geslachten, dan ziet de correlatietabel er als volgt uit:

		Bloedgroepen				Totaal
		O	A	B	AB	
Geslacht	M	/	/			
	V	/			/	
Totaal						

Per geslacht is er een tabel voor de bloedgroepen (men kan natuurlijk even goed zeggen: per bloedgroep is er een tabel voor de geslachten). Voor elke persoon kan men nu, op grond van de waargenomen kenmerken, een turfstreepje in het juiste vakje van de correlatietabel plaatsen. Bij wijze van voorbeeld zijn in de voorgaande tabel streepjes gezet voor de volgende vier personen: 1) Man — bloed-

groep A, 2) Vrouw — bloedgroep O, 3) Vrouw — bloedgroep AB, 4) Man — bloedgroep O. Het is duidelijk dat de tabel ook de totaalverdelingen van de twee categorische systemen bevat.

Wanneer de per element beschikbare waarnemingsparen de grootheden x en y betreffen, gaat men in principe op dezelfde wijze te werk. Men stelt van de ene grootheid (x) een frequentietabel op voor alle waarden van de tweede grootheid (y). De tabel bevat tevens de frequentietabellen van de grootheid y voor alle waarden van x, alsmede de totaalverdelingen van x en y. Evenals bij het opstellen van één frequentietabel zal men hierbij gewoonlijk de waarden van x, resp. y, direct in klassen indelen volgens de in 2.1.3 gegeven regels.

Voorbeeld

2.4. Tabel 2.4a bevat de waargenomen zuurstofgehalten (in vol%) van arterieel bloed (x) en veneus bloed (y) van 50 gezonde mannen. De gehalten van het arteriële bloed bezitten een spreidingsbreedte van 22,3 — 17,3 = 5 vol% en die van het veneuze bloed hebben een spreidingsbreedte van 16,1 — 10,6 = 5,5 vol%. Kiest men het aantal ongeveer als $\sqrt{50}$, d.i. 7 à 8, dan kan de klassebreedte op 7 of 8 vol% worden gesteld; wij hebben voor beide grootheden als klassebreedte 8 aangehouden. De klasse-indelingen zijn, daar de waarnemingen geen bijzonderheden (zoals opeenhopingen bij bepaalde waarden) vertonen,

Tabel 2.4. Correlatietabel: zuurstofgehalten (in vol%) van arterieel bloed (x) en veneus bloed (y) van 50 gezonde mannen

a. Waarnemingsparen

Persoon	Arterie x	Vena y	Persoon	Arterie x	Vena y	Persoon	Arterie x	Vena y
1	18,2	11,4	18	18,7	13,8	35	20,0	14,2
2	20,5	14,0	19	19,6	11,7	36	20,5	13,1
3	20,9	15,1	20	17,5	10,6	37	18,9	12,9
4	18,5	12,0	21	18,8	13,2	38	18,5	12,0
5	21,9	14,6	22	18,8	12,5	39	19,0	12,3
6	18,4	12,0	23	20,9	12,5	40	18,0	11,0
7	17,4	11,3	24	20,0	13,9	41	20,8	13,7
8	22,3	15,3	25	19,4	11,9	42	17,3	11,9
9	20,4	11,9	26	20,4	14,2	43	20,4	12,8
10	18,2	12,7	27	20,1	13,3	44	19,7	12,0
11	19,3	12,7	28	20,1	12,6	45	20,7	14,4
12	20,3	12,8	29	19,0	14,5	46	19,1	11,5
13	18,2	12,2	30	18,6	11,9	47	18,7	11,8
14	20,3	14,8	31	19,4	12,6	48	18,9	11,2
15	20,3	13,4	32	19,8	12,1	49	22,2	16,1
16	20,6	13,5	33	22,2	15,2	50	18,6	11,8
17	20,8	14,4	34	20,9	14,4			

Uit J. W. Pfeiffer, 'Introduction to the analysis of experimental data', cursus Biostatistics 202, Columbia University, U.S.A.

b. Correlatietabel

y \ x	16,6 –17,3	17,4 –18,1	18,2 –18,9	19,0 –19,7	19,8 –20,5	20,6 –21,3	21,4 –22,1	22,2 –22,9	f_y
15,7—16,4								1	1
14,9—15,6						1		2	3
14,1—14,8				1	3	3	1		8
13,3—14,0			1		4	2			7
12,5—13,2			4	2	4	1			11
11,7—12,4	1		7	4	2				14
10,9—11,6		2	2	1					5
10,1—10,8		1							1
f_x	1	3	14	8	13	7	1	3	50

vrij willekeurig genomen en staan in tabel 2.4b. Met het oog op het uitturven van de waarnemingsparen in de vakjes van deze correlatietabel kan men van elke klasse het best de laagste en de hoogste waarde vermelden.

2.1.5. SORTEERMETHODEN

Bij een groot aantal waarnemingen wordt het opstellen van een frequentieverdeling of correlatietabel door middel van turven een vervelend werk, waarbij fouten haast onvermijdelijk zijn. Men kan dan beter een sorteermethode gebruiken. Dit geldt in 't bijzonder, als men aan elk element meerdere waarnemingen heeft verricht (bijv. van elk individu de lengte, het gewicht, de leeftijd, het hemoglobinegehalte, enz. heeft vastgesteld). Het sorteren kan op verschillende manieren plaatsvinden.

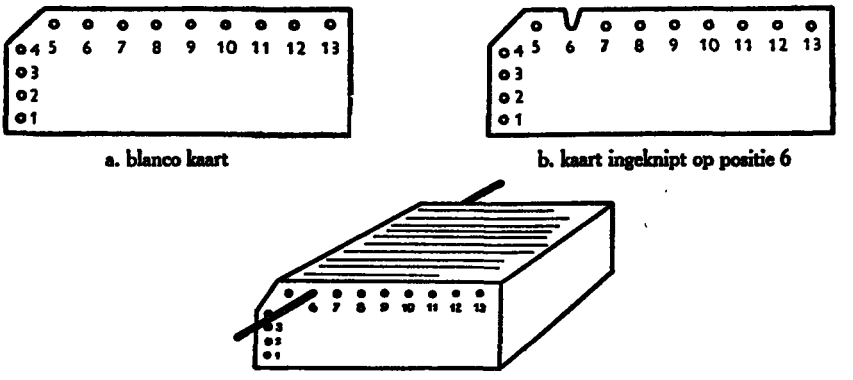
a. Sorteerkaartjes. Men vervaardigt kleine kaartjes, waarop alle gegevens per element systematisch, d.i. in een vaste volgorde, worden genoteerd. Men kan deze kaartjes bijv. uit millimeterpapier knippen of ze laten drukken (het best op zg. briefkaartkarton). Met deze kaartjes, waar in figuur 2.1 een voorbeeld van gegeven is, kan men nu zelf alle gewenste sorteringen verrichten.

	(1)	(2)	(3)	(4)	(5)	(6)
(7)	(8)	(9)	(10)	(11)	(12)	(13)
(14)	(15)	(16)	(17)	(18)	(19)	(20)
(21)	(22)	(23)	(24)	(25)	(26)	(27)

1 = Volgnummer
2 = Geslacht
3 = Leeftijd in jaren
4 = Welstandsklasse
5 = Lichaamslengte in cm
6 = Gewicht in kg
7 = Hemoglobinegehalte in g%
enz.

Figuur 2.1. Sorteerkaartje.

b. Randponskaarten. Dit zijn registratiekaarten, waarvan één of meer randen voorzien zijn van voorgeponste gaatjes op onderling gelijke afstanden. Wij beperken ons tot de eenvoudigste versie van dit systeem met één rij gaatjes[1]. Per gaatje kan men nu bijv. twee kenmerken aangeven, door al dan niet een deel van de kaartrand bij dit gaatje met een tang of ponsapparaat weg te knippen (zie figuur 2.2). Laat deze kenmerken bijv. 'jongen' en 'meisje' zijn en onderstel, dat bij een jongen gaatje 6 wordt ingeknipt en bij een meisje niet. Men kan dan sorteren door de kaarten in dezelfde stand te plaatsen (door het afgeknipte hoekje kan men zien of er kaarten verkeerd staan) en op 'positie' 6 een breinaald door de stapel kaarten te steken. De kaarten van de meisjes blijven aan de naald hangen, die van de jongens blijven staan. Men kan heel goed enige tientallen gaatjes in de randen aanbrengen, zonder dat het formaat van de kaart onhandig groot wordt.



a. blanco kaart

b. kaart ingeknipt op positie 6

c. stapel kaarten met doorgestoken pen, gereed voor het sorteren op positie 6

Figuur 2.2. Randponskaarten met naaldsortering [uit DE WOLFF].

Een categorisch systeem met meer dan twee categorieën kan worden aangebracht, door twee of meer aangrenzende gaatjes te combineren: gebruikt men n posities, dan heeft men 2^n mogelijkheden. Zo zou men bijv. de bloedgroepen als volgt in 2 posities kunnen onderbrengen:

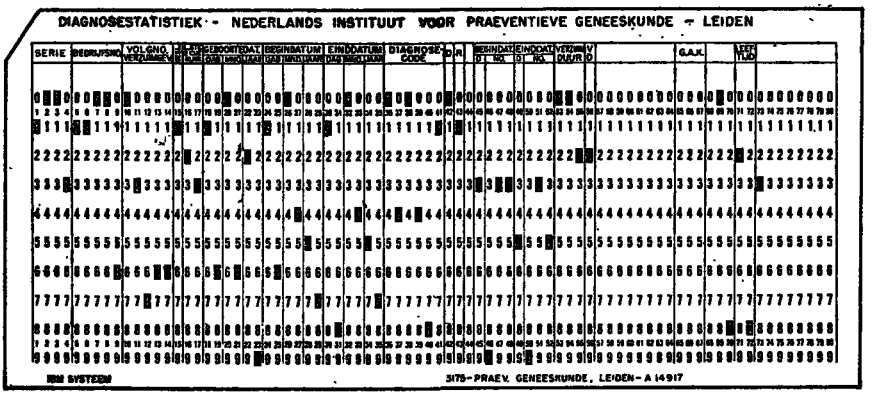
Bloedgroep	*Positie 1*	*Positie 2*
O	Open	Open
A	Open	Dicht
B	Dicht	Open
AB	Dicht	Dicht

[1] Er zijn echter vele verfijningen mogelijk. Voor verdere bijzonderheden wende men zich tot de firma's, die deze kaartsystemen leveren, zoals de N.V. Blikman & Sartorius te Amsterdam en de N.V. Ahrend-Globe te Hilversum.

Door met meerdere naalden tegelijk te werken, kan men tegelijkertijd op meerdere posities sorteren. Voor het verwerken van grote aantallen kaarten in één sorteergang kan een sorteerapparaat worden gebruikt.

c. Ponskaarten. Het ponskaartensysteem is een technisch volmaakte methode, waarbij het sorteren en tellen zeer snel automatisch wordt uitgevoerd. Men maakt hierbij gebruik van een kaart met gewoonlijk 80 kolommen (posities). In elke kolom kan door middel van een ponsmachine één der getallen 0, 1, 2, . . ., 9 worden geponst. Hierdoor kan in één positie een categorisch systeem of een variabele met ten hoogste tien categorieën (die de codenummers 0 t/m 9 ontvangen) worden ondergebracht (zie figuur 2.3). Door tevens gebruik te maken van twee zg. bovenponsingen — boven de genummerde hokjes — kan men in een kolom ook één der letters van het alfabet vastleggen. Combineert men twee posities, dan zijn er reeds 100 mogelijkheden (00 t/m 99), enz.

Bij het gebruik van ponskaarten dient men tevoren nauwkeurig vast te stellen welke gegevens men in de kaart wil opnemen, hoeveel categorieën of klassen elk gegeven telt, hoe deze gecodeerd worden, hoeveel posities daarvoor nodig zijn, enz. Op grond hiervan wordt de kaartindeling met bijbehorende codelijst gemaakt. Met deze lijst kunnen de gegevens in de juiste volgorde 'vercijferd' worden op een zg. ponsconcept, dat de basis voor de ponskaart vormt. Er zijn echter ook kaarten, waarop men de gegevens direct op de juiste plaats kan aanstrepen, waarna de ponsing vol-automatisch geschiedt (het zg. marksensing systeem). Men kan hierbij echter slechts 54 posities per kaart gebruiken.



Figuur 2.3. Ponskaart.

Wij zullen op de vele mogelijkheden, die het ponskaartensysteem ook voor de verdere bewerking biedt, niet ingaan. Bij incidenteel gebruik van het systeem kan men bij verschillende instellingen machinecapaciteit huren. Deze instellingen beschikken over deskundigen, die kunnen beoordelen of een project zich voor mechanische bewerking leent, wat de kosten hiervan zullen zijn, enz.

2.1.6. OPGAVEN

2.1. Stel uit onderstaande waarnemingen (hemoglobinegehalten in g% van jongens van 7 t/m 9 jaar) een frequentie- en somfrequentieverdeling samen:

14,0	13,4	13,6	13,0	13,1	13,7	13,0	13,4	13,8	13,3	12,9	13,3	13,1
13,6	13,0	13,1	12,9	13,4	13,0	13,2	13,0	12,6	12,7	12,8	13,2	12,9
12,1	12,4	12,2	13,0	12,8	12,6	12,8	12,7	12,1	12,6	11,8	12,7	
12,5	12,3	12,8	12,4	12,6	12,8	11,9	11,5	12,5	12,6	12,9	12,4	

2.2. Onderstaande waarnemingen betreffen de bezinkingssnelheid der erythrocyten na 45 minuten bij 68 vrouwen van 20 t/m 34 jaar. Maak een frequentie- en relatieve frequentieverdeling.

7	12	16	5	20	6	13	25	3	16	3	9	8	7	15	14	5
2	23	31	8	4	9	8	2	6	11	22	5	18	26	3	28	19
18	8	15	3	14	5	6	10	12	4	8	15	6	10	23	6	12
35	5	7	10	7	11	12	14	7	9	6	13	6	3	4	9	8

2.3. Gegeven zijn de volgende waarnemingen betreffende de schedellengte in mm (x) en het hersengewicht in gr (y) van 60 mannen tussen 20 en 60 jaar. Stel een correlatietabel samen.

x	y	x	y	x	y	x	y	x	y	x	y
173	1395	172	1341	160	1509	191	1752	170	1310	174	1456
193	1697	190	1490	172	1618	185	1668	181	1654	167	1412
156	1222	179	1478	196	1810	177	1351	167	1434	171	1367
188	1648	175	1340	174	1283	183	1569	178	1365	169	1271
176	1547	178	1556	176	1480	182	1745	174	1443	178	1425
160	1310	173	1481	183	1396	175	1214	186	1576	170	1534
172	1444	182	1479	180	1480	166	1250	199	1865	165	1339
166	1320	180	1549	179	1567	176	1535	171	1543	176	1462
175	1450	176	1656	193	1594	185	1392	176	1465	184	1469
189	1582	163	1190	170	1476	171	1425	179	1362	182	1517

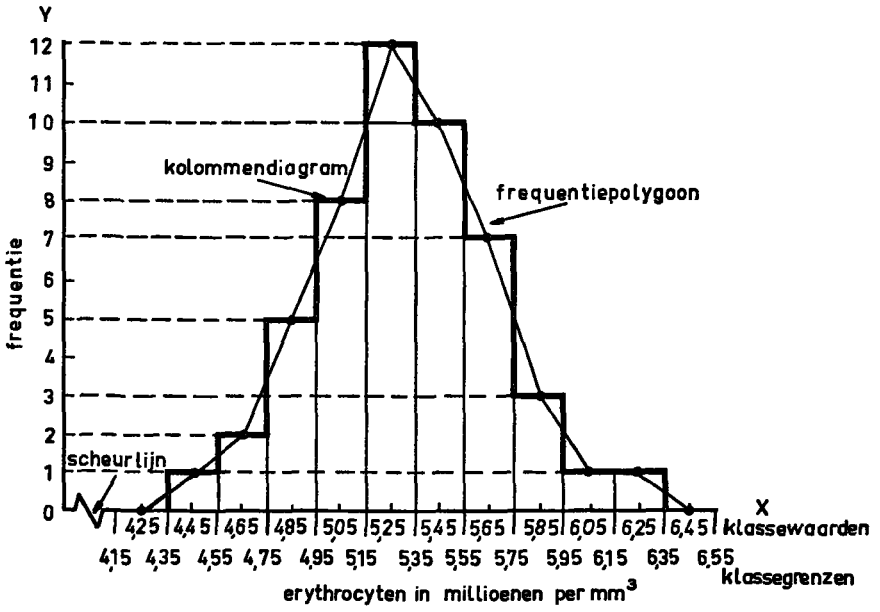
2.2. Grafieken [1])

2.2.1. HET KOLOMMENDIAGRAM

Men kan de frequentieverdeling met *constante klassebreedte* van een variabele als een kolommendiagram of histogram tekenen, door op de (horizontale) X-as de klassen af te zetten en op iedere klasse een rechthoek te plaatsen, waarvan de hoogte de klassefrequentie aangeeft,

[1] Zie eerst bijlage 2, par. 6.

gemeten volgens een schaalverdeling op de (verticale) Y-as. Deze schaal wordt zo gekozen, dat de 'hoogte' van het kolommendiagram (bij de klasse met de grootste frequentie) $\frac{2}{3}$ tot $\frac{3}{4}$ is van de 'breedte' (van de benedengrens van de laagste t/m de bovengrens van de hoogste klasse). Als voorbeeld is in figuur 2.4 het kolommendiagram van de frequentieverdeling in tabel 2.2b getekend.



Figuur 2.4. Kolommendiagram en frequentiepolygoon van de verdeling in tabel 2.2b.

Als de frequentieverdeling een categorisch systeem van attributen betreft tekent men gewoonlijk een *staafdiagram*. Hierin worden de attributen aequidistant, maar van elkaar gescheiden, op de X-as afgezet en wordt op elk attribuut een smalle kolom (of lijn) opgericht, waarvan de hoogte de (relatieve) frequentie van het attribuut weergeeft.

Als van de frequentieverdeling van een grootheid de *klassebreedte niet constant* is, dient men hier rekening mee te houden bij het vaststellen van de juiste kolomhoogte in het kolommendiagram. De frequenties van klassen met verschillende breedte zijn immers niet direct vergelijkbaar, zodat men moet werken met de *frequentiedichtheid*, d.i. de frequentie per eenheid van de beschouwde grootheid.

Voorbeeld

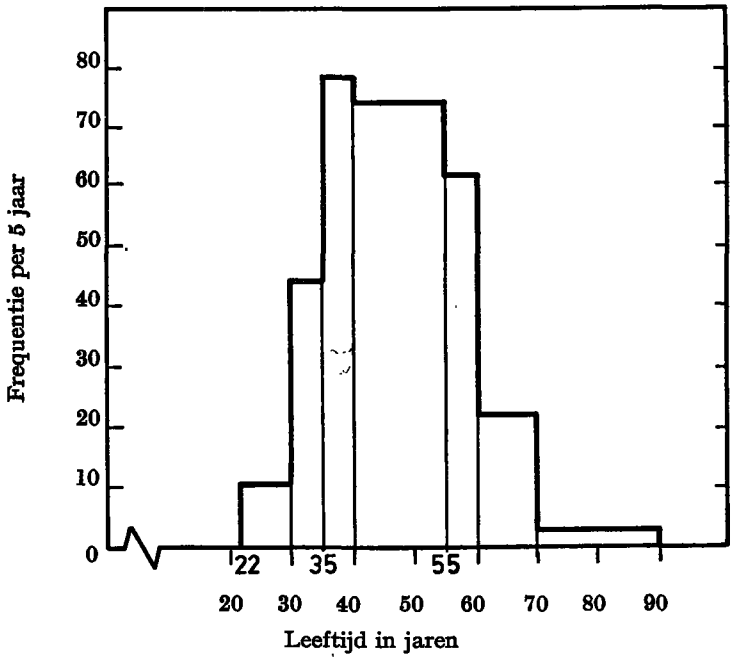
2.5. In tabel 2.5 is de leeftijdsverdeling opgenomen van 488 patiënten, behandeld voor cervix carcinoom.

Tabel 2.5. Patiënten behandeld voor cervix carcinoom+

Leeftijd	Klasse-Breedte	Aantal patiënten	
		Totaal	Per 5 jaar
22—29	8	18	11,25
30—34	5	45	45
35—39	5	79	79
40—54	15	225	75
55—59	5	63	63
60—69	10	45	22,5
70—89	20	13	3,25
Totaal		488	

+ Uit: Carcinoma of the Cervix Uteri, J. Am. M. Ass., June 25, 1932.

De leeftijdsklassen zijn van sterk uiteenlopende breedte zodat hun frequenties niet zonder meer kunnen worden vergeleken. Men verkrijgt een juist beeld door de frequentiedichtheid, bijv. per 5 jaar te nemen, d.i. door elke frequentie met 5/klassebreedte te vermenigvuldigen (zie de laatste kolom in tabel 2.5).



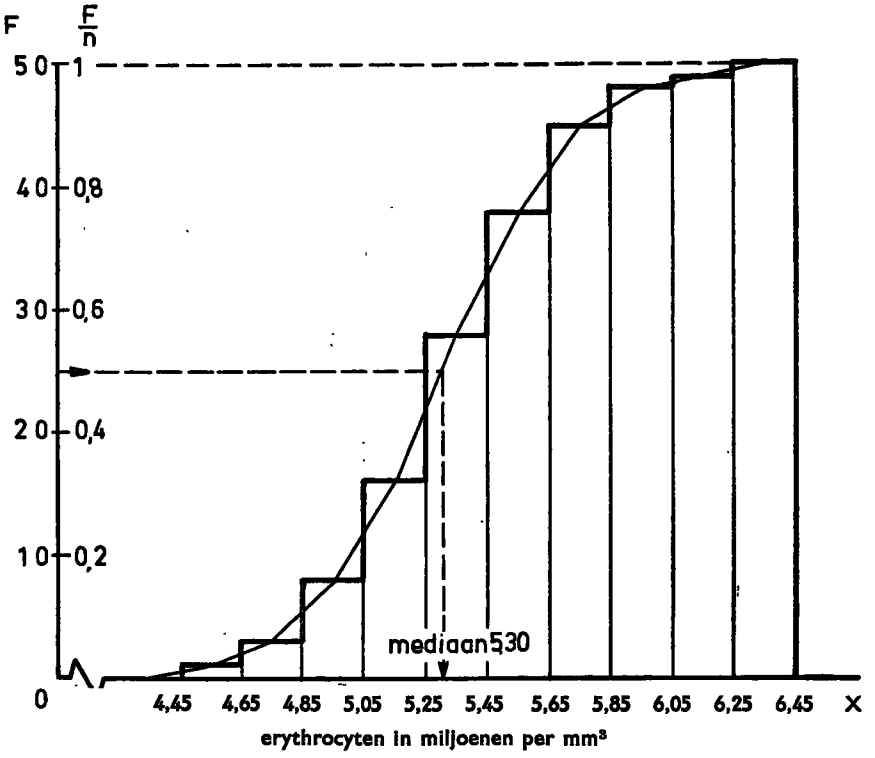
Figuur 2.5. Kolommendiagram van de verdeling in tabel 2.5.

In figuur 2.5 is een kolommendiagram van de leeftijdsverdeling getekend, waarin op de Y-as de frequentiedichtheid per 5 jaar is uitgezet. Evenals bij een kolommendiagram van een frequentieverdeling met constante klassebreedte zijn in deze figuur de oppervlakken van de kolommen op de (ongelijke) klassen evenredig met de klassefrequenties.

Als men uit de frequentieverdeling met constante klassebreedte van een variabele de somfrequentieverdeling heeft opgesteld, kan deze grafisch worden voorgesteld door een trappenlijn. Men richt hiertoe op de X-as op de bovengrenzen van de klassen kolommen op, waarvan de hoogte evenredig is aan de somfrequentie van de klasse waarbij de bovengrens behoort. In figuur 2.6 is de trappenlijn getekend, die behoort bij de somfrequentieverdeling in tabel 2.2b.

2.2.2. DE FREQUENTIEPOLYGOON

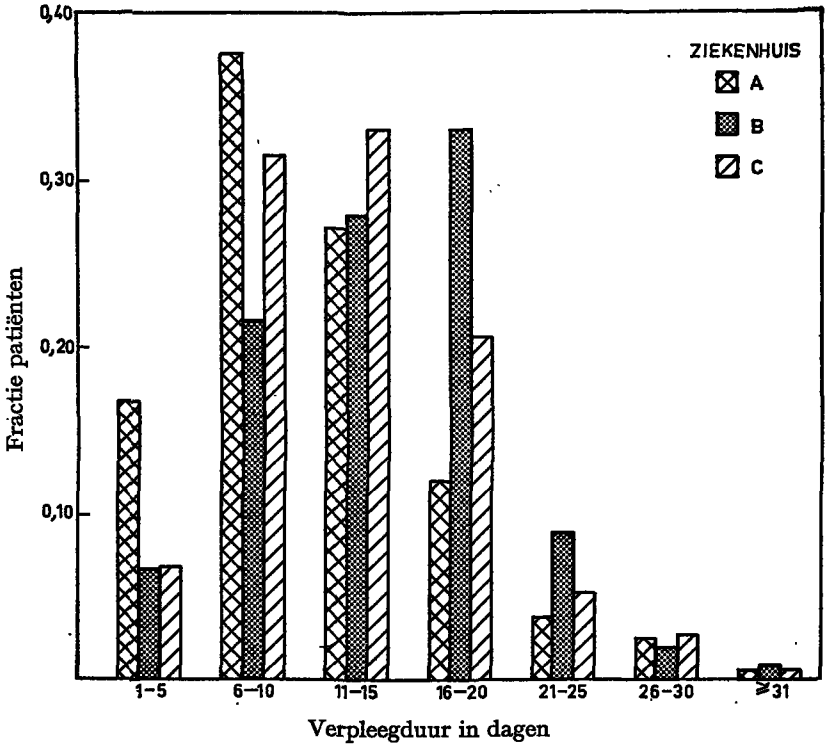
Bij het tekenen van de frequentiepolygoon van een frequentieverdeling met constante klassebreedte richt men op de X-as op elk klasse-



Figuur 2.6. Grafiek van de somfrequentieverdeling in tabel 2.2b (kolom 8).

midden een loodlijn op, waarvan de hoogte de klassefrequentie weergeeft. Vervolgens verbindt men de bovenpunten van deze ordinaten tot een gebroken lijn (polygoon). Hierbij worden ook de middens betrokken van de links en rechts toegevoegde klassen met frequentie nul (zie figuur 2.4). Op analoge wijze verkrijgt men via de ordinaten op de bovengrenzen van de klassen de somfrequentiepolygoon (zie figuur 2.6.)

In het algemeen verdient het kolommendiagram als grafiek van een frequentieverdeling de voorkeur boven de frequentiepolygoon. Niet alleen komt in het kolommendiagram het discrete karakter van de waarnemingen beter tot uitdrukking, maar verder is het beeld dat een frequentiepolygoon geeft altijd min of meer vertekend, omdat de oppervlakken onder de polygoon boven de klassen niet precies evenredig zijn aan de klassefrequenties. Als men echter beschikt over een groot aantal waarnemingen van een continue grootheid, zodat men een frequentieverdeling kan samenstellen die een groot aantal klassen met een relatief kleine klassebreedte bevat, kan men beter de frequentiepolygoon tekenen.



Figuur 2.7. Staafdiagrammen van de verdelingen van de verpleegduren in de gynaecologische afdelingen van 3 ziekenhuizen. Aantallen patiënten: A = 801, B = 288, C = 776.

Het komt vaak voor dat men twee of meer frequentieverdelingen grafisch wil vergelijken door ze op hetzelfde assenstelsel te tekenen. Wanneer de verdelingen op verschillende aantallen waarnemingen berusten, dient men hierbij vanzelfsprekend de *relatieve* frequentieverdelingen te gebruiken[1]. Kolommendiagrammen leveren dan gewoonlijk een onduidelijk beeld op, daar zij grotendeels op elkaar vallen. Wanneer het aantal waarnemingen groot en het aantal klassen niet te klein is kan men de relatieve (som)frequentiepolygonen tekenen. Figuur 2.15 geeft daar een voorbeeld van. In andere gevallen kan men beter gebruik maken van staafdiagrammen. In figuur 2.7 zijn op deze wijze de frequentieverdelingen van de verpleegduren in de gynaecologische afdelingen van drie ziekenhuizen vergeleken. Daartoe is op elke verpleegduurklasse voor elk ziekenhuis, een kolom opgericht, die de relatieve frequentie van de klasse representeert. Zowel de afzonderlijke verdelingen van de ziekenhuizen als de kenmerkende verschillen tussen deze verdelingen zijn nu direct af te lezen.

Wanneer de te vergelijken frequentieverdelingen berusten op steekproeven van grote omvang kan men nog wel met een grafische vergelijking volstaan (men dient dan bij de figuur de grootte van de steekproeven te vermelden). Voor kleinere steekproeven zal men — tenzij de verschillen tussen de verdelingen zeer evident zijn — langs mathematisch-statistische weg moeten nagaan, of de waargenomen verschillen nog toevallig kunnen zijn. Toetsingsmethoden die hiertoe kunnen worden gebruikt worden later besproken.

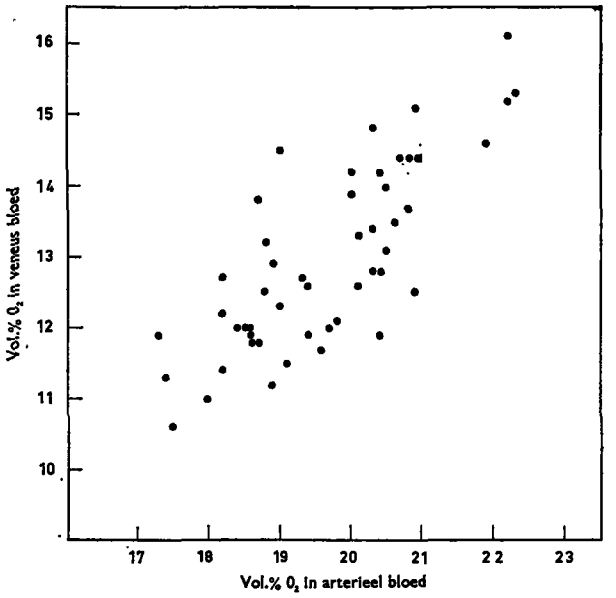
2.2.3. HET SPREIDINGSDIAGRAM

Het spreidingsdiagram kan worden samengesteld als men beschikt over de waarnemingsparen (x, y), dus als men per element over een waarneming van de grootheid x en over een waarneming van de grootheid y beschikt. In een assenstelsel met op de X-as een schaalverdeling van de grootheid x en op de Y-as een schaalverdeling van de grootheid y tekent men per element een punt met als abscis de waargenomen x-waarde en als ordinaat de waargenomen y-waarde van dat element. In figuur 2.8 is als voorbeeld het spreidingsdiagram van de waarnemingsparen in tabel 2.4 getekend.

2.2.4. HET VERLOOPDIAGRAM

Wanneer men het verloop van een grootheid in de tijd wil weer-

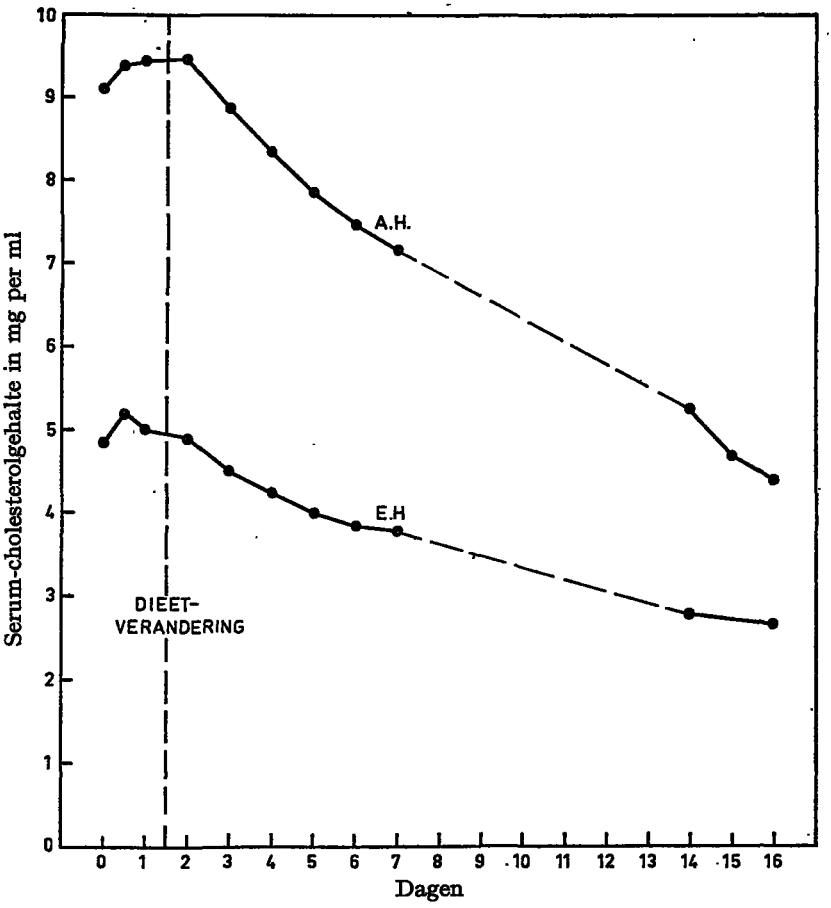
[1] Toch gebruikt men bij deze vergelijking nog vaak de absolute verdelingen. Enkele voorbeelden: N.T.v.G. (1957), blz. 747, id. (1958), blz. 2083, id. (1959), blz. 1430—1440; zie ook de grafieken betreffende verpleegduren, die als bijlage aan het rapport 'Waardering van medische gegevens uit ziekenhuizen' van de Gezondheidsorganisatie TNO (1958) zijn toegevoegd.



Figuur 2.8. Spreidingsdiagram van de waarnemingsparen in tabel 2.4a.

geven, zet men op de X-as de tijdstippen (x) en op de Y-as een schaalverdeling van de waargenomen grootheid (y). Elke waarneming kan dan door een punt (x, y) worden aangegeven. Belangrijk is, dat men de verhouding van de schalen op beide assen goed kiest, zodat noch een te smalle en hoge, noch een te brede en lage figuur ontstaat. Als de tijdstippen niet te ver uiteenlopen verbindt men de punten (x, y) gewoonlijk door lijnstukjes, teneinde hiermee globaal het verloop te accentueren. Is er echter in een regelmatige reeks tijdstippen een langere periode waarin geen waarnemingen zijn verricht, dan is het wenselijk dit duidelijk aan te geven door het laatste punt vóór en het eerste na deze periode door een stippellijn te verbinden.

In figuur 2.9 is het verloopdiagram van de serum-cholesterolgehalten van twee personen getekend. Figuur 2.10 betreft dezelfde gegevens, maar door verschillende onjuistheden ontstaat hier een misleidend beeld. De schaal op de X-as is te kort ten opzichte van die op de Y-as, terwijl bovendien de afstand tussen de 7-de en de 14-de dag is verkleind. Hierdoor verkrijgt men een overdreven indruk van de daling van de cholesterolgehalten, die nog wordt versterkt doordat de punten (x, y) zijn weggelaten en de lijnstukken tussen de punten van de 7-de en 14-de dag niet zijn gestippeld. De schaal op de Y-as is zonder enige aanduiding even beneden 3 mg per ml afge-



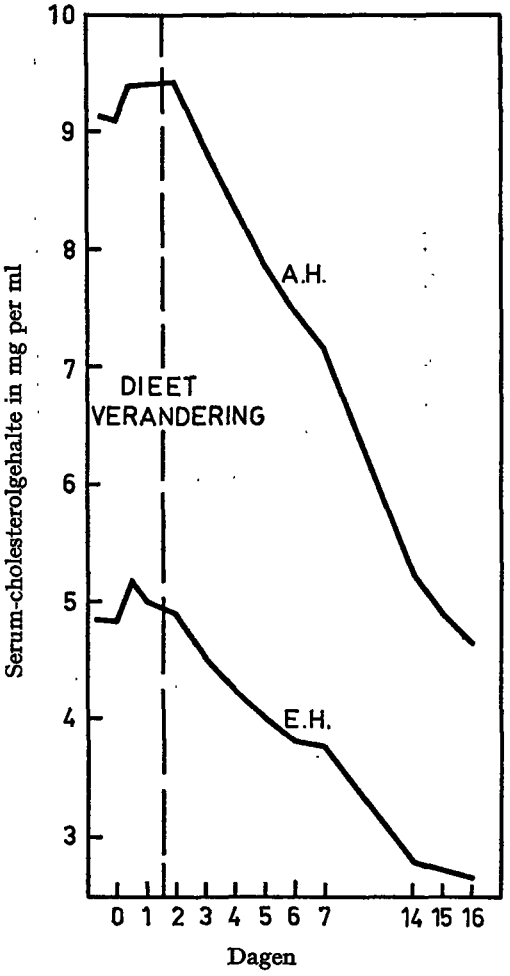
Figuur 2.9. Verloop van de serum-cholesterolgehalten bij twee mannelijke lijders aan familiaire hypercholesteraemie, vóór en na de overgang naar een vetvrij vegetarisch dieet, geheel vrij van cholesterol. Voor de dieetverandering was de cholesterolopname reeds betrekkelijk laag. Uit KEYS, Science 112 (1950), 79—81.

broken. Dit wekt de suggestie dat het gehalte van persoon E.H. tot 0 nadert en het vertekent de verhouding van de gehalten van de twee personen.

Men kan een verloopdiagram ook als staafdiagram tekenen. Figuur 2.11 geeft hier een voorbeeld van.

2.2.5. ANDERE GRAFIEKVORMEN

Er zijn nog verschillende andere grafiekvormen, die echter in hoofdzaak worden gebruikt voor pedagogische doeleinden, propa-



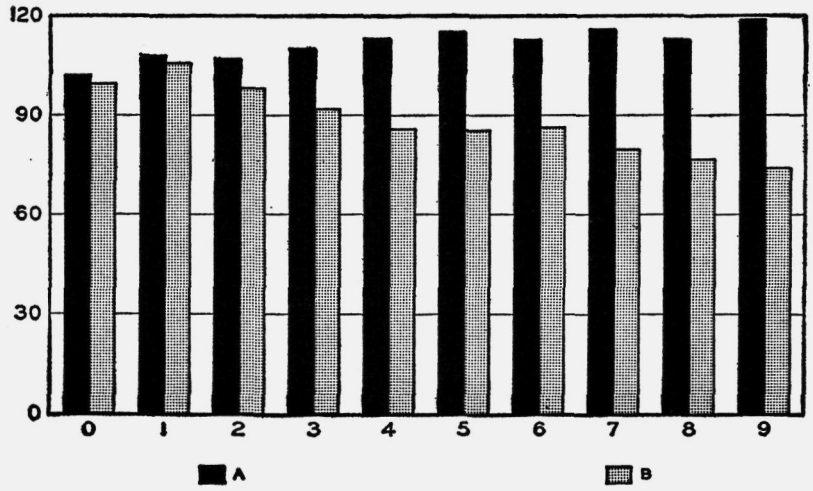
Figuur 2.10. Als figuur 2.9, maar onjuiste grafiek.

ganda, reclame, e.d. Wij volstaan daarom met enkele voorbeelden.

Een *cirkeldiagram* (zie figuur 2.12) leent zich goed voor het weergeven van een procentuele samenstelling. Bij de *beeldgrafieken* werkt men met figuurtjes, die van toepassing zijn op de grootheid die men grafisch voorstelt. De frequentie van een kenmerk of klasse wordt hier door een aantal van deze figuurtjes weergegeven, waarbij elk figuurtje een zeker aantal personen of objecten voorstelt. Het jaarlijks verschijnende Statistisch Zakboek van het Centraal Bureau voor de Statistiek bevat steeds een aantal zeer geslaagde beeldgrafieken (figuur 2.13 is aan het zakboek 1957 ontleend). *Cartogrammen* worden gebruikt om de geografische spreiding van een grootheid weer te geven. Op de kaart van het betreffende gebied geeft men de verschillende frequentie in de deelgebieden aan door een verschillende arcering. Ook kan men in de deelgebieden beeldgrafieken aanbrengen (het reeds genoemde Statistisch Zakboek bevat ook hier in de regel enkele voorbeelden van).

2.2.6. WENKEN EN WAARSCHUWINGEN

Voor het tekenen van grafieken gebruikt men bij voorkeur grafiekenpapier, dat in grote verscheidenheid verkrijgbaar is. Behalve het gewone millimeterpapier — met op de X- en Y-as een onderverdeling in 1 mm, met samenvattingen per 5 mm en per 10 mm — bestaan er ook

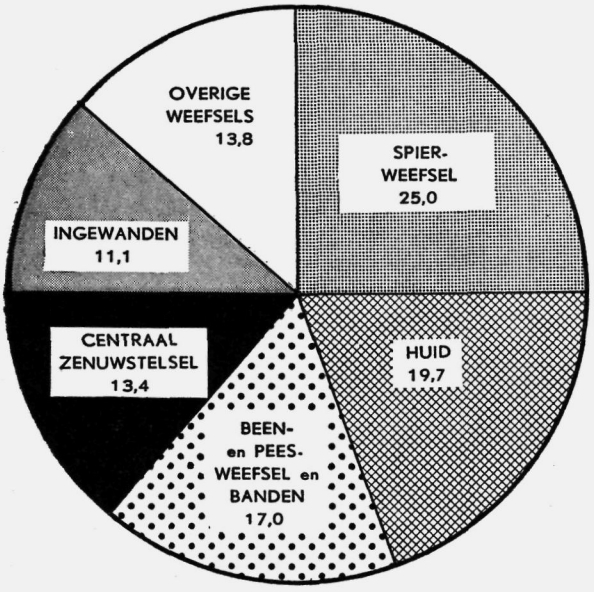


Figuur 2.11. Gemiddelde bloeddruk in mm Hg van twee groepen ratten gedurende 9 opeenvolgende weken. Groep A kreeg een synthetisch dieet met 0,1 g kaliumchloride, groep B ontving dezelfde voeding zonder kaliumchloride. Uit FREED en FRIEDMAN, Science, 112 (1950), 788—789.

papieren met afstanden van 2, 3, 4, 5 en 10 mm. Het 2 mm-papier is zeer geschikt voor reproductiedoeleinden. Het cm-papier leent zich bijv. voor het maken van grote grafieken voor lezingen of tentoonstellingen. Enkel of semi-logarithmisch papier heeft een aequidistante verdeling op één as, terwijl op de andere een logarithmische verdeling is aangebracht. Men kan dit papier bijv. gebruiken voor een grafische voorstelling van verschijnselen, waarbij in elke tijdseenheid ongeveer dezelfde procentuele toe- of afneming optreedt. Bovendien kan het worden gebruikt als men te maken heeft met een frequentieverdeling, waarvan de frequenties sterk uiteenlopen. Bij gebruik van aequidistant papier worden de afmetingen van de grafiek nl. onpraktisch groot, of de grafiek wordt (bij de lage frequenties) zeer onnauwkeurig. Tenslotte noemen wij nog het dubbel logarithmisch papier met op beide assen een logarithmische indeling. Voor een uitgebreide en met voorbeelden geïllustreerde bespreking van deze en andere grafiekpapieren verwijzen wij naar FERRO (1954).

Voorbeeld

2.6. Tabel 2.6 bevat het aantal nog niet genezen malariapatiënten, dat na resp. 1, 2, 3 en 4 maanden uit 1000 patiënten resteert. Per maand zien wij een vrij constante procentuele afneming (de procenten zijn verkregen door te nemen: 100 × aantal in een maand genezen patiënten/aantal patiënten aan het begin van de maand). Beschouwt men log A (zoals gewoonlijk gebruiken



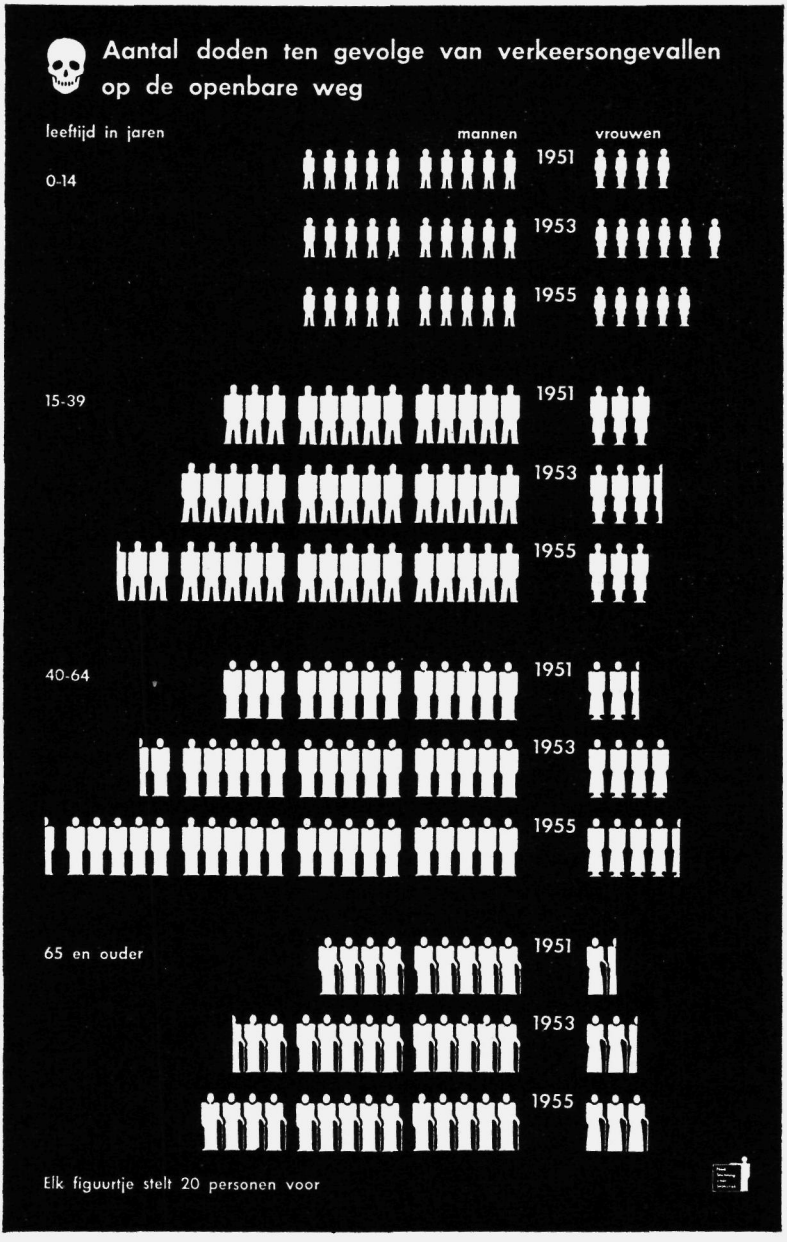
Figuur 2.12. Cirkeldiagram van de procentuele samenstelling van het totale lichaamsgewicht van de mens bij de geboorte. Uit CROXTON (1953).

wij logarithmen met grondtal 10, zie bijlage 2, par. 4) dan ziet men tussen de logarithmen een redelijk constant *verschil*. De gegevens laten zich dus goed voorstellen in een grafiek (figuur 2.14), waarin voor de tijd in maanden op de X-as een metrische en voor het aantal niet genezen patiënten op Y-as een logarithmische schaal is aangebracht. Het verloop van dit aantal in de tijd laat zich goed weergegeven door een op het oog getrokken rechte.

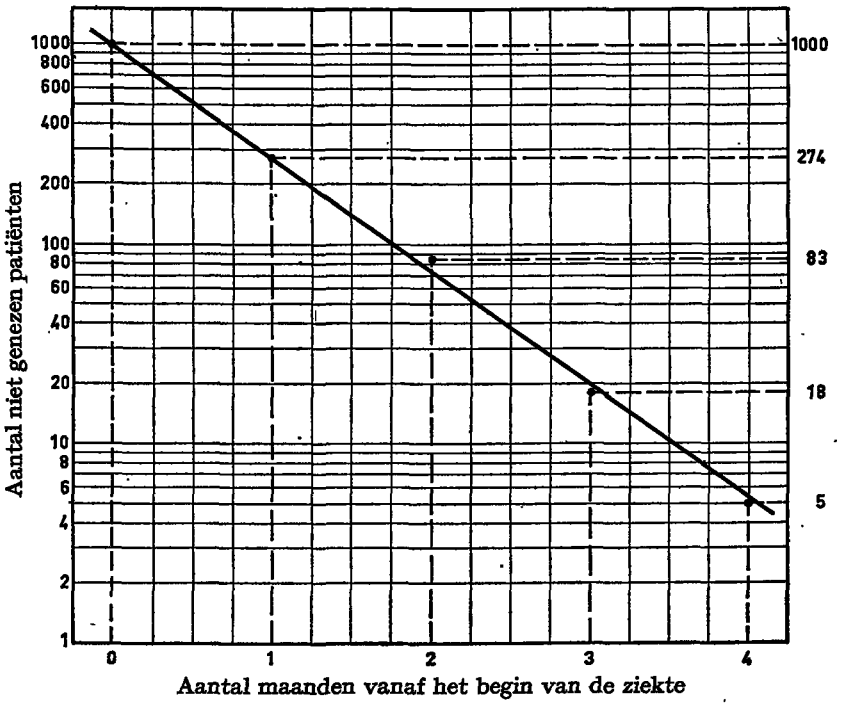
Tabel 2.6. Ziekteduren van malariapatiënten

Aantal maanden vanaf begin der ziekte	Aantal niet genezen patiënten A	Procentuele afneming	log A	Verschillen
0	1000		3,00	
		27%		0,56
1	274		2,44	
		30%		0,52
2	83		1,92	
		22%		0,66
3	18		1,26	
		28%		0,56
4	5		0,70	

Een grafiek behoort op zichzelf leesbaar te zijn en dient daartoe een volledige aanduiding der weergegeven grootheden, schaalverdelingen, enz. te bevatten. Soms laat men de aanduiding der schalen achterwege, waardoor de weergegeven uitkomsten zich geheel aan een verantwoorde interpretatie onttrekken. In advertenties treft men deze fout veel aan (zie bijv. N.T.v.G. 98 (1954), advertentie naast pag. 1360).



Figuur 2.13. Beeldgrafiek (Uit Statistisch Zakboek, C.B.S., 1957).

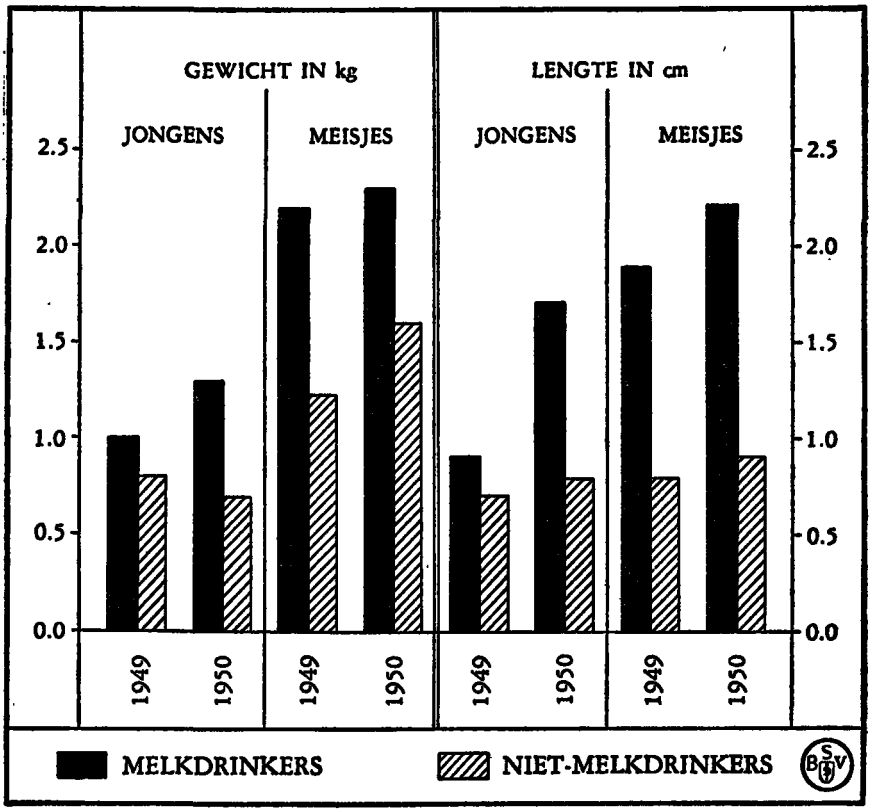


Figuur 2.14. Grafiek op semi-logarithmisch papier: ziekteduren van malariapatiënten.

Te veel lijnen in één grafiek leiden tot een onoverzichtelijk en ondoorzichtig geheel, waaruit men weinig of niets kan aflezen (zie figuur 1.5).

Van veel belang is verder de eenhedenkeuze op de beide assen. Door bij het tekenen van een kolommendiagram de klasse-indeling op de X-as te comprimeren en de frequentieschaal op de Y-as uit te rekken kan men de indruk wekken, dat de weergegeven verdeling een kleine variabiliteit en een sterke centrale tendentie vertoont. Doet men het tegengestelde, dan suggereert men een min of meer vlakke verdeling met een grote variabiliteit. Vandaar de in 2.2.1 gegeven aanwijzing, de 'hoogte' van de figuur op ongeveer $\frac{2}{3}$ à $\frac{3}{4}$ van de 'breedte' te kiezen.

Een bijzonder misleidend beeld kan ontstaan, als men de schaalverdeling op de Y-as niet bij het nulpunt laat beginnen. Figuur 2.15 demonstreert het effect, dat hiermee kan worden bereikt. De bedoeling is een indruk te geven van de verschillen tussen de gemiddelde lengte en het gemiddelde gewicht van groepen schoolkinderen (van $6\frac{1}{2}$ t/m $13\frac{1}{2}$ jaar), die al dan niet schoolmelk gebruikten. Weliswaar vertoont de schaal op de Y-as een nulpunt, maar men bedenke dat hier feitelijk



Figuur 2.15. Verschil in *toeneming* van gewicht en lengte tussen schoolmelk en niet-schoolmelk drinkende kinderen in de jaren 1949 en 1950 ten opzichte van 1937. (Uit 'Melk op School,' Centraal Schoolmelkcommité, Den Haag).

het gemiddelde gewicht, resp. de gemiddelde lengte, van 1937 zou moeten staan en dat deze gemiddelden wel in de buurt van 130 cm, resp. 27 kg zullen liggen.

Uit de voorgaande voorbeelden blijkt wel dat men een grafiek met zorg moet samenstellen en dat men een gegeven grafiek altijd kritisch moet bekijken. Grafieken kunnen een waardevolle steun vormen bij het bestuderen van cijfermateriaal, maar zij kunnen ook zeer misleidend zijn.

2.2.7. OPGAVEN

2.4. Teken het kolommendiagram van de verdeling in tabel 2.3b.

2.5. Teken het kolommendiagram en het cumulatieve kolommendiagram van de verdelingen in: a. Opgave 2.1, b. Opgave 2.2.

2.6. Teken een spreidingsdiagram van onderstaande gegevens betreffende het lichaamsgewicht in gr (x) en het gewicht van de kam in mg (y) van 13 Leghorn kuikens:

x	85	72	69	78	69	90	88	95	92	91	75	82	76
y	54	42	29	37	18	84	56	107	90	68	31	48	41

2.7. Teken een spreidingsdiagram van de gegevens in opgave 2.3.

2.8. Teken van de gegevens in opgave 11.6 een verloopdiagram.

2.3. Gemiddelden

2.3.1. DE MODUS EN DE MODALE KLASSE

De modus van een reeks waarden is de waarde met de grootste frequentie. De modale klasse van een frequentieverdeling met klassen van verschillende breedte is de klasse, waarvan het quotiënt frequentie/klassebreedte het grootst is. Bij een frequentieverdeling met constante klassebreedte is dit dus de klasse met de grootste frequentie. Men noemt de klassewaarde (het midden) van de modale klasse van een frequentieverdeling ook wel de modus van die verdeling. Zo is de modale klasse van de verdeling in tabel 2.3b de klasse 102,5—107,5 mm Hg.

Bij een klein aantal waarnemingen komt het herhaaldelijk voor dat er geen modus is (alle waarnemingen verschillen in waarde). Ook vindt men soms meerdere modi of een modus die toevallig op een weinig karakteristieke waarde valt. Bij een frequentieverdeling is de modale klasse afhankelijk van de klasse-indeling, die tot op zekere hoogte willekeurig is. De modus wordt dan ook weinig gebruikt.

2.3.2. DE MEDIAAN

De mediaan van een oneven aantal van $2n + 1$ naar grootte geordende waarnemingen is de waarde van de $(n + 1)^e$ waarneming. Bij een even aantal van $2n$ waarnemingen is de mediaan enig getal tussen de waarde van de n^e en de $(n + 1)^e$ waarneming; in de regel kiest men hiervoor de halve som (= het rekenkundig gemiddelde) van deze twee middelste waarnemingen. Om de mediaan te bepalen moet men dus eerst alle waarden naar grootte rangschikken.

Als men van een continu verdeelde grootheid slechts over een gegroepeerde frequentieverdeling beschikt, kan men in het algemeen slechts de klasse aangeven, waar de mediaan in ligt.

Voorbeelden

2.7. De mediaan van de onderstaande 9 waarnemingen is 130:

122 124 125 127 130 130 134 140 273

2.8. De mediaan van de onderstaande 8 waarnemingen is 40,5:

2 35 38 40 ↑ 41 42 46 52

2.9. De mediaan van de verdeling in voorbeeld 1.25 ligt in de klasse met de grenzen 13,25 en 13,75. Als men de waarnemingen in deze klasse kent, kan men ze rangschikken en de mediaan bepalen. Als men aanneemt dat de waarnemingen gelijkmatig over de klasse verspreid liggen, kan men de mediaan verkrijgen door lineaire interpolatie binnen deze klasse. Men vindt dan:

$$\text{mediaan} = 13{,}25 + \frac{95-83}{46}\ 0{,}5 = 13{,}38 = 13{,}4.$$

2.10. Men kan de mediaan ook bij benadering aflezen uit de somfrequentiepolygoon. Men trekt daartoe een horizontale lijn door het punt met relatieve somfrequentie 0,5 op de Y-as (zie de stippellijn in figuur 2.6). Deze lijn snijdt de polygoon in het punt, dat loodrecht ligt boven de waarde op de X-as, waar de mediaan ligt (controleer de uitkomst, mediaan = 5,30, door berekening).

2.3.3. HET REKENKUNDIG GEMIDDELDE

Niet-gegroepeerde waarnemingen. Het rekenkundig gemiddelde van een aantal waarnemingen is de som van hun waarden, gedeeld door hun aantal. Het wordt gewoonlijk 'het gemiddelde' genoemd en waar dit niet tot misverstand aanleiding kan geven, zullen wij deze gewoonte volgen.

Betreffen de waarnemingen de elementen van een *populatie*, dan wordt hun gemiddelde aangeduid met μ (de Griekse letter mu). Nu geldt

$$\mu_x = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{\Sigma x_i}{N}, \tag{2.1}$$

waarbij N de omvang der populatie is.

Betreffen de waarnemingen de elementen in een *steekproef*, dan wordt hun gemiddelde aangeduid met $\bar{x}$ en is

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\Sigma x_i}{n}, \tag{2.2}$$

waarbij n de omvang van de steekproef is.

Voorbeeld.

2.11. Het gemiddelde van de 9 waarnemingen in voorbeeld 2.7 is:

$$\bar{x} = \frac{122 + 124 + 125 + 127 + 130 + 130 + 134 + 140 + 273}{9} = \frac{1305}{9} = 145.$$

Frequentieverdelingen. Het gemiddelde van de frequentieverdeling van de variabele x op een populatie (kortweg: het gemiddelde van de *populatieverdeling* van x) is, als deze verdeling k klassen van constante breedte heeft:

[1] Zie voor het gebruik van het somteken Σ (Griekse hoofdletter sigma) bijlage 2, par. 1.

$$(2.3) \qquad \mu_x = \frac{f_1x_1 + f_2x_2 + \ldots + f_kx_k}{N} = \frac{\sum_{i=1}^{k} f_ix_i}{N} = \frac{\Sigma f_ix_i}{N}.$$

Hierin is x_i de klassewaarde en f_i de frequentie van de klasse i. Berekent men het gemiddelde uit de relatieve frequentieverdeling dan is

$$(2.4) \qquad \mu_x = \sum_{i=1}^{k} \frac{f_i}{N} x_i = \Sigma \frac{f_i}{N} x_i$$

Het gemiddelde van de frequentieverdeling van de variabele x op een steekproef (= het gemiddelde van de *waargenomen verdeling* van x) bij k klassen van dezelfde breedte is

$$(2.5) \qquad \bar{x} = \frac{\sum_{i=1}^{k} f_ix_i}{n} = \frac{\Sigma f_ix_i}{n}.$$

Als de klassen van de frequentieverdeling gevormd worden door de waargenomen waarden zelf (zodat elke klasse uit één waarde bestaat en bij een continu verdeelde grootheid de klassebreedte gelijk is aan het afrondingsinterval), levert de berekening van het gemiddelde uit de frequentieverdeling precies dezelfde uitkomst op als de berekening uit de ongegroepeerde waarnemingen.

Als de klassen bestaan uit intervallen, die gevormd zijn door een constant aantal waarden bijeen te nemen, is dit niet het geval. Berekent men het gemiddelde uit zo'n gegroepeerde frequentieverdeling, dan wordt elke klasse vertegenwoordigd door het klassemidden x_i. Dit zal gewoonlijk niet precies gelijk zijn aan het gemiddelde van de klasse, zodat bij deze berekening foutjes worden gemaakt. Als het aantal waarnemingen niet te klein en de gekozen klasse-indeling niet te grof is, zullen de foutjes in positieve en in negatieve richting elkaar grotendeels opheffen, zodat hun som dicht bij nul ligt. Het gemiddelde van de gegroepeerde verdeling zal dan slechts weinig verschillen van het gemiddelde van de ongegroepeerde waarnemingen.

Voorbeelden

2.12. Bij onderstaande berekeningen wordt aangenomen dat de gegeven frequentieverdeling een populatie betreft:

Frequentieverdeling

x	3	4	5	6	7	8	9	Totaal
f	1	2	6	8	4	3	1	$25 = N$
fx	3	8	30	48	28	24	9	$150 = \Sigma fx$

(2.3) $\mu_x = \frac{150}{25} = 6$

Relatieve frequentieverdeling

x	3	4	5	6	7	8	9	Totaal
$\frac{f}{N}$	0,04	0,08	0,24	0,32	0,16	0,12	0,04	$1{,}00 = \Sigma\frac{f}{N}$
$\frac{f}{N} \cdot x$	0,12	0,32	1,20	1,92	1,12	0,96	0,36	$6{,}00 = \Sigma\frac{f}{N}x$

(2.4) $$\mu_x = 6$$

2.13. Berekening van het gemiddelde van de verdeling in tabel 2.3b:

x	85	90	95	100	105	110	115	120	125	130	Totaal
f	1	3	4	8	14	13	9	5	2	1	$60 = n$
fx	85	270	380	800	1470	1430	1035	600	250	130	$6450 = \Sigma fx$

(2.5) $$\bar{x} = \frac{6450}{60} = 107{,}5$$

Berekent men het gemiddelde via de oorspronkelijke waarnemingen in tabel 2.3a dan vindt men met (2.2):

$$\bar{x} = \frac{100 + 111 + 105 + \ldots + 100 + 106}{60} = \frac{6441}{60} = 107{,}35.$$

De fout, die ontstaat door het gemiddelde uit de gegroepeerde verdeling te berekenen blijkt hier 0,15 te bedragen, d.i. ca. 0,14%.

Gewogen gemiddelden. Als men beschikt over twee steekproeven, resp. steekproef 1 van de omvang n_1 met gemiddelde $\bar{x}_1$ en steekproef 2 van de omvang n_2 met gemiddelde $\bar{x}_2$, dan is het gewogen gemiddelde van alle $n = n_1 + n_2$ waarnemingen

(2.6) $$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2},$$

daar: $n_1\bar{x}_1 = \sum_{i=1}^{n_1} x_i$, $n_2\bar{x}_2 = \sum_{j=1}^{n_2} x_j$ en $\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2} x_j = \sum_{m=1}^{n} x_m$.

Analoog geldt voor k steekproeven:

(2.7) $$\bar{x} = \frac{\sum_{i=1}^{k} n_i\bar{x}_i}{\sum_{i=1}^{k} n_i} = \Sigma w_i\bar{x}_i \qquad (w_i = \frac{n_i}{n},\ n = \Sigma n_i).$$

2.4. Spreidingsmaten

2.4.1. DE SPREIDINGSBREEDTE

De spreidingsbreedte R (naar het Engelse range) van een aantal waarnemingen is het verschil tussen de hoogste en laagste waarde. Wij hebben deze maat voor de variabiliteit reeds gebruikt bij het kiezen van de klassebreedte voor een gegroepeerde frequentieverdeling. De spreidingsbreedte heeft het nadeel, dat zij slechts op twee waarnemingen berust en dat haar interpretatie afhangt van de steekproefomvang: hoe groter deze is, des te groter is in het algemeen de spreidingsbreedte.

Voorbeeld

2.14. Beschouw de waarnemingen in tabel 2.3a. De spreidingsbreedten van de 10 kolommen (ieder met 6 waarnemingen) zijn, van links naar rechts:
22 24 29 31 26 8 31 20 19 26 (gemiddelde 23,6, mediaan 25).
De spreidingsbreedten van de 6 rijen (ieder met 10 waarnemingen) zijn, van boven naar beneden:
30 25 39 42 32 25 (gemiddelde 32,2, mediaan 31).
De spreidingsbreedte van alle 60 waarnemingen is 44.

Voor $n \leqslant 10$ kan de spreidingsbreedte, die direct uit de waarnemingen kan worden afgelezen, dienen voor het schatten van de waarde van een andere spreidingsmaat, de standaardafwijking (zie 2.4.4). Deze procedure wordt bijv. bij de kwaliteitscontrole in de industrie toegepast.

2.4.2. DE GEMIDDELDE AFWIJKING

Deze spreidingsmaat wordt tegenwoordig nog maar zelden gebruikt. Zij wordt gedefinieerd als

$$GA = \frac{\Sigma \mid x_i - \bar{x} \mid}{n} \tag{2.8}$$

en is het gemiddelde van de *absolute waarden* van de afwijkingen $x_i - \bar{x}$, dat zijn de waarden van deze verschillen, zonder dat daarbij op het teken wordt gelet [merk op dat $\Sigma(x_i - \bar{x}) = 0$].

2.4.3. FRACTIELEN EN HET KWARTIELENINTERVAL

Het fractiel α van een naar grootte geordende reeks elementen is de waarde waarbeneden een fractie α van alle waarnemingen ligt. Men duidt een fractiel aan door aan het symbool van de variabele deze fractie als index toe te kennen. Zo is bijv. $x_\alpha(x_{0,10})$ het fractiel α (0,10) van de verdeling van de grootheid x (men zegt ook wel: het 10e *percentiel*, resp. het 1e *deciel*). De fractielen $x_{0,25}$ en $x_{0,75}$ noemt men het eerste, resp. het derde *kwartiel*. Het tweede kwartiel, $x_{0,50}$, is de mediaan.

Evenals bij de mediaan doet zich bij fractielen de moeilijkheid voor, dat men bij een waargenomen verdeling x_α meestal niet precies kan aangeven. Fractielen kunnen echter vrij nauwkeurig (bijv. door lineaire interpolatie) worden bepaald, als de verdeling op een groot aantal waarnemingen berust. Als maat voor de spreiding kan men dan gebruiken

$$K = \frac{x_{0,75} - x_{0,25}}{2}, \tag{2.9}$$

het *kwartieleninterval* (halve kwartielsafstand). Ook deze spreidingsmaat wordt echter zelden gebruikt.

Fractielen spelen een belangrijke rol bij het tabelleren van (theoretische) continue verdelingen. Hier zullen wij in latere hoofdstukken vele voorbeelden van ontmoeten.

2.4.4. DE VARIANTIE EN DE STANDAARDAFWIJKING

Niet-gegroepeerde waarnemingen. De *variantie* van de populatieverdeling van de grootheid x, betreffende een *populatie* van N elementen wordt gedefinieerd als

$$\sigma_x^2 = \frac{\sum_{i=1}^{N} (x_i - \mu_x)^2}{N} = \frac{\Sigma(x_i - \mu_x)^2}{N}. \tag{2.10}$$

De wortel uit de variantie, σ (de Griekse letter sigma), heet de *standaardafwijking* of *spreiding*. Uit (2.10) blijkt, dat σ_x^2 en σ_x de variabiliteit van de waarnemingen t.o.v. hun gemiddelde μ_x meten. Wij zullen later zien dat de standaardafwijking, die in dezelfde eenheden is uitgedrukt als de waarnemingen zelf, het best kan worden gebruikt voor het *beschrijven* van de variabiliteit. De variantie speelt een belangrijke rol bij mathematische manipulaties.

De variantie van een *steekproef* van n waarnemingen van de grootheid x definiëren wij als

$$s_x^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1} = \frac{\Sigma(x_i - \bar{x})^2}{n - 1}. \tag{2.11}$$

Evenals bij het gemiddelde worden de variantie van een populatie en die van een steekproef door verschillende symbolen (σ^2 tegenover s^2) van elkaar onderscheiden. Verder blijkt dat bij de berekening van s^2 door $n - 1$ wordt gedeeld (en niet door n, zoals men zou verwachten). Wij komen hier in 13.3 op terug, maar willen nu reeds twee opmerkingen maken. De eerste betreft het geval dat de 'steekproef' uit één

waarneming bestaat. Deelt men dan bij de berekening van de variantie door 1, dan is de uitkomst $0/1 = 0$. Het is duidelijk dat dit in het algemeen een onjuiste schatting is van de populatievariantie. Deelt men echter door $n - 1 = 0$, dan is de uitkomst onbepaald (0/0), wat uitdrukt dat deze steekproef geen informatie omtrent de populatievariantie verschaft. Verder wijzen wij erop, dat de berekende n verschillen $x_i - \bar{x}$ niet alle onafhankelijk zijn, daar $\Sigma(x_i - \bar{x}) = 0$. Dientengevolge staat het n-de verschil vast, wanneer $n - 1$ verschillen zijn berekend. De noemer van formule (2.11) geeft dus het aantal onderling onafhankelijke waarden aan, waarop de teller, de *kwadratensom* $\Sigma(x_i - \bar{x})^2$, berust. Men noemt dit aantal het *aantal vrijheidsgraden* van de variantie (zie ook 9.2.3).

Voorbeeld

2.15. Voor onderstaande steekproeven van 6 waarnemingen zijn s^2 en s berekend.

Steekproef 1			Steekproef 2		
x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	−3	9	0	−4	16
3	−1	1	3	−1	1
4	0	0	4	0	0
5	+1	1	5	+1	1
5	+1	1	5	+1	1
6	+2	4	7	+3	9
24	0	16	24	0	28
Σx_i	$\Sigma(x_i - \bar{x})$	$\Sigma(x_i - \bar{x})^2$	Σx_i	$\Sigma(x_i - \bar{x})$	$\Sigma(x_i - \bar{x})^2$
$s_1^2 = \frac{16}{6-1} = 3{,}2$		$s_1 = 1{,}79$	$s_2^2 = \frac{28}{6-1} = 5{,}6$		$s_2 = 2{,}37$

De formules (2.10) en (2.11) kunnen worden herleid tot een vorm, die geschikter is voor het uitvoeren van berekeningen. Immers (zie bijlage 2, par. 1):

$$\Sigma(x_i - \mu)^2 = \Sigma x_i^2 - 2\mu \cdot \Sigma x_i + N\mu^2.$$

Daar $\mu = \Sigma x_i/N$ en $\mu N = \Sigma x_i$, volgt dat

$$\Sigma(x_i - \mu)^2 = \Sigma x_i^2 - 2\left\{\frac{\Sigma x_i}{N}\right\}\Sigma x_i + N\left\{\frac{\Sigma x_i}{N}\right\}^2 = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{N}$$

of

$$\Sigma(x_i - \mu)^2 = \Sigma x_i^2 - 2\mu\,(\mu N) + N\mu^2 = \Sigma x_i^2 - N\mu^2.$$

Men kan voor (2.10) ook schrijven

$$\sigma_x^2 = \frac{\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{N}}{N} = \frac{\Sigma x_i^2}{N} - \mu_x^2. \tag{2.12}$$

Op analoge wijze volgt uit (2.11)

$$s_x^2 = \frac{\Sigma x_i^2 - \frac{(\Sigma x_i)^2}{n}}{n-1}. \tag{2.13}$$

Bij het gebruik van deze formules behoeft men dus niet meer alle verschillen $x_i - \mu$, resp. $x_i - \bar{x}$, te berekenen, maar slechts de som van de waarnemingen (Σx_i) en de som van de kwadraten van de waarnemingen (Σx_i^2). Tevens vermijdt men hierdoor foutjes die ontstaan, als men bij het gebruik van (2.10) of (2.11) met een afgeronde waarde van μ_x, resp. van $\bar{x}$, werkt.

Voorbeeld

2.16. Past men (2.13) toe op de waarnemingen in steekproef 1 van voorbeeld 2.15, dan is:
$\Sigma x_i = 24$, $(\Sigma x_i)^2/n = (24)^2/6 = 96$, $\Sigma x_i^2 = (1)^2 + (3)^2 \ldots + (6)^2 = 112$,
zodat

$$s^2 = \frac{112 - 96}{5} = \frac{16}{5} = 3{,}2.$$

Frequentieverdelingen. De variantie van de populatieverdeling van de grootheid x (met k klassen van constante breedte) is

$$\sigma_x^2 = \frac{f_1(x_1 - \mu_x)^2 + f_2(x_2 - \mu_x)^2 + \ldots + f_k(x_k - \mu_x)^2}{N} = \frac{\Sigma f_i(x_i - \mu_x)^2}{N}. \tag{2.14}$$

Voor berekening van de populatievariantie kan men gebruik maken van de formule

$$\sigma_x^2 = \frac{\Sigma f_i x_i^2 - \frac{(\Sigma f_i x_i)^2}{N}}{N} = \frac{\Sigma f_i x_i^2}{N} - \mu_x^2. \tag{2.15}$$

Berekent men deze variantie uit de relatieve frequentieverdeling, dan geldt

$$\sigma_x^2 = \Sigma \frac{f_i}{N} \cdot x_i^2 - \mu_x^2. \tag{2.16}$$

Voorbeeld

2.17. Onderstaande berekeningen betreffen de verdeling in voorbeeld 2.12.

Frequentieverdeling

x	3	4	5	6	7	8	9	Totaal
f	1	2	6	8	4	3	1	$25 = N$
fx	3	8	30	48	28	24	9	$150 = \Sigma fx$
fx^2	9	32	150	288	196	192	81	$948 = \Sigma fx^2$

$$\sigma_x^2 = \frac{948 - \frac{(150)^2}{25}}{25} = 1{,}92$$

$$\sigma_x^2 = \frac{948}{25} - (6)^2 = 1{,}92$$

Relatieve frequentieverdeling

x	3	4	5	6	7	8	9	Totaal
$\frac{f}{N}$	0,04	0,08	0,24	0,32	0,16	0,12	0,04	$\Sigma\frac{f}{N} = 1$
$\frac{f}{N}x$	0,12	0,32	1,20	1,92	1,12	0,96	0,36	$\Sigma\frac{f}{N}x = 6$
$\frac{f}{N}x^2$	0,36	1,28	6,00	11,52	7,84	7,68	3,24	$\Sigma\frac{f}{N}x^2 = 37{,}92$

$$\sigma_x^2 = 37{,}92 - (6)^2 = 1{,}92$$

De variantie s_x^2 van de waargenomen verdeling van de grootheid x (met k klassen van constante breedte) is

$$s_x^2 = \frac{\Sigma f_i(x_i - \bar{x})^2}{n-1} = \frac{\Sigma f_i x_i^2 - \frac{(\Sigma f_i x_i)^2}{n}}{n-1}. \tag{2.17}$$

Zie voor een voorbeeld waarbij (2.17) wordt toegepast tabel 2.7 (linkerdeel).

2.4.5. OPGAVEN

2.9. Gegeven is de volgende steekproef:
3 2 1 1 1 3 1 4 5 1 2 4 3 2 3 5 4 2 5 1 2
Bepaal (zo mogelijk): a. de modus, b. de mediaan, c. het gemiddelde, d. de variantie, e. de spreiding.

2.10. Als opgave 2.9, maar voor de steekproef:
4 5 3 4 4 6 1 6 4 2 5 4 5 3 7 4 3 5 4 6 2 5 1 3 5 4 3 6 4 2 4

2.11. Bereken het gemiddelde en de spreiding van onderstaande relatieve populatieverdeling:

x	10	11	12	13	14	15	16	17	18	19	20
f/N	0,01	0,04	0,08	0,11	0,15	0,20	0,18	0,11	0,07	0,03	0,02

2.12. Bereken het gemiddelde en de variantie van de n gehele getallen 1, 2, 3, ..., n. Bedenk hierbij dat de som van deze getallen gelijk is aan $n(n+1)/2$ en dat de som van hun kwadraten $n(n+1)(2n+1)/6$ bedraagt.

2.5. Vereenvoudigde berekening van het gemiddelde en de standaardafwijking

2.5.1. ENKELE EIGENSCHAPPEN

Als $v_i = x_i - a$ (a is een constante) dan is

Eigenschap 1: $\bar{v} = \bar{x} - a$.

$$\text{Omdat } \bar{v} = \frac{\Sigma(x_i - a)}{n} = \frac{\Sigma x_i - na}{n} = \frac{\Sigma x_i}{n} - \frac{na}{n} = \bar{x} - a.$$

Eigenschap 2: $s_v^2 = s_x^2$, zodat $s_v = s_x$.

Want $\Sigma(v_i - \bar{v})^2 = \Sigma[(x_i - a) - (\bar{x} - a)]^2 = \Sigma(x_i - \bar{x})^2$.

Als $w_i = x_i/b$ (b is een constante) dan is

Eigenschap 3: $\bar{w} = \bar{x}/b$.

Omdat $\bar{w} = \dfrac{\Sigma x_i/b}{n} = \dfrac{\Sigma x_i}{bn} = \dfrac{\bar{x}}{b}$.

Eigenschap 4: $s_w^2 = s_x^2/b^2$, $s_w = s_x/b$.

Want $\Sigma(w_i - \bar{w})^2 = \Sigma(x_i/b - \bar{x}/b)^2 = \Sigma\left(\dfrac{x_i - \bar{x}}{b}\right)^2 = \dfrac{\Sigma(x_i - \bar{x})^2}{b^2}$.

Uit de voorgaande eigenschappen volgt dat voor de grootheid

$z_i = \dfrac{x_i - a}{b}$ geldt

Eigenschap 5: $\bar{z} = \dfrac{\bar{x} - a}{b}$, zodat $\bar{x} = a + b\bar{z}$.

Eigenschap 6: $s_z^2 = s_x^2/b^2$, $s_z = s_x/b$, zodat $s_x^2 = b^2 s_z^2$ en $s_x = b s_z$.

2.5.2. BEREKENINGSMETHODEN

Als men het gemiddelde en de spreiding van n niet-gegroepeerde waarnemingen wil berekenen, kunnen de eigenschappen 1 en 2 vaak worden toegepast ter vereenvoudiging van het rekenwerk. Bij het voorbeeld in onderstaande tabel is van de waarden van x de constante 1220 afgetrokken. De berekeningen zijn uitgevoerd met de waarden van $v = x - 1220$, zodat het kwadrateren vrijwel uit het hoofd kan geschieden:

x	1220	1218	1205	1226	1220	1214	1236	1221	1238	1225	1221	1232	Totaal
v	0	−2	−15	6	0	−6	16	1	18	5	1	12	$\Sigma v = 36$
v^2	0	4	225	36	0	36	256	1	324	25	1	144	$\Sigma v^2 = 1052$

Nu is $\bar{v} = 36/12 = 3$, zodat $\bar{x} = 3 + 1220 = 1223$. Verder vindt men:

$$s_v^2 = s_x^2 = \frac{1052 - (36)^2/12}{11} = 85{,}8182, \quad s_x = \sqrt{85{,}8182} = 9{,}26.$$

Beschikt men over een frequentieverdeling met constante klassebreedte c dan kan een hulpgrootheid u worden ingevoerd, waarvan het nulpunt ligt bij een centraal gelegen klassewaarde x_0 en die in iedere hogere klasse met 1 opklimt, resp. in iedere lagere klasse met 1 afdaalt. Tussen de grootheden x en u bestaat dan de relatie

$$u_i = \frac{x_i - x_0}{c} \qquad \text{(zodat } x_i = x_0 + cu_i\text{)}.$$

Uit de eigenschappen (5) en (6) volgt

$$\bar{u} = \frac{\bar{x} - x_0}{c} \qquad \text{(zodat } \bar{x} = x_0 + c\bar{u}\text{)}, \tag{2.18}$$

en

$$s_u = \frac{s_x}{c} \qquad \text{(zodat } s_x = cs_u \text{ en } s_x^2 = c^2 s_u^2\text{)}. \tag{2.19}$$

Als men het gemiddelde en de spreiding van de hulpgrootheid u heeft berekend, kunnen x en s_x direct worden bepaald. Tabel 2.7 toont aan dat door deze procedure het rekenwerk aanzienlijk wordt vereenvoudigd (vergelijk bijv. de kolommen fx^2 en fu^2). De keuze van x_0 bij een centraal gelegen klassewaarde biedt het voordeel, dat zo klein mogelijke producten fu en fu^2 ontstaan, daar de frequentie afneemt naarmate u numeriek groter wordt. Men kan x_0 echter ook elders kiezen.

Tabel 2.7. Berekening van het gemiddelde en de spreiding van een frequentieverdeling met constante klassebreedte, zonder en met gebruik van een hulpgrootheid (verdeling uit tabel 2.2b)

Klassewaarde x	Frequentie f	Berekening fx	Berekening $fx^2 = x \cdot fx$	Berekening met hulpgrootheid u	fu	fu^2
6,25	1	6,25	39,0625	+5	+ 5	25
6,05	1	6,05	36,6025	+4	+ 4	16
5,85	3	17,55	102,6675	+3	+ 9	27
5,65	7	39,55	223,4575	+2	+14	28
5,45	10	54,50	297,0250	+1	+10	10
5,25	12	63,00	330,7500	0	0	0 ($x_0 = 5{,}25$)
5,05	8	40,40	204,0200	—1	— 8	8
4,85	5	24,25	117,6125	—2	—10	20
4,65	2	9,30	43,2450	—3	— 6	18
4,45	1	4,45	19,8025	—4	— 4	16
Totaal	50 $\Sigma f = n$	265,30 Σfx	1414,2450 Σfx^2	$c = 0{,}2$	+14 Σfu	168 Σfu^2

$$\bar{x} = \frac{\Sigma fx}{n} = \frac{265{,}30}{50} = 5{,}31$$

$$s^2 = \frac{\Sigma fx^2 - \frac{(\Sigma fx)^2}{n}}{n-1} = \frac{1414{,}245 - \frac{(265{,}3)^2}{50}}{49}$$

$$= \frac{1414{,}2450 - 1407{,}6818}{49} = 0{,}1339$$

$$s = \sqrt{0{,}1339} = 0{,}366 = 0{,}37$$

$$\bar{u} = \frac{\Sigma fu}{n} = \frac{14}{50} = 0{,}28$$

$$\bar{x} = x_0 + c\bar{u} = 5{,}25 + 0{,}2(0{,}28) = 5{,}31$$

$$s_u^2 = \frac{\Sigma fu^2 - \frac{(\Sigma fu)^2}{n}}{n-1} = \frac{168 - \frac{(14)^2}{50}}{49}$$

$$= 3{,}3486 \quad s_u = 1{,}83$$

$$s_x = cs_u = 0{,}2\ (1{,}83) = 0{,}366 = 0{,}37$$

2.5.3. OPGAVEN

2.13. Uit het examen statistisch analyst 1956.
Een steekproef bestaat uit 11 waarnemingen. Kunt u aangeven, hoe de waarden van het gemiddelde, de mediaan en de standaardafwijking veranderen, indien achtereenvolgens: *a.* Alle waarnemingen met 10 worden verhoogd, *b.* Daarna alle waarnemingen met 2 worden vermenigvuldigd, *c.* Daarna de op twee na grootste waarneming met 11 wordt verhoogd. Indien u niet over voldoende gegevens beschikt om alle vragen te kunnen beantwoorden, wordt gevraagd aan te geven welk additioneel gegeven het vraagstuk oplosbaar zou maken.

2.14. Bepaal van onderstaande waarnemingen de mediaan en bereken $\bar{x}$ en s:

388	389	393	388	391	390	387	396	392	389	386	388	383
398	388	403	378	385	387	384	387	383	381	382	374	

2.15. Bepaal van de waarnemingen in opgave 2.2 de mediaan en het gemiddelde.

2.16. Bereken het gemiddelde en de spreiding van de waarnemingen in opgave 2.1: *a.* Uit de waarnemingen zelf, *b.* Uit de door u opgestelde frequentieverdeling.

2.17. Bereken het gemiddelde en de spreiding van de frequentieverdeling in tabel 2.3b.

2.18. Bereken het gemiddelde en de standaardafwijking van de drie frequentieverdelingen in tabel 2.8.

2.6. Karakterisering van een frequentieverdeling

Het gebruik van de in 2.4 en 2.5 behandelde kengetallen hangt af van de vorm van de frequentieverdeling van de bestudeerde grootheid en het doel van de karakterisering. Wij bespreken daarom eerst enkele typen van veel voorkomende frequentieverdelingen en vervolgens het gebruik van de kengetallen.

2.6.1. ENKELE VERDELINGSTYPEN

De frequentieverdeling van vele grootheden blijkt voor een populatie (of een grote steekproef daaruit) ééntoppig te zijn. De frequentie van één, min of meer centraal gelegen, klasse (de modale klasse) is hoger dan die der andere klassen en aan weerszijden van deze klasse zijn de frequenties lager, naarmate de klassen verder van dit centrum verwijderd zijn.

Deze ééntoppigheid is verklaarbaar[1]. Beschouwt men de wijze waarop de waarde van een bepaalde grootheid bij een element van een populatie tot stand is gekomen, dan blijkt dat meestal een klein aantal algemene factoren is opgetreden. Deze factoren hebben elk element op (vrijwel) dezelfde wijze beïnvloed. Zouden zij alleen aanwezig zijn geweest dan zouden zij tot (vrijwel) dezelfde waarde voor alle elementen hebben geleid. Naast deze algemene of primaire factoren

[1] Zie voor een uitgebreidere beschouwing KAPTEYN (1903), KAPTEYN en VAN UVEN (1916), DRION (1936), BAAS BECKING en DRION (1936) en VAN ITERSON (1950).

treden echter — gewoonlijk vele — andere factoren op, die geheel of grotendeels onderling onafhankelijk zijn en die *niet* op alle elementen op dezelfde wijze inwerken. Naar gelang van de omstandigheden zal elk van deze secundaire factoren het algemene resultaat kunnen verlagen of verhogen. Het zal echter slechts zelden voorkomen dat zij alle in dezelfde richting werken; meestal zullen sommige een negatieve en andere een positieve invloed uitoefenen, zodat het resulterende totale effect numeriek klein zal zijn. Soms zullen de positieve of de negatieve effecten duidelijk overheersen, zodat het algemene resultaat sterker wordt beïnvloed, enz. Bestaat een populatie uit een groot aantal elementen, dan zal men er dus betrekkelijk veel aantreffen met een waarde in de nabijheid van de door de primaire factoren tot stand gekomen uitkomst en daar zal zich dan het centrum van de verdeling bevinden. Hoe verder men zich van dit centrum verwijdert, des te kleiner zal het aantal waarden zijn.

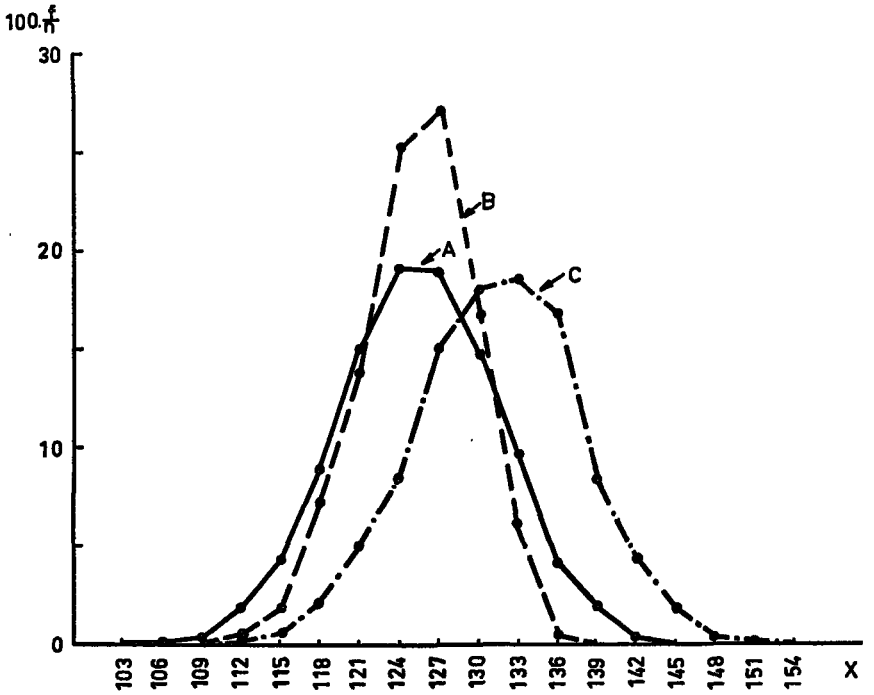
Het voorgaande verklaart ook hoe bijv. een *tweetoppige* verdeling kan optreden. Dit zal gebeuren als niet alle elementen van de populatie aan dezelfde algemene factoren onderhevig zijn. De uitkomsten concentreren zich dan om twee waarden en als deze niet te dicht bijeen liggen zullen zich twee toppen vormen (liggen zij dichter bijeen, dan verkrijgt men een verdeling met een vlak middenstuk). Men zegt dan wel dat de populatie heterogeen is. Slaagt men erin de beide soorten elementen van elkaar te onderscheiden, dan ontstaan twee homogene deelpopulaties die ieder een ééntoppige verdeling bezitten. Door te zoeken naar een criterium, dat ons in staat stelt een tweetoppige verdeling in twee ééntoppige te splitsen, komt men soms belangrijke eigenschappen van een populatie op het spoor.

Uit het voorgaande mag men niet afleiden, dat iedere ééntoppige verdeling op een homogene populatie betrekking heeft. Het is immers mogelijk dat men te doen heeft met een heterogene populatie, waarvan de toppen der deelpopulaties samenvallen of zo dicht bijeen liggen, dat ze in totaal slechts één top opleveren.

Sommige ééntoppige verdelingen zijn ongeveer symmetrisch t.o.v. hun centrum en vertonen een klokvorm. Zij gelijken op een continue theoretische verdeling, die in hoofdstuk 5 uitvoeriger wordt behandeld, een *normale verdeling* (zie de figuren 5.3 en 5.4). Men noemt dit type verdelingen *pseudo-normaal* en enkele voorbeelden ervan zijn: *a.* De frequentieverdelingen van de lichaamslengten van jongens in tabel 2.8 en figuur 2.16. *b.* De verdelingen van de diameters van erythrocyten in tabel 22.3. *c.* De verdeling van hemoglobinegehalten in tabel 10.1.

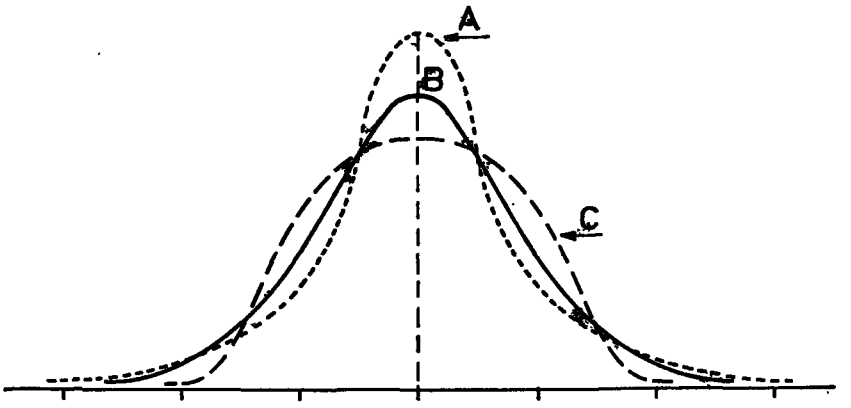
Tabel 2.8. Pseudo-normale verdelingen: lichaamslengten van jongens in cm

Lengte in cm x	A		B		C	
	f	$100\,f/n$	f	$100\,f/n$	f	$100\,f/n$
151					1	0,1
148					2	0,3
145	2	0,2			13	1,8
142	3	0,3	1	0,1	32	4,4
139	17	1,9	2	0,2	60	8,3
136	38	4,2	7	0,6	121	16,8
133	87	9,7	68	6,2	134	18,6
130	132	14,7	184	16,8	130	18,1
127	171	19,0	298	27,2	109	15,1
124	173	19,2	277	25,3	62	8,6
121	135	15,0	154	14,1	35	4,9
118	80	8,9	79	7,2	15	2,1
115	40	4,4	20	1,8	4	0,6
112	17	1,9	5	0,4	2	0,3
109	4	0,4	1	0,1		
106	1	0,1				
Totaal	900	99,9	1096	100,0	720	100,0
$\bar{x}$	125,5		125,4		131,4	
s_x	6,10		4,41		6,17	



Figuur 2.16. Frequentiepolygonen van de drie pseudonormale verdelingen in tabel 2.8.

Niet alle ééntoppige en ongeveer symmetrische verdelingen zijn echter pseudo-normaal. Zij kunnen ook te 'spits' zijn met een te grote concentratie bij het centrum (leptokurtosis) of te 'vlak' met te grote frequenties in de staarten (platykurtosis), zoals in figuur 2.17 is getekend.



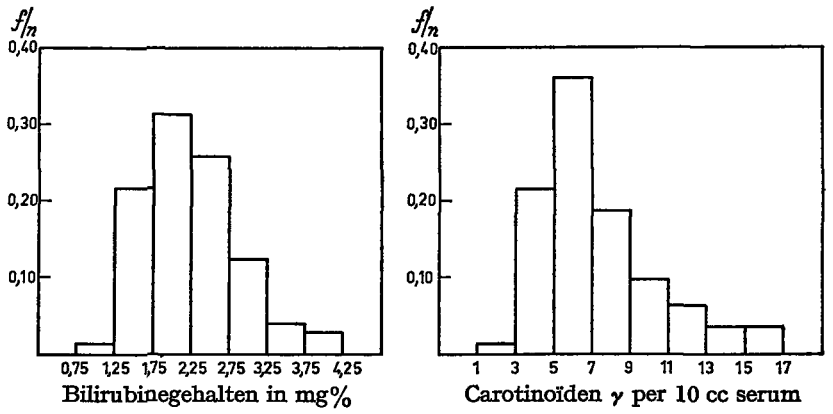
Figuur 2.17. Kurtosis: A = leptocurve (spitse kromme), B = normale kromme, C = platycurve (platte kromme).

Tabel 2.9 geeft een voorbeeld van een redelijk symmetrische, maar te spitse verdeling.

Tabel 2.9. Verdeling van de geboortegewichten (in hg) van 9651 pasgeboren jongens (uit de cursus 'Summarizing Experimental Data', Columbia University, U.S.A.)

x	5-8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40	41-44	45-48	49-52	53-56	57-60	61-64	Totaal
f	5	39	98	136	285	897	2240	2727	2051	770	310	73	17	2	1	9651

Andere ééntoppige verdelingen zijn asymmetrisch: de frequenties aan de hoge (d.i. grafisch gezien aan de rechter) zijde van de top nemen langzamer af dan die aan de lage (linker) zijde. Men spreekt dan van *scheefheid naar rechts* of van positieve scheefheid. Het tegengestelde geval, scheefheid naar links, blijkt in de natuur minder vaak voor te komen. Het overheersen van scheefheid naar rechts is begrijpelijk als men bedenkt dat vele grootheden (zoals bezinkingssnelheid der erythrocyten, serum bilirubinegehalte, vitamine C in urine, systolische bloeddruk) geen negatieve waarden kunnen aannemen, omdat het getal nul een natuurlijke grens naar links vormt. Figuur 2.18 geeft enkele voorbeelden van naar rechts scheve verdelingen. Tabel 2.10 geeft een voorbeeld van een naar links scheve verdeling.



Figuur 2.18. Scheefheid naar rechts.

Tabel 2.10. Frequentieverdelingen van de duur van de zwangerschap in weken, A = primiparae, B = multiparae. Gegevens: Kweekschool voor Vroedvrouwen, Amsterdam

Aantal weken	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	Totaal
$f(A)$	3	1	2	5	2	4	9	13	44	84	245	348	596	418	272	74	25	6	6	2157
$f(B)$	1	0	1	3	3	6	7	14	41	85	159	308	556	350	208	62	22	2	5	1833

Scheefheid naar rechts zal ook optreden, wanneer men weliswaar te doen heeft met een pseudo-normaal verdeelde grootheid x, maar door omstandigheden een functie van de vorm x^2 of ax^3 meet. Zijn bijv. de diameters van een groot aantal knikkers bij benadering normaal verdeeld, dan volgen de gewichten een naar rechts scheve verdeling.

In hoofdstuk 4 bespreken wij enkele discrete theoretische verdelingstypen, t.w. de hypergeometrische en binomiale verdelingen, die scheef naar links, symmetrisch en scheef naar rechts kunnen zijn en de POISSON-verdelingen, die een — vaak zeer sterke — scheefheid naar rechts vertonen.

Verschillende naar rechts scheve verdelingen zijn van het *log-normale type*, d.w.z. de *logarithmen* van de waarden van de grootheid zijn bij benadering normaal verdeeld. In de regel zijn in de biologie de logarithmen 'normaler' verdeeld dan de 'gewone' waarden, maar zoals uit tabel 2.11 blijkt is het verschil tussen de logarithmische en de arithmetrische schaal alleen merkbaar, als de spreidingsbreedte groot is t.o.v. de gemiddelde waarde. Zodoende zien wij dat bij A de afstanden tussen de logarithmen sterk afnemen, terwijl bij B de afneming slechts gering is.

Een indicatie tegen het nemen van logarithmen is het (mogelijk) voorkomen van de waarde 0. Men kan dan echter een kunstgreep toe-

Tabel 2.11. Twee getallenreeksen met hun logarithmen en de verschillen tussen deze logarithmen

A	Getallen	1	2	3	4	5	6
	Logarithmen	0,000	0,301	0,477	0,602	0,699	0,778
	Verschillen		0,301	0,176	0,125	0,097	0,079
B	Getallen	101	102	103	104	105	106
	Logarithmen	2,0043	2,0086	2,0128	2,0170	2,0212	2,0253
	Verschillen		0,0043	0,0042	0,0042	0,0042	0,0041

passen door de logarithme van $(x + 1)$ in plaats van de logarithme van x te nemen.

Een voorbeeld levert de verdeling van de bezinkingssnelheid van de erythrocyten in opgave 2.2. Deze verdeling is uitgesproken scheef naar rechts, maar als men van elke waarde de logarithme neemt en de frequentieverdeling van de logarithmen opstelt, verkrijgt men de pseudo-normale verdeling in tabel 2.12.

Tabel 2.12. Frequentieverdeling van de logarithmen van de bezinkingssnelheden in opgave 2.2

log x	0,15 t/a 0,45	0,45 t/a 0,60	0,60 t/a 0,75	0,75 t/a 0,90	0,90 t/a 1,05	1,05 t/a 1,20	1,20 t/a 1,35	1,35 t/a 1,50	1,50 t/a 1,65
f	2	5	8	12	15	12	7	6	1

2.6.2. KARAKTERISERING VAN PSEUDO-NORMALE VERDELINGEN

Zoals wij in hoofdstuk 5 zullen zien wordt een normaal verdeelde grootheid x volledig gekarakteriseerd door haar gemiddelde en haar standaardafwijking. Men kan bijv. voor elk interval van de vorm $\mu \pm k\sigma$ exact aangeven welke fractie van de waarden van x in dit interval ligt. In tabel 2.13 zijn enkele intervallen van deze vorm en de bijbehorende fracties gegeven (zij worden uitvoerig behandeld in 5.3.3). Uit deze tabel blijkt onder meer, dat *praktisch* alle waarden van een normaal verdeelde grootheid binnen het interval $\mu \pm 3{,}3\sigma$ liggen.

Tabel 2.13. De fractie van de waarden van een normaal verdeelde grootheid die in het interval $\mu \pm k\sigma$ ligt

k exact	k afgerond	fractie in interval $\mu \pm k\sigma$
1,000	1,0	0,67
1,645	1,65	0,90
1,960	2,0	0,95
2,576	2,6	0,99
3,291	3,3	0,999

Heeft men in een populatie (of, wat vaker voorkomt: in een steekproef) een pseudo-normale verdeling waargenomen en karakteriseert men deze door haar gemiddelde en haar standaardafwijking te geven,

dan gaat daarbij informatie verloren omtrent de optredende afwijkingen van de normaalvorm. Zijn deze afwijkingen echter klein, dan is dit verlies aan informatie ook klein en praktisch van weinig betekenis. Het gemiddelde en de standaardafwijking kunnen de verdeling dan bevredigend beschrijven, daar men van elk interval $\mu \pm k\sigma$ (resp. $\bar{x} \pm ks$) bij benadering weet welke fractie van de waarnemingen het bevat.

De verdelingen in tabel 2.8 kunnen als voorbeeld dienen. Zij zijn alle pseudo-normaal en hun gemiddelden en spreidingen zijn gegeven. Wij zien dat bij verdeling A het interval $\bar{x} \pm 3{,}3s$, d.i. het interval van $125{,}5 - 3{,}3\,(6{,}10) = 105{,}4$ tot $125{,}5 + 3{,}3\,(6{,}10) = 145{,}6$ inderdaad alle waarnemingen blijkt te bevatten. Voor verdeling C zou het interval $131{,}4 \pm 2\,(6{,}17)$, d.i. van ca. 119 tot 144, 95% der waarnemingen moeten bevatten. Hoewel deze uitkomst door de groepering niet precies te controleren is blijkt dit goed te kloppen, daar 683 van de 720 waarden, d.i. 94,9% in het interval 119,5 tot 143,5 liggen.

De karakterisering van pseudo-normale verdelingen door middel van hun gemiddelde en standaardafwijking verschaft direct een inzicht in de belangrijke verschillen tussen deze verdelingen. Zo zien wij voor de verdelingen in tabel 2.8, dat A en B vrijwel hetzelfde gemiddelde bezitten, maar dat het gemiddelde van C duidelijk (ca. 6 cm) hoger ligt. De verdelingen A en C bezitten ongeveer dezelfde spreiding, doch verdeling B vertoont een veel grotere variabiliteit.[1]

2.6.3. KARAKTERISERING VAN ANDERE VERDELINGSTYPEN

Ook voor niet-normale verdelingen kan men op grond van het gemiddelde en de standaardafwijking tot uitspraken komen omtrent het percentage, dat binnen of buiten bepaalde grenzen ligt. Voor een willekeurig verdeelde grootheid x met gemiddelde μ en spreiding σ geldt dan de ongelijkheid van TCHEBYCHEFF die luidt: buiten het interval $\mu \pm k\sigma$ $(k > 1)$ ligt *ten hoogste* een fractie $1/k^2$ van de waarden van x. Zo ligt dus altijd, onafhankelijk van de vorm van de verdeling, buiten het interval $\mu \pm 2\sigma$ ten hoogste een fractie $(\frac{1}{2})^2 = 0{,}25$ (25%) van de waarden en buiten het interval $\mu \pm 3\sigma$ ten hoogste een fractie $(\frac{1}{3})^2 = 1/9$ van de waarden. Ten hoogste — want de werkelijk buiten deze grenzen liggende fractie kan veel kleiner zijn. Indien de verdeling van de grootheid x ééntoppig en symmetrisch is geldt de ongelijkheid van CAMP en MEIDELL, die inhoudt dat buiten het interval $\mu \pm k\sigma$ ten hoogste een fractie $(2/3k)^2$ van de waarden van x ligt. In dit geval ligt bijv. buiten het interval $\mu \pm 2\sigma$ ten hoogste

[1] Wegens de grote omvang van de steekproeven en de duidelijke verschillen en overeenkomsten tussen de gemiddelden en standaardafwijkingen kan men deze conclusies wel zonder meer trekken. Zie ook 2.6.4.

een fractie van $(2/6)^2 = 1/9$ van de waarden en deze ongelijkheid blijkt dus reeds aanzienlijker scherper te zijn dan die van TCHEBYCHEFF.

Beide ongelijkheden zijn voornamelijk van theoretisch belang want niet-normale verdelingen kan men met een gemiddelde en een standaardafwijking slechts onvolledig en dus onbevredigend karakteriseren. Bij ééntoppige verdelingen kan men nog wel andere kengetallen toevoegen, die de scheefheid en of de kurtosis karakteriseren. Deze getallen zijn echter minder eenvoudig te interpreteren en te vergelijken dan gemiddelden en standaardafwijkingen. Wij bespreken ze in hoofdstuk 23 bij de methoden volgens welke men een waargenomen frequentieverdeling op normaliteit kan toetsen.

Bij ééntoppige scheve verdelingen kan men de centrale tendentie nog wel karakteriseren door het gemiddelde en de mediaan te geven. Als het aantal waarnemingen niet te klein is kan men, vooral indien uitbijters optreden, de verdeling redelijk beschrijven door een reeks fractielen (bijv. $x_{0,05}$, $x_{0,10}$, $x_{0,25}$, $x_{0,50}$ = de mediaan, $x_{0,75}$, $x_{0,90}$, $x_{0,95}$) te vermelden. Deze fractielen zeggen iets omtrent centrale tendentie én spreiding. Het vermelden van één gemiddelde en één spreidingsmaat kan misleidend zijn. HUFF (1954) geeft hier enkele sprekende voorbeelden van.

Uit het voorgaande blijkt wel, dat het niet eenvoudig is om niet-normale verdelingen op ondubbelzinnige wijze te karakteriseren. Men kan dergelijke verdelingen gewoonlijk beter in hun geheel geven, zodat ieder hun karakter zelf beoordelen kan.

2.6.4. POPULATIES EN STEEKPROEVEN

Men kan kengetallen zowel voor het karakteriseren van populaties als van steekproeven gebruiken. Men noemt een kengetal gebaseerd op de frequentieverdeling van een populatie een *parameter*. Een kengetal berekend uit de frequentieverdeling van een steekproef wordt een *steekproeffunctie* (of statistische grootheid) genoemd. Gewoonlijk gebruikt men als symbool voor een parameter een Griekse letter of een Latijnse hoofdletter, terwijl een steekproeffunctie met een kleine Latijnse letter wordt aangeduid. Dit onderscheid hebben wij reeds bij het gemiddelde (μ vs. $\bar{x}$) en de standaardafwijking (σ vs. s) leren kennen.

De parameters van een populatieverdeling zijn bepaalde vaste waarden, die de (constante) populatiesamenstelling karakteriseren. Omdat men echter uit één populatie vele verschillende steekproeven kan trekken kan de uit deze steekproeven berekende steekproeffunctie verschillende waarden aannemen.

In de volgende hoofdstukken zullen aan de variabiliteit van verschillende steekproeffuncties (bij het trekken van aselecte steekproeven) uitvoerige besprekingen worden gewijd. Wij willen er hier echter uitdrukkelijk op wijzen, dat men bij het trekken van con-

clusies op basis van de uitkomsten van een steekproef, deze variabiliteit niet mag verwaarlozen. Men kan dan ook in het algemeen kengetallen berustend op steekproeven, zoals steekproefgemiddelden, niet alleen 'op het oog' interpreteren of met elkaar vergelijken, maar dient daarbij gebruik te maken van de schattings- en toetsingsmethoden die de mathematische statistiek voor dit doel heeft ontwikkeld.

2.7. Kengetallen voor sterfte en ziekte

In 1.5.2 is reeds vermeld dat er voor het karakteriseren van sterfte- en ziektecijfers verschillende kengetallen zijn. Wij volstaan ermee deze kengetallen in te leiden, met verwijzing naar de literatuur waarin zij volledig worden besproken.

2.7.1. KENGETALLEN VOOR STERFTE

Het *ruwe sterftecijfer* (Eng. crude death rate) per 1000 inwoners van een bepaalde bevolking in een bepaald jaar wordt gedefinieerd als

$$1000 \cdot \frac{\text{aantal sterfgevallen}}{\text{gemiddelde bevolkingsomvang}} . \tag{2.20}$$

Voor de gemiddelde bevolkingsomvang neemt men het gemiddelde van het aantal inwoners op 1 januari en 31 december van het jaar. Berekent men het cijfer over een langere periode dan één jaar, dan moet de uitkomst door het aantal jaren van deze periode worden gedeeld.

Zoals de naam reeds zegt geeft het ruwe sterftecijfer slechts een zeer grove informatie over de sterfte. Het is zelfs voor een oriënterende vergelijking van de sterfte in verschillende jaren of landen ongeschikt, omdat geen rekening wordt gehouden met de verschillen in opbouw van de bevolkingen.

Men spreekt van *specifieke sterftecijfers* (Eng. specified death rates) wanneer men de sterfte (resp. de sterfte naar een bepaalde doodsoorzaak) voor bepaalde lagen van de bevolking beschrijft. Daar zowel mannen en vrouwen als personen van verschillende leeftijdklassen een verschillende sterfte (kunnen) vertonen, zal men in ieder geval steeds de sterfte per geslacht en per leeftijdsklasse beschouwen. Overeenkomstig (2.20) neemt men voor de berekening van een specifiek sterftecijfer per 1000 personen van een bepaalde laag van de bevolking in een bepaald jaar

$$1000 \cdot \frac{\text{aantal sterfgevallen in de laag}}{\text{gemiddelde omvang van de laag}} . \tag{2.21}$$

Bij betrekkelijk zeldzame doodsoorzaken geeft men dit cijfer vaak per 10.000 of 100.000 personen.

In tabel 2.14 is van twee fictieve populaties, A en B, de opbouw naar leeftijd (in klassen van 10 jaar) en de sterfte naar een bepaalde doodsoorzaak X in een bepaald jaar opgenomen. Het ruwe sterftecijfer is in beide populaties vrijwel gelijk en bedraagt 10, resp. 9,8 per duizend personen.

Tabel 2.14. Sterfte naar doodsoorzaak X in twee populaties in een bepaald jaar

Leeftijdsklasse	Bevolking A			Bevolking B		
	Gemiddelde bevolking	Sterfte	Sterfte p. 1000	Gemiddelde bevolking	Sterfte	Sterfte p. 1000
0– 9	2000	8	4	1000	2	2
10–19	2500	15	6	1500	6	4
20–29	1500	9	6	2000	10	5
30–39	1000	8	8	2000	12	6
40–49	700	7	10	1500	12	8
50–59	600	9	15	1000	10	10
60–69	400	10	25	600	12	20
70–79	200	9	45	400	12	30
80 en >	100	15	150	200	24	120
Totaal	9000	90	10	10200	100	9,80

De leeftijdsopbouw van de twee bevolkingen verschilt echter duidelijk: zo is bij A het percentage personen van 40 jaar en ouder ongeveer 22%, terwijl dit bij B ca. 36% bedraagt. Daar de sterfte bij doodsoorzaak X met toenemende leeftijd stijgt, geeft het ruwe sterftecijfer een onjuist beeld. De leeftijdsspecifieke cijfers laten zien, dat voor elke leeftijdsgroep de sterfte bij A hoger is dan die bij B.

Er zijn gevallen, waarin men behoefte gevoelt aan een samenvattend 'gemiddelde' voor elk der te vergelijken populaties, waarin dan bijv. de verschillen in leeftijds- en geslachtsopbouw zijn verdisconteerd.[1] Om zo'n samenvattend kengetal te bereiken staan twee methoden van *standaardisatie*, de directe en de indirecte, ter beschikking.

Bij de *directe standaardisatie* gaat men als volgt tewerk: (1) Bepaal het specifieke sterftecijfer per geslacht/per leeftijdsklasse, (2) Kies een standaardpopulatie met een bepaalde opbouw naar geslacht en leeftijd; bij het standaardiseren van sterftecijfers die in verschillende jaren in één land zijn waargenomen kiest men hiervoor vaak de opbouw in dat land in een bepaald jaar, (3) Bereken de totale sterfte voor de standaardpopulatie, uitgaande van de per laag bij (1) berekende specifieke sterftecijfers. In tabel 2.15 is deze procedure, met standaardisatie naar leeftijd, toegepast op de bevolkingen in tabel 2.14. Als standaardpopulatie is hier de bevolking B gekozen.

[1] De sterfte is vanzelfsprekend nog afhankelijk van vele andere factoren, maar hiervan zijn gewoonlijk geen bruikbare verdelingen voor de gehele bevolking aanwezig.

Tabel 2.15. Directe standaardisatie naar leeftijd voor de bevolkingen in tabel 2.14

Leeftijds-klasse	Sterfte p. 1000 bij A	Leeftijdsopbouw van standaard-populatie B	Verwachte sterfte bij A
0– 9	4	1000	4
10–19	6	1500	9
20–29	6	2000	12
30–39	8	2000	16
40–49	10	1500	15
50–59	15	1000	15
60–69	25	600	15
70–79	45	400	18
80 en >	150	200	30
Totaal	10	10200	134

Tabel 2.15 laat zien dat de verwachte totale sterfte bij A, op basis van de leeftijdsopbouw van B, 134 op de 10.200 personen of 13,1 per 1000 personen bedraagt. Bij gebruik van B als standaardpopulatie is de *index* (Eng. ratio) van de sterfte van A t.o.v. B, met uitschakeling van het verschil in leeftijdsopbouw, dus 134/100 = 1,34.

Afhankelijk van de gekozen standaardpopulatie zal men voor de gestandaardiseerde sterftecijfers verschillende uitkomsten verkrijgen. Zo wordt bij het voorbeeld bij gebruik van de leeftijdsopbouw van A als standaardpopulatie de sterfte-index van A t.o.v. B 90/65,1 = 1,38. De gestandaardiseerde sterftecijfers hebben dan ook geen absolute waarde, maar zij verschaffen slechts een beeld van de sterfteverhoudingen, dat van waarde is als men het verloop van een naar samenstelling veranderende bevolking over een reeks van jaren wil bestuderen.

De *indirecte standaardisatie* kan worden gebruikt als men van de te vergelijken populaties wel beschikt over de opbouw (bijv. naar geslacht en leeftijd) en de ruwe sterftecijfers, maar de specifieke sterftecijfers niet kent. Deze methode verloopt als volgt (wij gaan hierbij van het voorbeeld in tabel 2.14 uit, waarbij alleen naar leeftijd wordt gestandaardiseerd): (1) Kies een standaardpopulatie waarvan de (leeftijds)specifieke sterftecijfers [verder te noemen de standaardsterfte(cijfers)] bekend zijn, (2) Bereken voor elk van de te vergelijken populaties het (uiteraard fictieve) totale sterftecijfer, dat zou gelden als de standaardsterfte van toepassing zou zijn geweest. In tabel 2.16 is deze berekening voor het voorbeeld uitgevoerd. Deze tabel bevat eerst de volledige gegevens van de standaardpopulatie, t.w. de leeftijdsopbouw, het aantal sterfgevallen per leeftijdsklasse en de sterfte per 1000 personen per leeftijdsklasse, en vervolgens de berekening van de totaalsterfte bij A en B op basis van deze standaardsterftecijfers.

Tabel 2.16. Indirecte standaardisatie naar leeftijd voor de bevolkingen in tabel 2.14 op basis van standaardpopulatie S

Leeftijdsklasse	Standaardpopulatie (S)			Bevolking A		Bevolking B	
	Gemidd. bevolking	Sterfte	Sterfte p. 1000 (Standaardsterfte)	Gemidd. bevolking	Sterfte	Gemidd. bevolking	Sterfte
0– 9	1500	3	2	2000	4	1000	2
10–19	2000	10	5	2500	12,5	1500	7,5
20–29	2000	12	6	1500	9	2000	12
30–39	1500	12	8	1000	8	2000	16
40–49	1000	10	10	700	7	1500	15
50–59	800	12	15	600	9	1000	15
60–69	500	10	20	400	8	600	12
70–79	350	14	40	200	8	400	16
80 en >	150	21	140	100	14	200	28
Totaal	9800	104	10,61	9000	79,5	10200	123,5

(3) Indien de leeftijdsopbouw van bevolking A identiek zou zijn geweest aan die van de standaardpopulatie, dan zou het voor A berekende totale (= ruwe) sterftecijfer per 1000 personen gelijk zijn geweest aan dat van de standaardpopulatie. Het totale cijfer voor A is echter volgens (2.20):

$$1000 \cdot \frac{79,5}{9000} = 8,83 \text{ en dat van S: } 1000 \cdot \frac{104}{9800} = 10,61 \text{ (per 1000 pers.)}$$

Het verschil tussen deze cijfers moet dus op het verschil in leeftijdsopbouw van de twee populaties berusten. De invloed van dit verschil in samenstelling kan worden uitgeschakeld door de bij A waargenomen ruwe sterfte per 1000 personen te vermenigvuldigen met de correctiefactor

$$C_A = \frac{\text{totale standaardsterfte per 1000}}{\text{totale berekende sterfte bij A per 1000}} = \frac{10,61}{8,83} = 1,20,$$

zodat wij voor de gestandaardiseerde sterfte bij A krijgen

$$C_A \times \text{(waargenomen sterfte per 1000)} = 1,20 \times 10,0 = 12,0.$$

Voor B is het berekende ruwe sterftecijfer per 1000 personen gelijk aan $1000 \cdot \frac{123,5}{10200} = 12,11$. De correctiefactor is dus $\frac{10,61}{12,11} = 0,876$ en de gestandaardiseerde sterfte bij B is dus $0,876 \times 9,80 = 8,58$.

De sterfte-index van A t.o.v. B is dus, bij gebruik van de indirecte methode, gelijk aan 12,0/8,58 = 1,40.

Ook bij de indirecte standaardisatie hebben de gestandaardiseerde sterftecijfers geen absolute betekenis, daar zij mede afhangen van de

gekozen standaardsterfte. Hun verhouding is echter weer een bruikbare maat voor het vergelijken van verschillende populaties.

Wanneer een keuze tussen de directe en de indirecte methode mogelijk is, verdient in het algemeen de directe methode de voorkeur. Een uitzondering vormt de situatie, die zich bijv. bij het standaardiseren naar leeftijd van ziekteverzuimcijfers van verschillende bedrijven kan voordoen, waarbij de populaties van betrekkelijk kleine omvang zijn zodat slecht bezette of lege leeftijdsklassen voorkomen. Deze materie wordt uitvoerig besproken door DE GROOT en VAN LEEUWEN (1960).

Een andere maat die bij de analyse van sterftecijfers wordt gebruikt is het procentuele aandeel van een bepaalde doodsoorzaak in de totale sterfte van een bepaalde periode, dus

$$100 \cdot \frac{\text{aantal sterfgevallen tengevolge van bepaalde doodsoorzaak}}{\text{totaal aantal sterfgevallen}}.$$

In 1960 bedroeg bijv. de sterfte tengevolge van 'Hartziekten t.g.v. aderverkalking enz.' (B-26 van de kleine internationale lijst van doodsoorzaken van 1955) volgens het Statistisch Zakboek 1961 van het C.B.S.:

bij mannen: 11.123 sterfgevallen, op totaal 47.301 sterfgevallen, d.i. 23,52%,

bij vrouwen: 8.187 sterfgevallen, op totaal 40.185 sterfgevallen, d.i. 20,37%.

Ook dit kengetal wordt vaak specifiek (per geslacht en per leeftijdsgroep) berekend.

Voor de sterfte rondom de geboorte is een aantal speciale kengetallen beschikbaar, die gewoonlijk per jaar (per duizend) worden berekend. De belangrijkste hiervan zijn:

$$\left.\begin{matrix}\textit{Late foetale sterfte}\\ = \textit{doodgeboorten}\end{matrix}\right\} = 1000 \cdot \frac{\text{aantal doodgeborenen}}{\text{totaal aantal (levend- plus dood-) geborenen}}$$

(onder doodgeborenen vallen alle vruchten met een ouderdom van 28 of meer weken, die in de betrokken periode dood ter wereld zijn gekomen);

$$\left.\begin{matrix}\textit{Perinatale}\\ \textit{sterfte}\end{matrix}\right\} = 1000 \cdot \frac{\text{aantal doodgeborenen + overledenen in eerste week}}{\text{totaal aantal (levend- plus dood-) geborenen}}$$

$$\left.\begin{matrix}\textit{Neonatale}\\ \textit{sterfte}\end{matrix}\right\} = 1000 \cdot \frac{\text{aantal overledenen in de eerste 4 levensweken}}{\text{aantal levendgeborenen}}$$

$$\left.\textit{Zuigelingen sterfte}\right\} = 1000 \cdot \frac{\text{aantal overleden kinderen beneden 1 jaar}}{\text{aantal levendgeborenen}}$$

beter is, zoals o.a. ook in Nederland wordt gebruikt:

$$1000 \cdot \frac{\text{aantal overledenen beneden 1 j. onder de levendgeboren in een periode}}{\text{aantal levendgeborenen in die periode}}.$$

Wij laten de sterftetafels buiten beschouwing. Zij worden o.m. besproken door RIJKEN VAN OLST (1958), terwijl hun samenstelling beknopt wordt behandeld in BAKKER (1955, deel II), PEARL (1941) en HILL (1961).

2.7.2. KENGETALLEN VOOR ZIEKTE [1]

Bij het bestuderen van ziekte en ziekteverzuim is de situatie t.a.v. de kengetallen meer complex dan bij de sterfte. In de eerste plaats moet een onderscheid gemaakt worden tussen kengetallen die op een *tijdstip* en die op een *tijdvak* betrekking hebben. De eerstgenoemde verschaffen a.h.w. een moment-opname, de andere betreffen een bepaalde periode.

Bij het bestuderen van de ziekte in een tijdvak kan men zich bezig houden met (zieke) personen en (ziekte)gevallen. Daar elke ziekte een begin, een zekere duur en een einde heeft, zijn er kengetallen voor de frequentie en voor de ziekteduur. Daarnaast kan men frequentieverdelingen geven, zoals bijv. die van het aantal personen met 0, 1, 2, ... ziektegevallen in het tijdvak (het frequentie-patroon). Men kan hierbij dan nog uitgaan van de gevallen die in het tijdvak van onderzoek zijn begonnen, zijn voorgekomen of zijn geëindigd. Tenslotte bestaat ook hier de mogelijkheid om met proportionele kengetallen te werken, waarbij bijv. het voor een bepaalde ziekte gevonden aantal gevallen wordt uitgedrukt in een percentage van het totaal der ziektegevallen.

Een volledige bespreking van deze veelheid van maten valt buiten het kader van dit boek. Het Normalisatierapport van de Gezondheidsraad (Verslagen en mededelingen betreffende de volksgezondheid (1960), 345-408) behandelt deze materie (speciaal gericht op het ziekteverzuim) volledig en systematisch, waarbij tevens een voorbeeld wordt uitgewerkt en aanbevelingen omtrent de keuze van kengetallen worden verstrekt.[2] Verder bevat dit rapport voorbeelden van directe en indirecte standaardisatie van verzuimcijfers, alsmede aanwijzingen betreffende de hand- en mechanische bewerking van gegevens over ziekteverzuim.

2.7.3. LITERATUUR

De 'vital and health statistics' zijn in dit boek slechts beknopt ingeleid. Voor een volledige behandeling verwijzen wij naar de boeken

[1] Voor ongevallen, opname in een ziekenhuis, e.d. gelden overeenkomstige overwegingen.
[2] Zie hiervoor ook DE GROOT (1958).

van BENJAMIN (1959) en SWAROOP (1960), die volledig aan deze materie zijn gewijd. Ook in de boeken over epidemiologie wordt er gewoonlijk vrij veel aandacht aan gewijd. In WOLFENDEN (1954) vindt men een bijzonder volledige verhandeling over sterftetafels.

Tenslotte vermelden wij enkele belangrijke recente Nederlandse publicaties over sterftecijfers:

H. H. VAN GELDEREN — Pre-school child mortality in the Netherlands, Verhandeling 28, N.I.P.G., Leiden, 1955.

J. H. DE HAAS c.s. — Atlas van de Kindersterfte in Nederland, Van Gorcum en Comp., Assen, 1956.

B. J. VAN DEN BERG — Ongevallensterfte in Nederland, in het bijzonder bij kinderen, Verhandeling 47, N.I.P.G., Leiden, 1959.

D. HOOGENDOORN — De Zuigelingensterfte in Nederland, Van Gorcum en Comp., Assen, 1959.

HOOFDSTUK 3

KANSREKENING

3.1. Het begrip kans in het dagelijks leven

In het dagelijks leven heeft het begrip *kans* (synoniem: *waarschijnlijkheid*) een vage en in de regel subjectieve betekenis. Men zegt bijvoorbeeld, dat er slechts een zeer kleine kans is om de leeftijd van 100 jaar te bereiken, of dat het zeer waarschijnlijk is dat het morgen goed weer zal zijn. Deze en soortgelijke uitspraken hebben betrekking op gebeurtenissen of verschijnselen, waarvan de uitkomst niet met zekerheid bekend is. Kan zo'n gebeurtenis twee uitkomsten, A en B, hebben en zeggen wij, dat de uitkomst A waarschijnlijk zal plaats vinden (meer kans heeft dan uitkomst B), dan bedoelen wij hiermee dat de ervaring heeft uitgewezen dat in het verleden onder soortgelijke omstandigheden als die waarin wij nu een oordeel uitspreken, de uitkomst A een grotere frequentie bezat dan de uitkomst B. Feitelijk hebben wij dan in gedachten het quotiënt

$$\frac{\text{aantal malen dat } A \text{ is opgetreden}}{\text{aantal malen dat } A \text{ of } B \text{ is voorgekomen}}$$

bepaald en geconstateerd, dat dit groter is dan $\frac{1}{2}$. Het is duidelijk, dat onze uitspraak ten gunste van A aan stelligheid zal winnen, naarmate het quotiënt dichter bij 1 komt te liggen. De kansrekening stelt ons in staat dit intuïtieve waarschijnlijkheidsbegrip te precizeren, door deze kansen in getallen uit te drukken.

3.2. Kansdefinitie

Wij gaan uit van de volgende definitie: de kans dat bij een aselecte trekking uit een populatie van eindige omvang het kenmerk A optreedt (of korter gezegd: de kans op (van) A) is het frequentiequotiënt

$$P(A) = \frac{f_A}{N} \tag{3.1}$$

(P van lat. probabilitas, f_A = de frequentie van kenmerk A op de populatie, N = de omvang van de populatie).

Uit deze definitie volgt

$$0 \leqslant P(A) \leqslant 1,$$

$P(A) = 0$ als kenmerk A niet voorkomt,

$P(A) = 1$ als kenmerk A uitsluitend voorkomt.

Voorbeelden

3.1. Beschouw de populatie in voorbeeld 1.22. Bij een aselecte trekking is de kans op een arbeider met bloedgroep A 0,445.

3.2. Trekt men uit een spel met 52 kaarten aselect één kaart, dan is de kans op een harten gelijk aan $13/52 = 1/4$. De kans op een aas is $4/52 = 1/13$. De kans op een 'zwarte' kaart is $26/52 = 1/2$.

De definitie kan ook worden toegepast wanneer men worpen met een geldstuk of een dobbelsteen beschouwt. Men kan bijv. één worp met een geldstuk zien als een aselecte trekking van één element uit een populatie van de (willekeurige) omvang N, bestaande uit elementen die het kenmerk 'kruis' of het kenmerk 'munt' dragen. Indien men veronderstelt, dat het geldstuk 'zuiver' is, dan bevat deze populatie evenveel elementen met het kenmerk kruis als met het kenmerk munt. De kans op het gooien van kruis is dan volgens (3.1) gelijk aan $\frac{1}{2}N/N = \frac{1}{2}$. Een reeks van n worpen, onder dezelfde omstandigheden verricht, kan worden gezien als een aselecte steekproef *met teruglegging* van n elementen uit deze populatie.

Bij vraagstukjes betreffende zuivere geldstukken en dobbelstenen kan men ook gebruik maken van de z.g. *klassieke kansdefinitie*:
Als een gebeurtenis kan plaatsvinden op N elkaar uitsluitende en gelijkwaardige manieren en als a van deze uitkomsten het kenmerk A dragen, is de kans op A gelijk aan a/N.

Voorbeeld

3.3. Als men met een dobbelsteen werpt zijn zes uitkomsten mogelijk: elk der zes zijden kan bovenliggen. Deze uitkomsten sluiten elkaar uit: twee of meer zijden kunnen niet tegelijkertijd verschijnen. De veronderstelling dat de dobbelsteen zuiver is maakt de zes uitkomsten gelijkwaardig. De kans op het werpen van zes ogen is dus bij een zuivere dobbelsteen 1/6.

3.3. Kansverdelingen

3.3.1. INLEIDING

Beschouw een populatie bestaande uit 10 knikkers, waarvan er 4 wit en 6 zwart zijn. De populatieverdeling en de relatieve populatieverdeling van het categorisch systeem van kenmerken dat gevormd wordt door de kleuren van de knikkers zijn in tabel 3.1 samengevat.

Tabel 3.1

	Kleur		Totaal
	Wit	Zwart	
Populatieverdeling	4	6	10
Relatieve populatieverdeling	0,4	0,6	1

Beschouw vervolgens een populatie bestaande uit 10 ballen met verschillende gewichten. De populatieverdeling en de relatieve populatieverdeling van de grootheid x, het gewicht van deze ballen, zijn in tabel 3.2 gegeven.

Tabel 3.2

	x = gewicht in gr				Totaal
	3	4	5	6	
Populatieverdeling	1	3	4	2	10
Relatieve populatieverdeling	0,1	0,3	0,4	0,2	1

Het is duidelijk dat bij aselecte trekking van 1 knikker uit de in tabel 3.1 gegeven populatie de kansen van de twee kleuren gegeven worden door de relatieve populatieverdeling. Evenzo geldt dat bij aselecte trekking van 1 bal uit de in tabel 3.2 gegeven populatie de kansen van de vier waarden van de grootheid x in de relatieve frequentieverdeling staan. Bij aselecte trekking van één element uit een populatie van eindige omvang is de *kansverdeling* van een categorisch systeem of van een discreet verdeelde variabele grootheid[1] dus de relatieve frequentieverdeling van het systeem, resp. van de grootheid, op de populatie.

Men noemt een grootheid die een kansverdeling volgt een *stochastische grootheid*. In navolging van VAN DANTZIG geeft men het stochastische karakter van een grootheid aan door het symbool waarmee deze grootheid wordt aangeduid te onderstrepen. Hetzelfde symbool zonder onderstreping dient dan om waarden aan te geven, die deze grootheid heeft aangenomen of kan aannemen. Volgens deze notatie betekent dus $P(\underline{x} = 2)$: de kans dat de stochastische grootheid $\underline{x}$ de waarde 2 aanneemt, en $P(\underline{x} = x_i)$: de kans dat de stochastische grootheid $\underline{x}$ de waarde x_i aanneemt (waarin x_i staat voor één van de mogelijke waarden van de grootheid). Vaak gebruikt men $P(\underline{x} = x)$ in plaats van $P(\underline{x} = x_i)$. In gevallen waarin dit geen misverstand kan opleveren gebruiken wij soms een nog kortere notatie door voor $P(\underline{x} = x_i)$ te schrijven $P(x_i)$, zodat bijvoorbeeld $P(\underline{x} = 2)$ vervangen wordt door $P(2)$.

Voorbeelden

3.4. Veronderstel dat men bij het werpen met een zuiver geldstuk aan de uitkomst kruis de waarde $x = 0$ en aan de uitkomst munt de waarde $x = 1$ toekent. De kansverdeling van de grootheid $\underline{x}$ kan dan worden samengevat in de volgende formule:

$$P(\underline{x} = x_i) = 1/2 \qquad (x_i = 0, 1),$$

te lezen als: de kans dat de stochastische grootheid $\underline{x}$ (de uitkomst van één worp met het geldstuk) de waarde x_i aanneemt is gelijk aan 1/2 wanneer x_i nul en 1/2 wanneer x_i één is.

[1] Zie ook 5.3.1.

3.5. De kansverdeling van de grootheid $\underline{x}$, het aantal ogen bij één worp met een zuivere dobbelsteen, is:

x_i	1	2	3	4	5	6
$P(\underline{x} = x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Voor deze kansverdeling kan men ook schrijven

$$P(\underline{x} = x_i) = 1/6 \qquad (x_i = 1, 2, \ldots, 6).$$

3.3.2. GEMIDDELDE EN VARIANTIE VAN EEN KANSVERDELING

Wij zullen in het vervolg herhaaldelijk gebruik maken van het gemiddelde en de variantie van de kansverdeling van een stochastische grootheid. Daar een kansverdeling een relatieve frequentieverdeling van een populatie is, waarin voor de bij de waarde x_i behorende relatieve frequentie f_i/N de kans $P(\underline{x} = x_i)$ in de plaats is gesteld, volgt uit formule (2.4) dat

$$\mu_x = \Sigma \frac{f_i}{N} \cdot x_i = \Sigma P(\underline{x} = x_i) \cdot x_i = \Sigma P(x_i) \cdot x_i. \tag{3.2}$$

Men noemt dit gemiddelde ook de (*mathematische*) *verwachting* van de grootheid $\underline{x}$.

De variantie van de kansverdeling van $\underline{x}$ volgt op analoge wijze uit formule (2.16):

$$\begin{aligned} \sigma_x^2 &= \Sigma \frac{f_i}{N} \cdot x_i^2 - \mu_x^2 = \Sigma P(\underline{x} = x_i) \cdot x_i^2 - \mu_x^2 \\ &= \Sigma P(x_i) \cdot x_i^2 - \mu_x^2. \end{aligned} \tag{3.3}$$

Voorbeeld

3.6. Voor de kansverdeling van de grootheid $\underline{x}$ in voorbeeld 3.5 vindt men dus door toepassing van (3.2) en (3.3):

$$\mu_x = \tfrac{1}{6}(1) + \tfrac{1}{6}(2) + \ldots + \tfrac{1}{6}(6) = \frac{21}{6} = 3{,}5,$$

$$\sigma_x^2 = \tfrac{1}{6}(1)^2 + \tfrac{1}{6}(2)^2 + \ldots + \tfrac{1}{6}(6)^2 - (3{,}5)^2 = \frac{17{,}5}{6} = 2{,}9167.$$

De in 2.5.1 behandelde eigenschappen betreffende gemiddelden en standaardafwijkingen gelden ook voor kansverdelingen (kijk deze eigenschappen na!). Bezit de grootheid $\underline{x}$ een kansverdeling met gemiddelde μ_x en variantie σ_x^2 en zijn A en B constanten, dan bezit de grootheid $\underline{x} - A$ een kansverdeling met gemiddelde $\mu_x - A$ en met variantie σ_x^2 (eigenschappen 1 en 2). De grootheid $\underline{x}/B$ bezit dan een kansverdeling met gemiddelde μ_x/B en met variantie σ_x^2/B^2 (eigenschappen 3 en 4).

Een stochastische grootheid $\underline{x}$ wordt *gestandaardiseerd*, wanneer men haar met haar gemiddelde vermindert en vervolgens door haar

spreiding deelt. Volgens de eigenschappen 5 en 6 heeft de gestandaardiseerde grootheid

$$\frac{\underline{x} - \mu_x}{\sigma_x} \tag{3.4}$$

als gemiddelde 0 en als variantie 1.

3.4. Kansregels

Er zijn verschillende regels voor het rekenen met kansen. Wij behandelen alleen de twee regels die in dit boek worden gebruikt.

3.4.1. DE SOMREGEL

De somregel luidt: wanneer men een aantal kenmerken of waarden van een kansverdeling samenvoegt, dan is de kans op het samengestelde kenmerk gelijk aan de som van de kansen van de samenstellende kenmerken of waarden.

Deze regel volgt onmiddellijk uit de kansdefinitie (3.1), daar immers

$$P(A \text{ of } B \text{ of } C) = \frac{f_A + f_B + f_C}{N} = \frac{f_A}{N} + \frac{f_B}{N} + \frac{f_C}{N} = P(A) + P(B) + P(C).$$

De kans op het werpen van een even aantal ogen met een zuivere dobbelsteen is dus

$$P(\underline{x} = 2 \text{ of } 4 \text{ of } 6) = P(\underline{x} = 2) + P(\underline{x} = 4) + P(\underline{x} = 6) = 3/6.$$

Voor de populatie in tabel 3.2 is de kans dat men een bal met een gewicht van 5 gram of minder verkrijgt gelijk aan 0,8.

3.4.2. DE PRODUCTREGEL

De productregel luidt: de kans op het tegelijk optreden van twee of meer kenmerken of waarden, die tot *onafhankelijke* kansverdelingen behoren, is gelijk aan het product van de kansen van ieder van deze kenmerken of waarden afzonderlijk. Zijn dus de stochastische grootheden $\underline{x}$ en $\underline{y}$ onafhankelijk, dan geldt dat

$$P(\underline{x} = x_i \text{ en } \underline{y} = y_j) = P(\underline{x} = x_i) \cdot P(\underline{y} = y_j).$$

Beschouw de twee volgende populaties:

	Zwarte knikkers	Witte knikkers	Totaal
Populatie I	Z_1	W_1	N_1
Populatie II	Z_2	W_2	N_2

Bij aselecte trekking uit populatie I is de kans op een zwarte knikker $P(Z \mid \mathrm{I}) = Z_1/N_1$. Bij aselecte trekking uit populatie II is de kans op

een zwarte knikker $P(Z \mid \text{II}) = Z_2/N_2$. De uitkomst van de trekking uit populatie I oefent geen invloed uit op de uitkomst van de trekking uit populatie II: de kansen behoren dus tot onafhankelijke verdelingen. De productregel stelt nu dat de kans dat beide trekkingen een zwarte knikker opleveren gelijk is aan $P(Z|\text{I}) \cdot P(Z|\text{II}) = \frac{Z_1}{N_1} \cdot \frac{Z_2}{N_2}$.

De juistheid van deze uitkomst blijkt, als wij een derde, denkbeeldige populatie beschouwen, die bestaat uit alle mogelijke paren knikkers waarvan het eerste lid uit populatie I en het tweede uit populatie II komt. Er zijn N_1N_2 van deze paren mogelijk, daar men bij iedere knikker uit de eerste populatie successievelijk iedere knikker uit de tweede kan kiezen. Hieronder bevinden zich (volgens dezelfde redenering) Z_1Z_2 zwarte paren. In de plaats van een aselecte trekking van één knikker uit iedere werkelijke populatie kan men de aselecte trekking van één paar knikkers uit de denkbeeldige populatie stellen en de kans dat hierbij een paar zwarte knikkers verschijnt is volgens (3.1)

$$\frac{Z_1Z_2}{N_1N_2} = \frac{Z_1}{N_1} \cdot \frac{Z_2}{N_2},$$

de uitkomst die de productregel oplevert. Wij laten het aan de lezer over de voorgaande redenering tot drie of meer populaties uit te breiden. Verder vermelden wij dat ook het omgekeerde geldt: blijkt voor twee (of meer) kansverdelingen de productregel op te gaan, dan volgt daaruit dat deze kansverdelingen onafhankelijk (men zegt ook wel: stochastisch onafhankelijk) zijn.

Men kan de productregel ook voor allerlei typische kansrekenvraagstukjes gebruiken. Stel bijv. dat men werpt met een zuiver geldstuk, werpt met een zuivere dobbelsteen én aselect een kaart trekt (uit een spel van 52 kaarten). De kans dat de uitkomst munt, zes ogen en een aas tegelijkertijd optreden is dan volgens de productregel gelijk aan $\frac{1}{2} \times \frac{1}{6} \times \frac{4}{52} = \frac{1}{156}$.

3.4.3. AFLEIDING VAN KANSVERDELINGEN

Door eerst de productregel en vervolgens de somregel toe te passen kan men uit twee of meer onafhankelijke kansverdelingen nieuwe kansverdelingen afleiden. Wij geven enkele voorbeelden die laten zien welke procedure men daarbij kan volgen.

Voorbeelden

3.7. Uit elk van de populaties in tabel 3.3 trekt men aselect één kind. Welke kansverdeling volgt de grootheid z, het aantal jongens in deze 'steekproef' van twee kinderen?

Tabel 3.3

	Populaties I	II	
Jongens	2	3	5
Meisjes	18	7	25
	20	10	30

In tabel 3.4 is de gevraagde kansverdeling berekend. De grootheid $\underline{z}$ kan de waarden 0, 1 en 2 aannemen. Bij A is aangegeven hoe deze waarden tot stand kunnen komen. Bij B is met de productregel uit de kansverdelingen van de twee populaties (die langs de randen van de tabel staan) berekend welke kansen bij de in A gegeven mogelijkheden behoren: deze tabel bevat de z.g. *simultane* kansverdeling van de beide populaties. Tenslotte volgt uit de tabellen A en B de kansverdeling van $\underline{z}$ (waarbij de kans van $\underline{z} = 1$ door toepassing van de somregel wordt verkregen).

Tabel 3.4. Kansverdeling van de grootheid $\underline{z}$, het aantal jongens bij aselecte trekking van één kind uit elk van de populaties in tabel 3.3.

A. *Mogelijkheden*

Populatie II \ Populatie I	J	M
J	$z=2$	$z=1$
M	$z=1$	$z=0$

B. *Simultane kansverdeling*

Populatie II \ Populatie I	J	M	
J	0,03	0,27	0,3
M	0,07	0,63	0,7
	0,1	0,9	1

C. *Kansverdeling van* $\underline{z}$

z_i	0	1	2	Totaal
$P(\underline{z} = z_i)$	0,63	0,34	0,03	1

Volgens de somregel, toegepast op de simultane verdeling, is $P(\underline{z} = 1) = 0{,}27 + 0{,}07 = 0{,}34$.

3.8. Uit elk van de populaties in tabel 3.5 trekt men aselect één element. Welke kansverdeling volgt de som van de waarden van de twee elementen de grootheid $\underline{z} = \underline{x} + \underline{y}$?

Tabel 3.5

Populatie I

x_i	0	1	2	3
$P(\underline{x} = x_i)$	0,2	0,3	0,4	0,1

Populatie II

y_j	2	4	6
$P(\underline{y} = y_j)$	0,2	0,6	0,2

De kansverdeling van $\underline{z}$ is in tabel 3.6 volgens het in tabel 3.4 gebruikte schema afgeleid. Bij A staan weer de mogelijkheden en bij B de bijbehorende kansen berekend met de productregel. Herhaalde toepassing van de somregel op de simultane verdeling in B levert de kansverdeling van de grootheid $\underline{z}$ op.

Tabel 3.6. Kansverdeling van de grootheid $\underline{z} = \underline{x} + \underline{y}$, d.i. de som van de waarden van de elementen bij trekking van één element uit elk van de populaties, waarvan de verdelingen in tabel 3.5 staan

A. *Mogelijkheden*

Populatie II: y \ Populatie I: x	0	1	2	3
2	2	3	4	5
4	4	5	6	7
6	6	7	8	9

B. *Simultane verdeling van $\underline{x}$ en $\underline{y}$*

Populatie II: y \ Populatie I: x	0	1	2	3	
2	0,04	0,06	0,08	0,02	0,2
4	0,12	0,18	0,24	0,06	0,6
6	0,04	0,06	0,08	0,02	0,2
	0,2	0,3	0,4	0,1	1

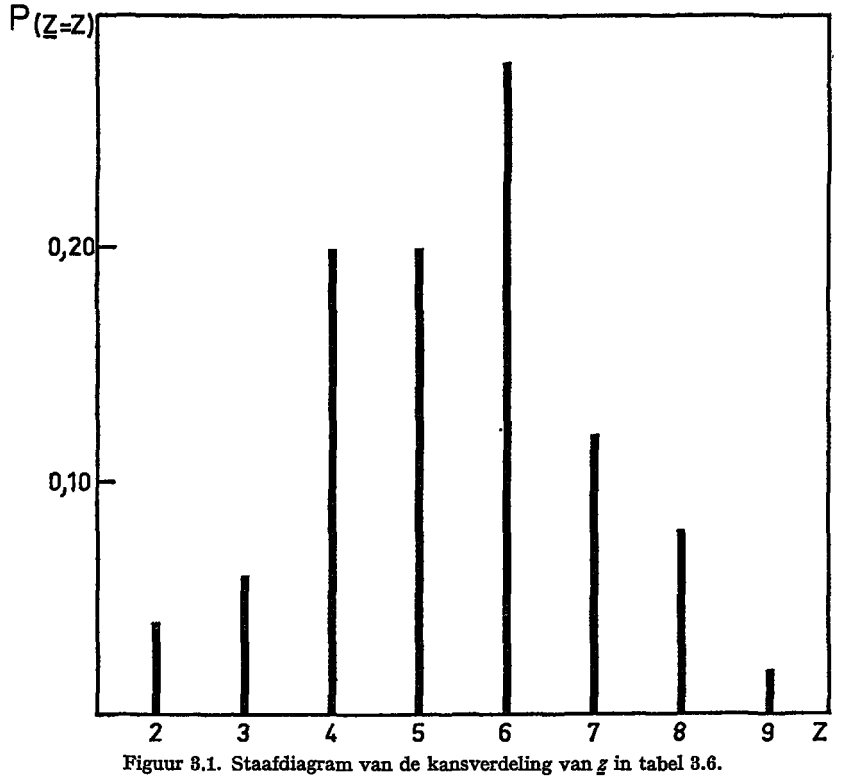
C. *Kansverdeling van $\underline{z}$*

z_i	2	3	4	5	6	7	8	9
$P(\underline{z} = z_i)$	0,04	0,06	0,20	0,20	0,28	0,12	0,08	0,02

Bijvoorbeeld:
$P(\underline{z} = 5)$
$= 0{,}02 + 0{,}18$
$= 0{,}20$

Het is duidelijk dat men de kansverdeling van het *gemiddelde* van de waarden van de twee elementen verkrijgt, indien men in tabel 3.6C de waarden z_i vervangt door de waarden $z_i/2$.

Als grafische voorstelling van een kansverdeling gebruikt men gewoonlijk een staafdiagram en ook wel een kolommendiagram. In figuur 3.1 is van de kansverdeling in tabel 3.6C het staafdiagram getekend.



Figuur 3.1. Staafdiagram van de kansverdeling van $\underline{z}$ in tabel 3.6.

3.4.4.* GEMIDDELDE EN VARIANTIE VAN EEN AFGELEIDE KANSVERDELING

Indien uit de grootheden $\underline{x}$ en $\underline{y}$ een nieuwe kansverdeling wordt afgeleid door te nemen

$$\underline{z} = A\underline{x} + B\underline{y}$$

(waarbij A en B constanten zijn), dan geldt

(3.5) $$\mu_z = A\mu_x + B\mu_y$$

en

(3.6) $$\sigma_z^2 = A^2\sigma_x^2 + B^2\sigma_y^2.$$

Beschouw voorbeeld 3.8. Uit tabel 3.5 vindt men met de formules (3.2) resp. (3.3): $\mu_x = 1{,}4$, $\sigma_x^2 = 0{,}84$, $\mu_y = 4$, $\sigma_y^2 = 1{,}60$. Daar $\underline{z} = \underline{x} + \underline{y}$ is $A = B = 1$ en volgt uit (3.5)

$$\mu_z = \mu_x + \mu_y = 1{,}4 + 4{,}0 = 5{,}4$$

en uit (3.6)

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 = 0{,}84 + 1{,}60 = 2{,}44.$$

Wij laten aan de lezer over om te verifiëren dat men bij berekening van μ_z en σ_z^2 uit de kansverdeling in tabel 3.6 dezelfde uitkomsten vindt.

Bij voorbeeld 3.8 geldt voor de kansverdeling van het gemiddelde van de waarden van de twee elementen

$$\underline{v} = \underline{z}/2 = (\underline{x} + \underline{y})/2 = \underline{x}/2 + \underline{y}/2.$$

In dit geval is dus $A = B = \frac{1}{2}$, zodat volgens (3.5)

$$\mu_v = 1{,}4/2 + 4{,}0/2 = 2{,}7$$

en volgens (3.6)

$$\sigma_v^2 = 0{,}84/(2)^2 + 1{,}60/(2)^2 = 0{,}61.$$

Leidt men uit de grootheden $\underline{x}$ en $\underline{y}$ een nieuwe kansverdeling af door te nemen

$$\underline{w} = A\underline{x} - B\underline{y}$$

dan geldt dat

(3.7) $$\mu_w = A\mu_x - B\mu_y$$

en

(3.8) $$\sigma_w^2 = A^2\sigma_x^2 + B^2\sigma_y^2.$$

Het voorgaande kan direct tot het afleiden van een kansverdeling uit meer dan twee onafhankelijke kansverdelingen worden uitgebreid. Zo geldt bijvoorbeeld dat de grootheid $\underline{u} = \underline{x} + \underline{y} - \underline{z}$ als gemiddelde heeft $\mu_u = \mu_x + \mu_y - \mu_z$ en als variantie $\sigma_u^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2$.

3.4.5. OPGAVEN[1]

3.1. Populatie I bestaat uit 5 rode en 15 witte ballen, populatie II telt 6 rode en 24 witte ballen. Men trekt uit elke populatie aselect één bal. Hoe groot is de kans dat beide ballen dezelfde kleur hebben?

3.2. Populatie I bestaat uit 4 witte, 5 rode, en 6 zwarte ballen en populatie II uit 5 witte, 6 rode en 7 zwarte ballen. Men trekt aselect uit elke populatie één bal. Hoe groot is de kans dat de twee getrokken ballen dezelfde kleur hebben?

3.3. Men heeft twee dobbelstenen. De eerste heeft 2 blauwe en 4 rode zijden. De tweede heeft eveneens blauwe en rode zijden, maar de verdeling daarvan is onbekend. Gegeven is echter dat de kans dat bij een worp met beide stenen dezelfde kleur bovenkomt gelijk is aan 14/36. Hoe zijn de kleuren van de tweede dobbelsteen verdeeld?

3.4. Men werpt met twee zuivere dobbelstenen. *a.* Hoe groot is de kans op een paar (= zelfde aantal ogen op beide stenen)? *b.* Leid de kansverdeling van de som van het aantal ogen op beide stenen (z) af. *c.* Geef μ_z en σ_z^2. *d.* Teken van de kansverdeling van $\underline{z}$ een staafdiagram.

3.5. Men werpt met 4 zuivere dobbelstenen. *a.* Bepaal de kansverdeling van de som van het aantal ogen op de vier stenen. *b.* Bereken het gemiddelde en de variantie van deze kansverdeling. *c.* Teken van de kansverdeling een staafdiagram. Wat valt u hierbij op?

3.6. Twee onafhankelijke grootheden $\underline{x}$ en $\underline{y}$ bezitten de volgende kansverdelingen:

x_i	2	3	4	5	6
$P(\underline{x} = x_i)$	0,1	0,2	0,3	0,2	0,2

y_j	0	1	2
$P(\underline{y} = y_j)$	0,6	0,3	0,1

a. Bereken μ_x, σ_x^2, μ_y en σ_y^2.
b. Geef de kansverdeling van de grootheid $2\underline{x}$ en het gemiddelde en de variantie van deze verdeling.
c. Bepaal de kansverdeling van $\underline{z} = \underline{x} + \underline{y}$, alsmede μ_z en σ_z^2.
d. Bepaal de kansverdeling van $\underline{w} = \underline{x} - \underline{y}$, alsmede μ_w en σ_w^2.
e. Bepaal de kansverdeling van $\underline{v} = 2\underline{x} - \underline{y}$, alsmede μ_v en σ_v^2.

3.7. Een populatie bestaat uit 9 zwarte en 16 witte ballen. Een tweede populatie bestaat uit 4 zwarte en 9 witte ballen. Men trekt aselect één bal uit de eerste populatie, plaatst deze in de tweede en trekt daaruit vervolgens aselect één bal. Hoe groot is de kans dat: *a.* De laatstgetrokken bal zwart is? *b.* Wit is? *c.* Beide ballen dezelfde kleur hebben? *d.* Beide ballen verschillend van kleur zijn? *e.* Beide ballen zwart zijn? *f.* Ten minste één bal zwart is? *g.* Ten hoogste één bal wit is?

3.8. Een bak bevat 5 zwarte en 4 witte ballen. Zij worden één voor één aselect getrokken. De eerste drie ballen worden in een zwarte en de overige ballen in een witte doos geplaatst. Hoe groot is de kans dat het aantal zwarte ballen in de zwarte doos plus het aantal witte ballen in de witte doos gelijk is aan 5?

3.9. In een collectie van 12 lampen bevinden zich 2 defecte exemplaren. Iemand beproeft aselect één voor één de lampen, totdat beide defecten zijn opgespoord. Hoe groot is de kans dat hij de tweede defecte lamp bij de tiende keer proberen vindt?

[1] Het is gewenst dat men de opgaven 3.1 t/m 3.6 maakt alvorens verder te lezen.

3.5. Kansverdelingen van steekproeffuncties

3.5.1. STEEKPROEFFUNCTIES

In 2.6.4 definieerden wij een *steekproeffunctie* (of statistische grootheid) als een waarde waardoor een steekproef kan worden gekarakteriseerd. Men noemt de verzameling van alle waarden van een steekproeffunctie, die door toepassing van een zeker steekproefvoorschrift uit een populatie kan ontstaan, de *steekproevenruimte.*

Wij brengen nog even in herinnering dat bij trekking met teruglegging elk element, nadat daaraan een waarneming is verricht, weer aan de populatie wordt toegevoegd, zodat deze populatie tijdens het trekken niet van samenstelling verandert. Bij het trekken zonder teruglegging wordt daarentegen een getrokken element niet weer aan de populatie toegevoegd, zodat deze van trekking tot trekking verandert.

Voorbeelden

3.9. Trekt men aselect (met of zonder teruglegging) een steekproef van 4 elementen uit de populatie in tabel 3.1 en kiest men als steekproeffunctie x = het aantal witte knikkers in de steekproef, dan bestaat de steekproevenruimte uit de waarden $x = 0, 1, 2, 3, 4$. Kiest men als steekproeffunctie $p = x/4$, d.i. de fractie witte knikkers in de steekproef, dan bestaat de steekproevenruimte uit de waarden $p = 0, 0{,}25, 0{,}50, 0{,}75$ en 1.

3.10. Trekt men aselect met teruglegging twee ballen uit de populatie in tabel 3.2 en kiest men als steekproeffunctie het gemiddelde van de gewichten van de ballen in deze steekproef, dan bestaat de steekproevenruimte uit de waarden $\bar{x} = 3; 3{,}5; 4; 4{,}5; 5; 5{,}5; 6$ (ga dit na!). Kiest men als steekproeffunctie de spreidingsbreedte van de gewichten, dan bestaat de steekproevenruimte uit de waarden 0, 1, 2, 3.

3.11. Trekt men uit de populatie in tabel 3.2 zonder teruglegging twee ballen, dan bestaat de steekproevenruimte van het gemiddelde uit de waarden 3,5; 4; 4,5; 5; 5,5; 6 en die van de spreidingsbreedte uit de waarden 0, 1, 2, 3.

In deze paragraaf houden wij ons verder bezig met de kansverdelingen van steekproeffuncties, die berusten op een steekproef van n elementen die *aselect* worden getrokken uit een populatie van N elementen waarvan de samenstelling bekend is. De twee trekkingsmethoden (met en zonder teruglegging) worden daarbij afzonderlijk behandeld.

Wij merken hier reeds op dat in de praktijk het omgekeerde probleem aan de orde komt. Men krijgt dan bijvoorbeeld te maken met problemen, waarbij conclusies over de (onbekende) populatiesamenstelling moeten worden getrokken op basis van de waarde, die een steekproeffunctie voor *één* aselecte steekproef heeft aangenomen. In de hoofdstukken 6 en 7 zullen wij zien, dat dit soort problemen kan worden opgelost, indien men de kansverdeling van de betrokken steekproeffunctie voor verschillende (hypothetische) populatiesamenstellingen kan berekenen.

3.5.2. STEEKPROEVEN MET TERUGLEGGING

Wanneer een steekproef van n elementen aselect met teruglegging wordt getrokken, heeft men bij elke trekking met dezelfde populatie te doen. De kansverdeling van een steekproeffunctie kan dan in principe volgens de in 3.4.3 besproken procedure worden afgeleid. Wij geven hiervan een aantal eenvoudige voorbeelden.

Voorbeelden

3.12. Men trekt uit de populatie in tabel 3.1 aselect met teruglegging twee knikkers. Welke kansverdeling volgt de grootheid $\underset{\sim}{x}$, het aantal witte knikkers in deze steekproef?

Trekking 1 èn trekking 2 geschieden uit een populatie, die bestaat uit 4 zwarte en 6 witte knikkers. Bij toepassing van de productregel kan men weer twee vierkanten vormen (zie tabel 3.7), waarin de mogelijke uitkomsten van de trekking met de daarbij behorende waarden van de grootheid $\underset{\sim}{x}$ zijn opgenomen (A) en de bijbehorende kansen zijn berekend (B). Uit A en B kan weer de kansverdeling van $\underset{\sim}{x}$ worden opgesteld (C).

Tabel 3.7. Kansverdeling van de grootheid $\underset{\sim}{x}$, het aantal witte knikkers in een aselect met teruglegging getrokken steekproef van 2 knikkers uit de populatie in tabel 3.1

A. *Waarden van* $\underset{\sim}{x}$

Trekking 2 \ Trekking 1	Wit = 1	Zwart = 2
Wit = 1	W — W $x = 2$	Z — W $x = 1$
Zwart = 0	W — Z $x = 1$	Z — Z $x = 0$

B. *Simultane kansverdeling*

Trekking 2 \ Trekking 1	W	Z	
W	0,16	0,24	0,4
Z	0,24	0,36	0,6
	0,4	0,6	1

C. *Kansverdeling van* $\underset{\sim}{x}$

x_i	0	1	2
$P(\underset{\sim}{x} = x_i)$	0,36	0,48	0,16

3.13. Men trekt uit de populatie in tabel 3.2 aselect met teruglegging een steekproef van 2 ballen. Welke kansverdeling volgt de grootheid $\underset{\sim}{\bar{x}}$, het gemiddelde van de gewichten van de ballen in de steekproef?

In tabel 3.8 is volgens de inmiddels gebruikelijke procedure de gevraagde kansverdeling berekend. Bij A staan de sommen van de gewichten die mogelijk zijn en bij B de bij deze sommen behorende kansen. De gemiddelden in C volgen direct uit de sommen, door deze door 2 te delen.

Tabel 3.8. Kansverdeling van de grootheid $\bar{x}$, het gemiddelde gewicht van 2 ballen die aselect met teruglegging worden getrokken uit de populatie in tabel 3.2

A. *Waarden van* $S = x_1 + x_2$

	x_2 \ x_1	Trekking 1 3	4	5	6
	3	6	7	8	9
Trek-	4	7	8	9	10
king 2	5	8	9	10	11
	6	9	10	11	12

B. *Simultane kansverdeling*

		Trekking 1 3	4	5	6	
	3	0,01	0,03	0,04	0,02	0,1
Trek-	4	0,03	0,09	0,12	0,06	0,3
king 2	5	0,04	0,12	0,16	0,08	0,4
	6	0,02	0,06	0,08	0,04	0,2
		0,1	0,3	0,4	0,2	1

C. *Kansverdelingen van* $\underline{S}$ *en* $\underline{\bar{x}} = \underline{S}/2$

S_i	6	7	8	9	10	11	12
$\bar{x}_i$	3	$3\frac{1}{2}$	4	$4\frac{1}{2}$	5	$5\frac{1}{2}$	6
$P(\underline{S} = S_i) = P(\underline{\bar{x}} = \bar{x}_i)$	0,01	0,06	0,17	0,28	0,28	0,16	0,04

3.14. Welke kansverdeling volgt de spreidingsbreedte voor de in voorbeeld 3.13 genoemde steekproef?

Voor het berekenen van deze kansverdeling dient men in tabel 3.8 het vierkantje A te vervangen door het nevenstaande, waarin van elke mogelijkheid de spreidingsbreedte (R = verschil tussen de hoogste en laagste waarde) is opgenomen. De bij deze waarden van R behorende kansen staan reeds in tabel 3.8B, zodat —met de somregel — de volgende kansverdeling van $\underline{R}$ verschijnt:

Waarden van R

	x_2 \ x_1	Trekking 1 3	4	5	6
	3	0	1	2	3
Trek-	4	1	0	1	2
king 2	5	2	1	0	1
	6	3	2	1	0

R_i	0	1	2	3
$P(\underline{R} = R_i)$	0,30	0,46	0,20	0,04

[Bijv.: $P(R = 0) = 0{,}01 + 0{,}09 + 0{,}16 + 0{,}04 = 0{,}30$].

3.15. Het is duidelijk dat de kansverdeling van het gemiddelde gewicht van 3 ballen, die aselect met teruglegging worden getrokken uit de populatie in tabel 3.2, kan worden afgeleid uit de kansverdeling bij het trekken van 1 bal en de in voorbeeld 3.13 bepaalde kansverdeling van de som van 2 ballen. Deze afleiding is in tabel 3.9 uitgevoerd en behoeft o.i. geen verder commentaar.

Tabel 3.9. Kansverdeling van het gemiddelde gewicht van 3 ballen, die aselect met teruglegging worden getrokken uit de populatie in tabel 3.2

A. *Waarden van* $S = x_1 + x_2 + x_3$

$x_1 + x_2$ \ x_3	3	4	5	6
6	9	10	11	12
7	10	11	12	13
8	11	12	13	14
9	12	13	14	15
10	13	14	15	16
11	14	15	16	17
12	15	16	17	18

B. *Simultane kansverdeling*

$x_1 + x_2$ \ x_3	3	4	5	6	
6	0,001	0,003	0,004	0,002	0,01
7	0,006	0,018	0,024	0,012	0,06
8	0,017	0,051	0,068	0,034	0,17
9	0,028	0,084	0,112	0,056	0,28
10	0,028	0,084	0,112	0,056	0,28
11	0,016	0,048	0,064	0,032	0,16
12	0,004	0,012	0,016	0,008	0,04
	0,1	0,3	0,4	0,2	1

C. *Kansverdelingen van* $\underline{S}$ *en* $\underline{\bar{x}} = \underline{S}/3$

S_i	9	10	11	12	13	14	15	16	17	18
$\bar{x}_i$	3	$3\frac{1}{3}$	$3\frac{2}{3}$	4	$4\frac{1}{3}$	$4\frac{2}{3}$	5	$5\frac{1}{3}$	$5\frac{2}{3}$	6
$P(\underline{S} = S_i) = P(\underline{\bar{x}} = \bar{x}_i)$	0,001	0,009	0,039	0,105	0,192	0,246	0,220	0,132	0,048	0,008

*Opmerking**

Voor de populatie in tabel 3.2 is $\mu_x = 4{,}7$ en $\sigma_x^2 = 0{,}81$. Voor de in voorbeeld 3.13 afgeleide kansverdeling van het gemiddelde gewicht van 2 ballen, die aselect met teruglegging uit deze populatie worden getrokken, geldt volgens formule (3.5)

$$\mu_{\bar{x}} = (\mu_x + \mu_x)/2 = 2\mu_x/2 = \mu_x = 4{,}7.$$

Voor de variantie van deze kansverdeling vindt men met formule (3.6)

$$\sigma_{\bar{x}}^2 = (\sigma_x^2 + \sigma_x^2)/(2)^2 = 2\sigma_x^2/4 = \sigma_x^2/2 = 0{,}405.$$

Voor de in voorbeeld 3.15 afgeleide kansverdeling van het gemiddelde gewicht van 3 aselect met teruglegging getrokken ballen vindt men op deze wijze

$$\mu_{\bar{x}} = (\mu_x + \mu_x + \mu_x)/3 = 3\mu_x/3 = \mu_x = 4{,}7$$

en

$$\sigma_{\bar{x}} = (\sigma_x^2 + \sigma_x^2 + \sigma_x^2)/(3)^2 = 3\sigma_x^2/9 = \sigma_x^2/3 = 0{,}27$$

(Controleer deze uitkomsten door de gemiddelden en varianties uit de kansverdelingen in tabel 3.8, resp. 3.9 te berekenen.)

Uit het voorgaande volgt dat bij *trekking met teruglegging* voor de kansverdeling van het gemiddelde van een aselecte steekproef van n elementen geldt

(3.9) $$\mu_{\bar{x}} = \mu_x$$

(3.10) $$\sigma_{\bar{x}}^2 = \sigma_x^2/n.$$

Het is duidelijk dat de procedure voor het afleiden van kansverdelingen van steekproeffuncties, zoals deze bij de voorafgaande voorbeelden is gebruikt, bijzonder tijdrovend wordt wanneer de populaties een groot aantal waarden of categorieën tellen, resp. wanneer de steekproefomvang groot is. Wij komen hier in 3.5.4 op terug.

3.5.3. STEEKPROEVEN ZONDER TERUGLEGGING

Wanneer uit een populatie aselect zonder teruglegging een steekproef van n elementen wordt getrokken, kan de productregel niet zonder meer worden toegepast. Bij eenvoudige vraagstukjes kan men de regel nog wel toepasbaar maken, door bij elke trekking rekening te houden met de veranderingen, die de populatie door de voorafgaande trekkingen heeft ondergaan.

Trekt men bijvoorbeeld aselect zonder teruglegging 2 knikkers uit de populatie in tabel 3.1 en wordt gevraagd hoe groot de kans is dat zij beide wit zijn, dan kan men als volgt redeneren: 1. De kans op een witte knikker is bij de eerste trekking gelijk aan 4/10. 2. Indien de eerst getrokken knikker wit is blijven er 9 knikkers over, waarvan er $4 - 1 = 3$ wit zijn. De kans dat de tweede trekking eveneens een witte knikker oplevert is dus 3/9. 3. De kans dat men beide keren een witte knikker trekt is dus (daar nu de twee kansverdelingen onafhankelijk zijn) gelijk aan $4/10 \times 3/9 = 12/90$.

De lezer kan zelf vaststellen, dat het omslachtig is (en spoedig tot fouten leidt) om bijv. langs deze weg te berekenen, hoe groot de kans is op 3 witte knikkers, wanneer men aselect zonder teruglegging 5 knikkers uit de populatie in tabel 3.1 trekt.

Bij de berekening van kansverdelingen van steekproeffuncties bij het trekken van een steekproef zonder teruglegging kan men het best overgaan tot de reeds in 3.4.2 gevolgde procedure: het opstellen van de (denkbeeldige) populatie, die bestaat uit alle steekproeven van n elementen, die mogelijk zijn. Wij geven hier eveneens een aantal voorbeelden van.

Voorbeelden[1]

3.16. Men trekt uit de populatie in tabel 3.1 aselect zonder teruglegging een steekproef van 2 knikkers. Welke kansverdeling volgt het aantal witte knikkers in deze steekproef?

Uit formule (3) in bijlage 2 (par. 2) volgt dat in totaal $C_2^{10} = 10!/2!8! = 10 \times 9/2 \times 1 = 45$ steekproeven van 2 knikkers kunnen worden samengesteld. Voor het aantal manieren, waarop een steekproef met 0, 1 of 2 witte knikkers tot stand kan komen vindt men met formule (4):

[1] Lees eerst bijlage 2, paragraaf 2.

0 witte knikkers:	$C_0^4 \times C_2^6 = 15$
1 witte knikker :	$C_1^4 \times C_1^6 = 24$
2 witte knikkers:	$C_2^4 \times C_0^6 = 6$
Totaal	45

De gevraagde kansverdeling van de grootheid $\underline{x}$, het aantal witte knikkers in de steekproef, is dus

x_i	0	1	2
$P(\underline{x} = x_i)$	$\frac{15}{45}$	$\frac{24}{45}$	$\frac{6}{45}$
	0,333	0,533	0,133

3.17. Een populatie bestaat uit 6 elementen, die resp. de waarden 1, 2, 3, 4, 5 en 6 dragen. Men trekt aselect zonder teruglegging een steekproef van 3 elementen. Welke kansverdeling volgt de grootheid $\underline{S}$, de som van de waarden van de elementen in de steekproef?

In tabel 3.10 zijn bij A de $C_3^6 = 6!/3!3! = (6 \times 5 \times 4)/(3 \times 2 \times 1) = 20$ mogelijke steekproeven van 3 elementen — bij trekking zonder teruglegging — systematisch opgeschreven. Het eerste vak bevat alle combinaties die met 1 en 2 beginnen, op de tweede regel die met 1 en 3, enz.), het tweede vak alle combinaties die met 2 beginnen, enz. Steekproeven die dezelfde waarde van $\underline{S}$ opleveren zijn hierbij onder elkaar geplaatst, zodat onmiddellijk kan worden vastgesteld hoe frequent elke waarde van $\underline{S}$ is (men noemt deze frequentieverdeling wel de (theoretische) *steekproevenverdeling* (Eng. sampling distribution) van de grootheid $\underline{S}$). De relatieve steekproevenverdeling bij B is de kansverdeling van $\underline{S}$.

Tabel 3.10. Kansverdeling van de grootheid $\underline{S}$, de som van de waarden van een aselect zonder teruglegging getrokken steekproef van 3 elementen, uit een populatie met de waarden 1, 2, 3, 4, 5 en 6

A. *De 20 combinaties van 3 elementen*

Som	6	7	8	9	10	11	12	13	14	15
	123	124	125	126						
			134	135	136					
					145	146				
							156			
				234	235	236				
						245	246			
								256		
							345	346		
									356	
										456
f	1	1	2	3	3	3	3	2	1	1

B. *Kansverdeling van* $\underline{S}$

S_i	6	7	8	9	10	11	12	13	14	15
$P(\underline{S} = S_i)$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{2}{20}$	$\frac{1}{20}$	$\frac{1}{20}$

3.18. Men trekt uit de populatie in tabel 3.2 aselect zonder teruglegging een steekproef van 2 ballen. Welke kansverdeling volgt het gemiddelde gewicht van deze steekproef?

Er zijn $C_2^{10} = 45$ steekproeven mogelijk en deze zijn in tabel 3.11 bij A opgeschreven, tezamen met hun som S. De steekproevenverdeling van $\underline{S}$ en $\underline{\bar{x}} = \underline{S}/2$ volgt uit tabel A door de frequentie van elke waarde van $\underline{S}$ te tellen en de relatieve steekproevenverdeling levert de gevraagde kansverdeling.

Tabel 3.11. Kansverdeling van de grootheid $\underline{\bar{x}}$, het gemiddelde gewicht van 2 ballen die aselect zonder teruglegging worden getrokken uit de populatie in tabel 3.2

A. *Steekproeven en steekproefsommen*

Trekking 1 \ Trekking 2	3	4	4	4	5	5	5	5	6	6
3		7	7	7	8	8	8	8	9	9
4			8	8	9	9	9	9	10	10
4				8	9	9	9	9	10	10
4					9	9	9	9	10	10
5						10	10	10	11	11
5							10	10	11	11
5								10	11	11
5									11	11
6										12
6										

B. *Steekproevenverdeling en kansverdeling van* $\underline{S}$ *en* $\underline{\bar{x}}$

S_i	7	8	9	10	11	12
$\bar{x}_i$	$3\frac{1}{2}$	4	$4\frac{1}{2}$	5	$5\frac{1}{2}$	6
f	3	7	14	12	8	1
$P(\underline{S} = S_i) = P(\underline{\bar{x}} = \bar{x}_i)$	3/45 0,067	7/45 0,156	14/45 0,310	12/45 0,267	8/45 0,178	1/45 0,022

3.19. Wanneer men in tabel 3.11A de spreidingsbreedten in plaats van de steekproefsommen opneemt, kan de kansverdeling van $\underline{R}$ worden berekend. Wij laten het aan de lezer over na te gaan dat deze kansverdeling als volgt luidt:

R_i	0	1	2	3
$P(\underline{R} = R_i)$	10/45 0,222	23/45 0,511	10/45 0,222	2/45 0,044

3.20. Men trekt aselect zonder teruglegging een steekproef van 3 ballen uit de populatie in tabel 3.2. Welke kansverdeling volgt het gemiddelde gewicht van deze steekproef?

Wij laten het aan de lezer over de $C_3^{10} = 120$ mogelijkheden op te schrijven en zodoende de kansverdeling in tabel 3.12 af te leiden.

Tabel 3.12. Kansverdeling van de grootheid $\bar{x}$, het gemiddelde gewicht van 3 ballen die aselect zonder teruglegging worden getrokken uit de populatie in tabel 3.2

$\bar{x}_i$	$3\frac{2}{3}$	4	$4\frac{1}{3}$	$4\frac{2}{3}$	5	$5\frac{1}{3}$	$5\frac{2}{3}$
$P(\bar{x} = \bar{x}_i)$	3/120 0,025	13/120 0,108	24/120 0,200	32/120 0,267	29/120 0,241	15/120 0,125	4/120 0,033

*Opmerking**

Trekt men aselect *zonder teruglegging* uit een populatie van N elementen met gemiddelde μ_x en variantie σ_x^2 een steekproef van n elementen, dan geldt voor de kansverdeling van de grootheid $\bar{x}$, het steekproefgemiddelde, dat

(3.11) $$\mu_{\bar{x}} = \mu_x$$

en

(3.12) $$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n} \cdot \frac{N-n}{N-1}.$$

Wij laten hier de bewijzen achterwege [men vindt deze bijv. in WILKS (1949)]. Merk op, dat voor $n \ll N$ geldt dat $N - n \approx N - 1$, dus $(N-n)/(N-1) \approx 1$, zodat formule (3.12) benaderd kan worden door formule (3.10), die geldt bij trekking met teruglegging.

3.5.4. TOEPASSING

Bij de voorbeelden in 3.5.2 en 3.5.3 zijn zowel de populaties als de steekproeven van zeer kleine omvang, zodat de kansverdelingen van de steekproeffuncties nog zonder veel moeite kunnen worden bepaald. Hierbij blijkt reeds dat de berekeningen bij trekking met teruglegging — omdat daarbij de productregel kan worden gebruikt — eenvoudiger zijn dan die bij trekking zonder teruglegging (men vergelijke bijv. de berekeningen bij de voorbeelden 3.15 en 3.20)

In de praktijk heeft men vrijwel altijd te maken met steekproeven die zonder teruglegging zijn getrokken, terwijl gewoonlijk de populatie èn de steekproefomvang (veel) groter zijn dan die bij de voor-

beelden. De exacte berekening van de kansverdeling van een steekproeffunctie (via de steekproevenverdeling, dus door het opschrijven van alle mogelijke steekproeven van de omvang n) wordt dan praktisch onuitvoerbaar. Dit blijkt uit tabel 3.13 waarin voor enkele (nog niet eens zo grote) populaties het (soms reeds formidabele) aantal steekproeven bij enkele waarden van n is opgenomen.

Tabel 3.13. Aantal mogelijke steekproeven bij trekking zonder teruglegging

Populatie-omvang N	Steekproef-omvang n	Aantal steekproeven C_n^N
10	5	252
20	5	15504
50	3	19600
50	5	2118760
100	3	161700
100	5	75287520

De zaak wordt echter reeds eenvoudiger als men bedenkt dat bij praktijkproblemen de steekproefomvang vaak (zeer) klein is met betrekking tot de populatie-omvang. Men kan dan immers de kansverdeling die geldt zonder teruglegging bevredigend benaderen door de kansverdeling die met teruglegging geldt en die eenvoudiger te berekenen is.

Verder gaat men bij het berekenen van de kansverdeling van een steekproeffunctie op één van de volgende manieren te werk:

1. Men berekent exact de kansverdeling voor een reeks waarden van N en n (bij trekking zonder teruglegging), resp. voor een reeks waarden van n (bij trekking met teruglegging). Soms kan men hierbij de berekeningen vereenvoudigen, doordat de kansverdeling voor $n + 1$ vrij eenvoudig uit die voor n kan worden afgeleid. Men bestudeert de vorm van deze verdeling en onderzoekt of, en zo ja onder welke voorwaarden, zij door een theoretische verdeling (zoals bijv. een normale verdeling) kan worden benaderd.

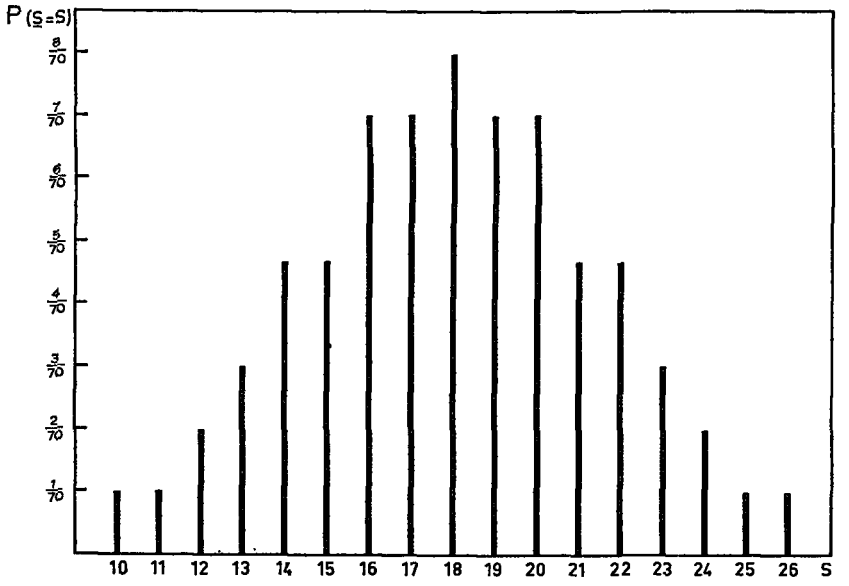
Voorbeeld

3.21. In voorbeeld 3.17 berekenden wij de kansverdeling van de grootheid $\underline{S}$ de som van de waarden van $n = 3$ elementen, die aselect worden getrokken uit een populatie van $N = 6$ elementen met de waarden 1, 2, . . ., 6. Deze kansverdeling (zie tabel 3.10) blijkt symmetrisch te zijn ten opzichte van haar gemiddelde en een zekere centrale tendentie te bezitten.

Tabel 3.14 geeft de kansverdeling van $\underline{S}$ voor $N = 8$ ($x = 1, 2, \ldots, 8$) en $n = 4$ (wij geven de lezer in overweging deze kansverdeling zelf te berekenen). Ook nu ziet men een symmetrische verdeling en deze begint, zoals uit figuur 3.2 blijkt, reeds enigszins op een normale verdeling te gelijken.

Tabel 3.14. Kansverdeling van de grootheid $\underline{S}$, de som van de waarden van een aselect zonder teruglegging getrokken steekproef van 4 elementen, uit een populatie met de waarden 1, 2, . . ., 8.

S	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
$P(\underline{S} = S)$	$\frac{1}{70}$	$\frac{1}{70}$	$\frac{2}{70}$	$\frac{3}{70}$	$\frac{5}{70}$	$\frac{5}{70}$	$\frac{7}{70}$	$\frac{7}{70}$	$\frac{8}{70}$	$\frac{7}{70}$	$\frac{7}{70}$	$\frac{5}{70}$	$\frac{5}{70}$	$\frac{3}{70}$	$\frac{2}{70}$	$\frac{1}{70}$	$\frac{1}{70}$



Figuur 3.2. Staafdiagram van de kansverdeling in tabel 3.14.

Het blijkt nu dat men reeds voor $N = 10$ ($x = 1, 2, \ldots, 10$) en $n = 5$ de kansverdeling van $\underline{S}$ bevredigend door een normale verdeling (met als gemiddelde μ_S en als spreiding σ_S) kan benaderen. Deze benadering wordt steeds beter, naarmate N en $n = \frac{1}{2}N$ groter worden. Verder blijkt zij voor grotere waarden van N en n ook op te gaan wanneer $n \neq \frac{1}{2}N$ is.

2. Men gaat uit van een populatieverdeling van een bepaald type en leidt mathematisch een formule af die de kansverdeling van de steekproeffunctie specificeert (d.w.z.: waarmee men voor een bepaalde populatieverdeling van dit type en de steekproefomvang n de kans op elke waarde van de steekproeffunctie kan berekenen).

Voorbeelden

3.22. In hoofdstuk 4 gaan wij uit van een populatieverdeling met twee categorieën (noem deze A en B), een z.g. dichotomie. Wij leiden dan formules af voor de kansverdeling van de grootheid $\underline{x}$, het aantal elementen met ken-

merk A, resp. van de grootheid $\underline{p} = \underline{x}/n$, de fractie elementen met kenmerk A, in een aselecte steekproef van n elementen, eerst bij trekking zonder en vervolgens bij trekking met teruglegging. Tevens demonstreren wij dat de — eenvoudiger te berekenen — kansverdeling bij trekking met teruglegging als benadering van die bij trekking zonder teruglegging kan worden gebruikt; mits $n \ll N$ (lees: de steekproefomvang aanzienlijk kleiner is dan de populatieomvang). Tenslotte blijkt in hoofdstuk 5 dat de eerstgenoemde verdeling onder bepaalde voorwaarden op haar beurt weer benaderd kan worden door een normale verdeling. Dit proces van successieve benaderingen ontmoet men in de statistiek voortdurend.

3.23. In hoofdstuk 13 gaan wij uit van een grootheid $\underline{x}$, die op een populatie een normale verdeling bezit met als parameters μ_x en σ_x. Wij geven daar een specificatie van de kansverdeling van het gemiddelde van een aselecte steekproef van n elementen uit deze populatie, die langs mathematische weg is afgeleid. Deze kansverdeling geldt bij benadering wanneer de populatieverdeling pseudo-normaal is.

Het komt voor dat men de kansverdeling van een steekproeffunctie niet volgens één der zojuist aangegeven methoden kan bepalen. Men neemt dan zijn toevlucht tot de empirie, d.w.z. men trekt aselect een groot aantal steekproeven van de gewenste grootte uit de te bestuderen populatie, bepaalt daaruit de *waargenomen* steekproevenverdeling van de steekproeffunctie en onderzoekt of deze door een of andere theoretische verdeling kan worden benaderd. De grote hoeveelheid werk die deze z.g. Monte Carlo methode met zich mee brengt is sinds de uitvinding van de electronische rekenmachine niet meer onoverkomelijk.

De Monte Carlo methode wordt overigens ook vaak gebruikt voor het aantonen van statistische wetmatigheden, waarvan een exact bewijs voor de student onnodig of te lastig wordt geacht. Wij gebruiken haar daartoe in 13.8.

3.5.5. OPGAVEN

3.10. Uit een populatie waarvan de elementen de waarden $x = 0, 0, 1, 1, 2$ dragen trekt men aselect een steekproef van 2 elementen. Bepaal de kansverdeling van $\underline{x}$ bij trekking (a) met teruglegging, (*b*) zonder teruglegging. Geef van beide kansverdelingen het gemiddelde en de variantie.

3.11. Als 3.10, maar voor een steekproef van 4 elementen. Teken van de kansverdelingen staafdiagrammen.

3.12. Bepaal voor een aselect zonder teruglegging getrokken steekproef van 2 letters uit de letters in MISSISSIPPI:

a. De kansverdeling van de steekproeffunctie $\underline{s}$, het aantal keren dat de letter S per steekproef voorkomt.

b. De kansverdeling van de steekproeffunctie $\underline{w}$, het aantal keren dat de letter I per steekproef voorkomt.

3.13. Een populatie bestaat uit 8 elementen met de waarden 1, 2, . . ., 8. Bepaal de kansverdeling van de grootheid $\underline{S}$, de som van de waarden van de elementen in een aselect zonder teruglegging getrokken steekproef van 3 elementen. Geef het gemiddelde en de variantie van deze kansverdeling (maak gebruik van de uitkomsten van opgave 2.12 en de formules (3.11) en (3.12), daarbij bedenkend dat $S = \Sigma x = n\bar{x}$).

3.6. Overschrijdingskansen

3.6.1. DEFINITIES

Beschouw de discreet verdeelde stochastische grootheid $\underline{x}$. Wij voeren hiervoor de volgende begrippen in:

a. De *linkse (overschrijdings)kans* van de waarde x_i is de kans, dat de grootheid $\underline{x}$ de waarde x_i of een lagere waarde aanneemt:

$$P_L(x_i) = P(\underline{x} \leqslant x_i).$$

b. De *rechtse (overschrijdings)kans* van de waarde x_i is de kans, dat de grootheid $\underline{x}$ de waarde x_i of een hogere waarde aanneemt:

$$P_R(x_i) = P(\underline{x} \geqslant x_i).$$

c. De *tweezijdige overschrijdingskans* van de waarde x_i is gelijk aan de som van de kansen van de waarden van de grootheid $\underline{x}$, die een hoogstens even grote kans bezitten als x_i. Wij gebruiken voor deze kans het symbool $P_D(x_i)$.

Is de kansverdeling van $\underline{x}$ *symmetrisch*, dan is de tweezijdige overschrijdingskans van x_i de kans, dat de grootheid $\underline{x}$ de waarde x_i of een minstens even ver van μ_x gelegen waarde aanneemt, d.w.z.
$P_D(x_i) = 2P_L(x_i)$ voor $x_i < \mu_x$ en $P_D(x_i) = 2P_R(x_i)$ voor $x_i > \mu_x$.
Wij verduidelijken dit met voorbeeld 3.25.

Voorbeelden

3.24. Gegeven is de volgende kansverdeling:

x_i	0	1	2	3	4	5	6	7	8	9
$P(\underline{x} = x_i)$	0,01	0,03	0,06	0,10	0,17	0,26	0,21	0,10	0,04	0,02

Voor deze verdeling geldt:

$P_L(2) = P(\underline{x} < 2) = P(\underline{x} = 0) + P(\underline{x} = 1) + P(\underline{x} = 2)$
$= 0{,}01 + 0{,}03 + 0{,}06 = 0{,}10,$
$P_R(2) = 1 - P_L(1) = 0{,}96,$
$P_R(6) = P(\underline{x} > 6) = 0{,}21 + 0{,}10 + 0{,}04 + 0{,}02 = 0{,}37,$
$P_D(8) = 0{,}04 + 0{,}02 + 0{,}03 + 0{,}01 = 0{,}10,$
$P_D(7) = 0{,}10 + 0{,}04 + 0{,}02 + 0{,}10 + 0{,}06 + 0{,}03 + 0{,}01 = 0{,}36.$

3.25. Gegeven is de volgende symmetrische kansverdeling:

x_i	3	4	5	6	7	8	9	10	11
$P(\underline{x} = x_i)$	0,01	0,04	0,10	0,20	0,30	0,20	0,10	0,04	0,01

Voor deze verdeling geldt:

$P_L(5) = P(\underline{x} < 5) = 0{,}10 + 0{,}04 + 0{,}01 = 0{,}15,$
$P_R(9) = P(\underline{x} > 9) = 0{,}10 + 0{,}04 + 0{,}01 = 0{,}15.$

Wegens de symmetrie van de verdeling is $\mu_x = 7$ .Wij zien dat
$P_L(5) = P_L(\mu_x - 2) = P_R(9) = P_R(\mu_x + 2)$. Hieruit volgt:
$P_D(5) = 2P_L(5) = P_D(9) = 2\,P_R(9) = 2(0{,}15) = 0{,}30.$

3.6.2. OPGAVEN

3.14. Geef voor de verdeling in voorbeeld 3.24: $P_L(3)$, $P_R(8)$, $P_D(1)$, $P_D(2)$.
3.15. Geef voor de verdeling in voorbeeld 3.25: $P_D(10)$, $P_D(8)$.

II

BEGINSELEN VAN DE MATHEMATISCHE STATISTIEK

'It may be interesting to remark here on the origins of the theory of statistics. Certain areas of biological experimentation reached a point where what are now called statistical methods were imperative if further progress was to be made. The essentials of statistical theory were then evolved by the biologists themselves. This parallels the natural history of almost any branch of abstract knowledge, but it is nevertheless curious in the case of statistics. For the theory of statistics appears to be a very natural development of the theory of probability, which is several hundred years old; somehow it was almost completely overlooked by workers in that field.'

A. M. MOOD

HOOFDSTUK 4

ENKELE KANSVERDELINGEN

4.1. Inleiding: steekproeven uit een dichotomie

In de hoofdstukken 4, 6 en 7, die in dit deel van het boek zijn samengebracht, behandelen wij de *gedachtengang van de mathematische statistiek*.[1] Wij gaan daarbij uit van een populatie met de eenvoudigste verdeling die mogelijk is, de *dichotomie*, waarbij de elementen hetzij kenmerk A, hetzij kenmerk B bezitten. Voorbeelden hiervan vormen de paren kenmerken: gevaccineerd — niet gevaccineerd, goedgekeurd — afgekeurd, genezen — gestorven, struma — geen struma.

Een dichotome populatie (die zelf ook dichotomie wordt genoemd) bestaande uit N elementen, waarvan er N_1 het kenmerk A dragen, kan worden gekarakteriseerd door de parameter $P(A) = N_1/N$, de fractie elementen met kenmerk A, daar vanzelfsprekend $P(B) = 1 - P(A)$. Gewoonlijk gebruikt men kortweg het symbool P voor $P(A)$ en het symbool Q voor $P(B)$. Een populatie bestaande uit 120355 personen, waarvan er 73417 gevaccineerd zijn, wordt dus gekarakteriseerd door de fractie[2] gevaccineerden, $P = 73417/120355 = 0{,}6100$.

In dit hoofdstuk bestuderen wij de variatie, die optreedt in steekproeven van n elementen die aselect uit een dichotomie worden getrokken. De samenstelling van zo'n steekproef kan worden gekarakteriseerd door de *steekproeffunctie* (statistische grootheid)

x = het aantal elementen met kenmerk A, of
$p = x/n$ = de fractie elementen met kenmerk A.

Volgens de in hoofdstuk 3 behandelde methoden leiden wij de *kansverdeling* (steekproevenverdeling) van de grootheden x en p af en wel in

4.2 — bij aselecte trekking zonder teruglegging,
4.3 — bij aselecte trekking met teruglegging.

Wij laten daarbij zien dat de eerstgenoemde kansverdeling (de z.g. hypergeometrische verdeling) door de laatstgenoemde (de z.g. binomiale verdeling) kan worden benaderd, indien de steekproefomvang n voldoende klein is ten opzichte van de populatie-omvang N. Deze kansverdelingen kunnen worden weergegeven in *formules*, waarmee

[1] Hoofdstuk 5 vormt ten dele een verlengstuk van hoofdstuk 4. Zie 4.5.

[2] Het percentage gevaccineerden is $100\,P = 61{,}00\%$. In de statistiek werkt men liever met de fractie P, daar deze tevens de kans op de uitkomst A bij aselecte trekking van één element vertegenwoordigt.

men de kans op een bepaalde waarde van $\underline{x}$ (resp. $\underline{p}$) voor elke gegeven dichotomie en elke steekproefomvang kan berekenen.

De kennis die wij in hoofdstuk 4 omtrent de kansverdeling van $\underline{x}$ en $\underline{p}$ verwerven, stelt ons in staat om tot een oplossing te komen van het probleem dat wij in de praktijk ontmoeten: het doen van een uitspraak omtrent de waarde van de parameter P op grond van de waarde die de grootheid $\underline{x}$ (of $\underline{p}$) voor één aselecte steekproef van n elementen heeft aangenomen. Hierbij kan men onderscheid maken tussen toetsingsproblemen en schattingsproblemen[1].

Bij een *toetsingsprobleem* wil men, op basis van de steekproefuitkomst onderzoeken (toetsen), welke van enkele onderstellingen (hypothesen) betreffende de waarde van P juist kan worden geacht. Veronderstel bijv. dat een steekproef van 100 personen, die aselect zonder teruglegging is getrokken uit een dichotome populatie van grote omvang, 38 personen bevat die gevaccineerd zijn tegen pokken. Is deze uitkomst nog te verenigen met de hypothese dat de fractie gevaccineerden in deze populatie $P = 0{,}5$ bedraagt, of kan worden geconcludeerd dat $P < 0{,}5$ is? Dit probleem — en andere problemen van dit type — kan men oplossen door middel van de statistische *toetsingstheorie*, die in hoofdstuk 6 wordt ingeleid.

Bij een *schattingsprobleem* wil men, op basis van de steekproefuitkomst, een schatting geven van de waarde van de parameter P, d.i. de waarden van P aangeven, waarmee deze uitkomst (bijv. 38 gevaccineerden in de steekproef van 100) nog te verenigen is. Dit probleem — en andere problemen van dit type — kan men oplossen door gebruik te maken van de statistische *schattingstheorie*, die in hoofdstuk 7 wordt ingeleid.

Het behoeft nauwelijks betoog dat men op grond van één aselecte steekproef slechts een uitspraak omtrent de populatie kan doen, waaraan een zekere reserve verbonden is. De toetsings- en schattingstheorie verschaffen de mogelijkheid om deze reserve te kwantificeren, d.i. om aan te geven welk risico men loopt dat de uitspraak die men doet fout is.

Voorlopig beperken wij ons tot dichotome populaties. Later gaan wij over tot toepassingen van de mathematische statistiek op populaties waarvan de verdeling minder eenvoudig is.

4.2. Hypergeometrische verdelingen

4.2.1. AFLEIDING, GEMIDDELDE EN VARIANTIE[2]

Wij beschouwen een steekproef van n elementen, die *zonder* teruglegging gekozen wordt uit een dichotomie en die x elementen met ken-

[1] In deze inleiding laten wij bewust een scherpere formulering achterwege.

[2] Bestudeer zo nodig nogmaals par. 2 van bijlage 2 en de daarbij behorende opgaven.

merk A bevat. In totaal zijn C_n^N verschillende steekproeven van n elementen mogelijk. De x elementen A kunnen uit de N_1 elementen A in de populatie op $C_x^{N_1}$ manieren worden gekozen. De $n - x$ elementen B kunnen op $C_{n-x}^{N-N_1}$ manieren tot stand komen. Een steekproef met x A's en $n - x$ B's kan dus op $C_x^{N_1} \cdot C_{n-x}^{N-N_1}$ manieren ontstaan. Bij aselecte trekking geldt volgens (3.1)

$$P(\underline{x} = x) = \frac{C_x^{N_1} \cdot C_{n-x}^{N-N_1}}{C_n^N}. \tag{4.1}$$

Met (4.1) kan men de kans op elke mogelijke waarde van $\underline{x}$ berekenen, d.w.z. deze formule specificeert een kansverdeling, die *hypergeometrische verdeling* wordt genoemd. Bewezen kan worden, dat voor deze verdeling geldt

$$\mu_x = \frac{nN_1}{N} = Pn \qquad [\text{waarin } P = P(A) = \frac{N_1}{N}], \tag{4.2}$$

en

$$\sigma_x^2 = \frac{nN_1(N - N_1)(N - n)}{N^2(N - 1)} = PQn\frac{N - n}{N - 1} \tag{4.3}$$

$$[\text{waarin } Q = P(B) = 1 - P].$$

Voorbeeld

4.1. Een populatie bestaat uit 3 jongens en 7 meisjes. Volgens (4.1) is de kansverdeling van de grootheid $\underline{x}$, het aantal jongens in een aselecte steekproef zonder teruglegging van 3 kinderen (bedenk dat $N_1 = 3$, $N = 10$, $N - N_1 = N_2 = 7$, $n = 3$, $P(J) = P = 3/10$, $P(M) = Q = 7/10$):

x_i	0	1	2	3
$P(\underline{x} = x_i)$	$\frac{C_0^3 \cdot C_3^7}{C_3^{10}} = \frac{35}{120}$	$\frac{C_1^3 \cdot C_2^7}{C_3^{10}} = \frac{63}{120}$	$\frac{C_2^3 \cdot C_1^7}{C_3^{10}} = \frac{21}{120}$	$\frac{C_3^3 \cdot C_0^7}{C_3^{10}} = \frac{1}{120}$

Uit (4.2) volgt: $\mu_x = Pn = 0{,}3(3) = 0{,}9$.

Uit (4.3) volgt: $\sigma_x^2 = PQn\frac{N - n}{N - 1} = (0{,}3)\ (0{,}7)3 \cdot \frac{10 - 3}{10 - 1} = 0{,}49$, $\sigma_x = 0{,}7$.

4.2.2. TOEPASSING

In de praktijk is een hypergeometrische verdeling bij grote omvang van de populatie en/of steekproef moeilijk hanteerbaar. Onderstel dat men een steekproef van 50 exemplaren zonder teruglegging trekt uit een populatie van 1000 exemplaren, waarbij 350 met kenmerk A en

650 met kenmerk B. De kans op x exemplaren A in zo'n steekproef is volgens (4.1):

$$P(\underline{x} = x) = \frac{C_x^{350} \cdot C_{50-x}^{650}}{C_{50}^{1000}}.$$

Stel verder dat men vraagt naar de kans op ten hoogste 10 exemplaren A, dus $P_L(10) = P(0) + P(1) + \ldots + P(10)$ (vragen van deze soort doen zich in de statistiek veel voor). Men zou dan elf kansen moeten uitrekenen, zoals

$$P(10) = \frac{C_{10}^{350} \cdot C_{40}^{650}}{C_{50}^{1000}} = \frac{\frac{350!}{10!340!} \cdot \frac{650!}{40!610!}}{\frac{1000!}{50!950!}}.$$

Het is duidelijk, dat deze procedure zoveel rekenwerk vergt dat zij praktisch onuitvoerbaar wordt.

In de praktijk trekt men meestal steekproeven zonder teruglegging waarbij de steekproefomvang in de regel klein is ten opzichte van de populatie-omvang. De elementen die tijdens het trekken uit de populatie verdwijnen brengen slechts een onbetekenende verandering in de populatiesamenstelling teweeg zodat men zonder bezwaar kan doen alsof met teruglegging getrokken is. Wij komen hier in 4.3.4 op terug.

4.2.3. OPGAVEN

4.1. Bereken μ_x en σ_x^2 uit de kansverdeling in voorbeeld 4.1 en verifieer dat de gevonden waarden overeenstemmen met de volgens (4.2), resp. (4.3) berekende.

4.2. Een bak bevat 3 zwarte en 4 witte ballen. Geef de kansverdeling van de grootheid $\underline{x}$, het aantal zwarte ballen in een aselect zonder teruglegging getrokken steekproef van 3 ballen. Bepaal μ_x en σ_x^2.

4.3. Uit een groep van 15 personen wordt een commissie van 5 personen gevormd, die omtrent een bepaald project advies moet uitbrengen. In de groep bevinden zich 4 personen, waarvan bekend is dat zij tegen de uitvoering van het project zijn. Van de overigen kan worden aangenomen, dat zij er voor zijn. Hoe groot is de kans, dat de commissie het project met meerderheid van stemmen aanbeveelt, als de commissie aselect wordt gekozen?

4.4. Een groep van 30 personen bevat tweemaal zoveel mannen als vrouwen. Men kiest aselect 5 personen uit deze groep. Hoe groot is de kans, dat zich daarin bevinden: *a.* 3 vrouwen, *b.* ten hoogste 2 vrouwen, *c.* ten minste 3 vrouwen?

4.5 Men heeft twee vazen. De eerste bevat 4 zwarte en 2 witte ballen, de tweede bevat 3 zwarte en 4 witte ballen. *a.* Men trekt aselect zonder teruglegging 2 ballen uit de eerste vaas en vervolgens 3 ballen uit de tweede vaas. Geef de kansverdeling van de grootheid $\underline{v}$, het aantal zwarte ballen in de steekproef van 5 ballen die men op deze wijze verkrijgt.

b. Men trekt aselect zonder teruglegging 2 ballen uit de eerste vaas en plaatst deze ongezien in de tweede. Vervolgens trekt men aselect zonder teruglegging 5 ballen uit de tweede vaas. Geef de kansverdeling van de grootheid $\underline{w}$, het aantal zwarte ballen in deze steekproef.

c. Men voegt alle ballen in één vaas bijeen en trekt daaruit aselect zonder teruglegging 5 ballen. Geef de kansverdeling van de grootheid $\underline{z}$, het aantal zwarte ballen in deze steekproef.

4.3. Binomiale verdelingen

4.3.1. AFLEIDING, GEMIDDELDE EN VARIANTIE

Wij beschouwen nogmaals de dichotome populatie in 4.1, die N_1 elementen met kenmerk A en N_2 $(= N - N_1)$ elementen met kenmerk B bevat. Onderstel dat hieruit een aselecte steekproef van n elementen *met* teruglegging wordt getrokken en dat deze x elementen met kenmerk A bevat.

Volgens (3.1) is bij elke aselecte trekking: $P(A) = P = N_1/N$ en $P(B) = Q = 1 - P$. De kans dat men bij trekking van n elementen eerst x elementen A en vervolgens $n - x$ elementen B verkrijgt

$$\underbrace{AAA \ldots\ldots A}_{x \text{ keer}} \qquad \underbrace{BBB \ldots\ldots B}_{n - x \text{ keer}}$$

is volgens de productregel $P^x Q^{n-x}$. De elementen kunnen echter ook in een andere volgorde verschijnen, bijv. als

$$\underbrace{ABAABAA \ldots A}_{\substack{x-2\ A\text{'s} \\ 2\ B\text{'s}}} \qquad \underbrace{BBABABB \ldots B}_{\substack{2\ A\text{'s} \\ n-x-2\ B\text{'s}}}$$

Elk dezer volgorden bezit volgens de productregel dezelfde kans $P^x Q^{n-x}$. In totaal zijn er C_x^n van deze volgorden mogelijk (zie bijlage 2, par. 2), zodat volgens de somregel

$$(4.4) \qquad P(\underline{x} = x) = C_x^n \cdot P^x Q^{n-x} = \frac{n!}{x!\ (n-x)!} P^x Q^{n-x}.$$

De kansverdeling die door (4.4) gespecificeerd wordt noemt men *binomiale verdeling*.[1] Bewezen kan worden, dat voor zo'n verdeling geldt

$$(4.5) \qquad \mu_x = Pn,$$

en

$$(4.6) \qquad \sigma_x^2 = PQn.$$

[1] De naam van deze verdeling is ontleend aan het feit, dat de kansen van de verdeling gelijk zijn aan de termen van het binomium van NEWTON. C_x^n noemt men de binomiaalcoëfficiënt.

Als steekproeffunctie kan men ook kiezen $\underline{p} = \frac{\underline{x}}{n}$, d.i. de fractie elementen met kenmerk A in de steekproef. De grootheid $\underline{p}$ volgt dezelfde kansverdeling als $\underline{x}$, maar volgens de eigenschappen 3 en 4 in 2.5.1 is

$$\mu_p = P, \tag{4.7}$$

en

$$\sigma_p^2 = \frac{PQ}{n}. \tag{4.8}$$

De binomiaalcoëfficiënten C_x^n kunnen direct worden afgelezen uit de bekende driehoek van PASCAL, waarvan wij in tabel 4.1 een deel weergeven. Op elke rij staan de coëfficiënten, die behoren bij de waarde van n, die vóór de rij is gegeven en wel zo, dat men van links naar rechts (of omgekeerd) C_x^n kan aflezen voor resp. $x = 0, 1, 2, \ldots, n$. De eerste en laatste coëfficiënt van een rij zijn steeds gelijk aan 1, de overige kunnen worden verkregen door de links en rechts erboven staande coëfficiënten van de voorgaande rij te sommeren. Elke rij kan zodoende uit de voorgaande worden afgeleid.

Tabel 4.1. Driehoek van PASCAL

n																		2^n
							1											
1						1		1										2
2					1		2		1									4
3				1		3		3		1								8
4			1		4		6		4		1							16
5		1		5		10		10		5		1						32
6	1		6		15		20		15		6		1					64
7		1		7		21		35		35		21		7		1		128
8		1		8		28		56		70		56		28		8		1	256
9	1		9		36		84		126		126		84		36		9		1	512
enz.																				

Voor $P = Q = \frac{1}{2}$ gaat (4.4) over in $P(\underline{x} = x) = C_x^n \cdot (\frac{1}{2})^x (\frac{1}{2})^{n-x}$, zodat

$$P(\underline{x} = x) = C_x^n \cdot (\tfrac{1}{2})^n = \frac{C_x^n}{2^n} = C_x^n \cdot 2^{-n} \quad (P = Q = \tfrac{1}{2}). \tag{4.9}$$

Daar $\Sigma C_x^n = 2^n$ kunnen de kansen in dit geval direct uit de driehoek van PASCAL worden afgelezen.

Voorbeelden

4.2. Uit de populatie in voorbeeld 4.1 wordt aselect met teruglegging een steekproef van 3 kinderen getrokken. De kansverdeling van $\underline{x}$, het aantal jon-

gens in de steekproef (resp. van $\underline{p}$, de fractie jongens in de steekproef), luidt volgens (4.4):

x_i	p_i	Berekening	$P(\underline{x} = x_i)$ $P(\underline{p} = p_i)$
0	0	$C_0^3 \cdot (3/10)^0\,(7/10)^3 = \frac{3!}{0!3!} \cdot \frac{1 \cdot 7^3}{10^3}$	0,343
1	1/3	$C_1^3 \cdot (3/10)^1\,(7/10)^2 = \frac{3!}{1!2!} \cdot \frac{3 \cdot 7^2}{10^3}$	0,441
2	2/3	$C_2^3 \cdot (3/10)^2\,(7/10)^1 = \frac{3!}{2!1!} \cdot \frac{3^2 \cdot 7}{10^3}$	0,189
3	1	$C_3^3 \cdot (3/10)^3\,(7/10)^0 = \frac{3!}{3!0!} \cdot \frac{3^3 \cdot 1}{10^3}$	0,027
			1,000

Uit (4.5) volgt: $\mu_x = 0{,}3\ (3) = 0{,}9$ en uit (4.7): $\mu_p = 0{,}3$.

Uit (4.6) volgt: $\sigma_x^2 = (0{,}3)\ (0{,}7)3 = 0{,}63$, uit (4.8): $\sigma_p^2 = (0{,}3)\ (0{,}7)/3 = 0{,}07$.

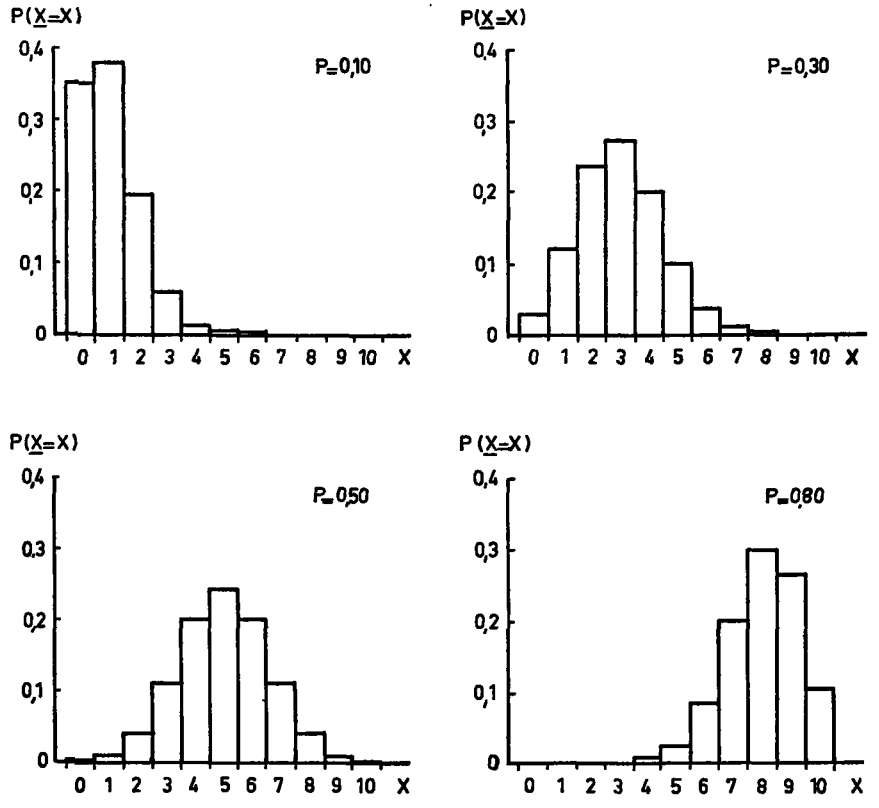
4.3. In tabel 4.2 zijn voor een aantal waarden van P de binomiale verdelingen: $P(\underline{x} = x) = P(\underline{p} = p) = C_x^{10} \cdot P^x Q^{10-x}$ opgenomen. Dit zijn dus de kansverdelingen van $\underline{x}$ (het aantal elementen met kenmerk A), resp. van $\underline{p}$ (de fractie elementen met kenmerk A) in een met teruglegging getrokken aselecte steekproef uit een dichotome populatie, waarin de fractie elementen met kenmerk A de boven de verdeling opgenomen waarde P bezit. Merk op dat alleen de verdeling met $P = \frac{1}{2} = Q$ symmetrisch is en dat voor $P \neq Q$ een asymmetrie optreedt, die sterker wordt naarmate P en Q meer verschillen.

In figuur 4.1 zijn vier binomiale verdelingen uit tabel 4.2 als kolommendiagrammen getekend. Voor $P = 0{,}1$ is de verdeling om lage waarden van $\underline{x}$ ($\underline{p}$) geconcentreerd, met (zeer) lage kansen op hoge waarden: $P(\underline{x} > 4) =$

Tabel 4.2. Enkele binomiale verdelingen met $n = 10$

x	$p = \frac{x}{10}$	P = fractie elementen met kenmerk A in de populatie							
		0,01	0,04	0,30	0,50	0,60	0,80	0,90	0,92
0	0	0,90438	0,66483	0,02825	0,00098	0,00010	0,00000	0,00000	0,00000
1	0,1	0,09135	0,27701	0,12106	0,00977	0,00157	0,00000	0,00000	0,00000
2	0,2	0,00415	0,05194	0,23347	0,04395	0,01062	0,00007	0,00000	0,00000
3	0,3	0,00011	0,00577	0,26683	0,11719	0,04247	0,00079	0,00001	0,00000
4	0,4	0,00000	0,00042	0,20012	0,20508	0,11148	0,00551	0,00014	0,00004
5	0,5	0,00000	0,00002	0,10292	0,24609	0,20066	0,02642	0,00149	0,00054
6	0,6	0,00000	0,00000	0,03676	0,20508	0,25082	0,08808	0,01116	0,00522
7	0,7	0,00000	0,00000	0,00900	0,11719	0,21499	0,20133	0,05740	0,03427
8	0,8	0,00000	0,00000	0,00145	0,04395	0,12093	0,30199	0,19371	0,14781
9	0,9	0,00000	0,00000	0,00014	0,00977	0,04031	0,26844	0,38742	0,37773
10	1	0,00000	0,00000	0,00001	0,00098	0,00605	0,10737	0,34868	0,43439

Kansen in 5 decimalen nauwkeurig. De som der kansen per verdeling kan door afronding iets van 1 afwijken.



Figuur 4.1. Kolommendiagrammen van de binomiale verdelingen met $n = 10$ en $P = 0{,}10$ 0,30, 0,50 en 0,80.

$= 0{,}01280$, $P(\underline{x} > 5) = 0{,}00164$. Men noemt een asymmetrische verdeling van dit type *scheef naar rechts*. Voor $P = 0{,}3$ is de spreiding van de verdeling groter en de scheefheid naar rechts reeds veel minder uitgesproken. Voor $P = \frac{1}{2}$, is de spreiding het grootst: de verdeling is symmetrisch met modus $= \mu_x = 5$. Voor $P = 0{,}8$ treedt *scheefheid naar links* op.

4.4. In tabel 4.3 zijn, voor dezelfde waarden van P als in tabel 4.2, de binomiale verdelingen met $n = 20$: $P(\underline{x} = x) = P(\underline{p} = p) = C_x^{20} \cdot P^x Q^{20-x}$ gegeven. Merk op dat bij deze grotere steekproefomvang voor een bepaalde $P \neq \frac{1}{2}$ de scheefheid van de binomiale verdeling minder sterk is. Wij komen hier in 5.1 op terug.

4.3.2. TABELLEN VAN BINOMIALE VERDELINGEN

Als n groot is levert het berekenen van binomiale kansverdelingen met (4.4) of (4.9) veel rekenwerk op. Dit kan althans enigszins worden

Tabel 4.3. Enkele binomiale verdelingen met $n = 20$

x	$p=\frac{x}{20}$	P = fractie elementen met kenmerk A in de populatie							
		0,01	0,04	0,30	0,50	0,60	0,80	0,90	0,92
0	0	0,81791	0,44200	0,00080	0,00000	0,00000	0,00000	0,00000	0,00000
1	0,05	0,16523	0,36834	0,00684	0,00002	0,00000	0,00000	0,00000	0,00000
2	0,10	0,01586	0,14580	0,02785	0,00018	0,00000	0,00000	0,00000	0,00000
3	0,15	0,00096	0,03645	0,07160	0,00109	0,00004	0,00000	0,00000	0,00000
4	0,20	0,00004	0,00645	0,13042	0,00462	0,00027	0,00000	0,00000	0,00000
5	0,25	0,00000	0,00086	0,17886	0,01479	0,00129	0,00000	0,00000	0,00000
6	0,30	0,00000	0,00009	0,19164	0,03696	0,00485	0,00000	0,00000	0,00000
7	0,35	0,00000	0,00001	0,16426	0,07393	0,01456	0,00001	0,00000	0,00000
8	0,40	0,00000	0,00000	0,11440	0,12013	0,03550	0,00009	0,00000	0,00000
9	0,45	0,00000	0,00000	0,06537	0,16018	0,07099	0,00046	0,00000	0,00000
10	0,50	0,00000	0,00000	0,03082	0,17620	0,11714	0,00203	0,00001	0,00000
11	0,55	0,00000	0,00000	0,01201	0,16018	0,15974	0,00739	0,00005	0,00001
12	0,60	0,00000	0,00000	0,00386	0,12013	0,17971	0,02216	0,00036	0,00008
13	0,65	0,00000	0,00000	0,00102	0,07393	0,16588	0,05455	0,00197	0,00055
14	0,70	0,00000	0,00000	0,00022	0,03696	0,12441	0,10910	0,00887	0,00316
15	0,75	0,00000	0,00000	0,00004	0,01479	0,07465	0,17456	0,03192	0,01454
16	0,80	0,00000	0,00000	0,00001	0,00462	0,03499	0,21820	0,08978	0,05227
17	0,85	0,00000	0,00000	0,00000	0,00109	0,01235	0,20536	0,19012	0,14144
18	0,90	0,00000	0,00000	0,00000	0,00018	0,00309	0,13691	0,28518	0,27109
19	0,95	0,00000	0,00000	0,00000	0,00002	0,00049	0,05765	0,27017	0,32816
20	1	0,00000	0,00000	0,00000	0,00000	0,00004	0,01153	0,12158	0,18869

verminderd door gebruik te maken van een zg. *recursieformule*, die de verhouding van twee opeenvolgende kansen geeft:

$$(4.10) \qquad \frac{P(\underline{x} = x+1)}{P(\underline{x} = x)} = \frac{C_{x+1}^{n} \cdot P^{x+1} Q^{n-x-1}}{C_{x}^{n} \cdot P^{x} Q^{n-x}} = \frac{n-x}{x+1} \cdot \frac{P}{Q}.$$

Men berekent één der kansen (bij voorkeur de grootste) met (4.4) of (4.9). De overige kansen volgen successievelijk uit de verhoudingsgetallen.

In vele gevallen kan men echter gebruik maken van tabellen, die voor verschillende waarden van P en n de (cumulatieve) binomiale verdelingen direct geven.

a. Voor waarden van n tot 150 en $P = 0{,}01(0{,}01)0{,}99$[1]: Office of Chief of Ordnance (1952).

b. Iets minder gedetailleerd voor waarden van n tot 150 en geselecteerde waarden van n tot 1000: Harvard University (1955). Deze tabellen bevatten ook verdelingen voor waarden van P die een veelvoud van 1/12 en 1/16 zijn.

c. Voor $n = 2(1)49$ en $P = 0{,}01(0{,}01)0{,}50$: National Bureau of Standards (1950).

d. Voor $n = 50(5)100$ en dezelfde waarden van P als onder *c*: Romig (1953).

[1] Dit wil zeggen: vanaf 0,01 met stappen van 0,01 opklimmend tot en met 0,99.

e. Voor $P = \frac{1}{2}$ en n t/m 200: VAN WIJNGAARDEN (1950).
f. Voor $n = 5(5)30$ en verschillende waarden van P: PEARSON & HARTLEY (1954), tabel 37.

Binomiale kansen kunnen soms betrekkelijk eenvoudig bij benadering worden berekend. Wij behandelen twee benaderingsmethoden: 1. De POISSON-benadering, die kan worden toegepast als P (zeer) klein en n (zeer) groot is (zie 4.4.1). 2. De normale benadering die bruikbaar is als n (vrij) groot is en P niet te veel van $\frac{1}{2}$ verschilt (zie 5.5.4).

4.3.3.* AANPASSING VAN EEN BINOMIALE VERDELING

Het kan voorkomen, dat men wil onderzoeken of een bepaalde populatie binomiaal verdeeld is. Als men over een aselecte steekproef uit deze populatie beschikt, kan de daarbij waargenomen verdeling worden vergeleken met de zg. best-passende binomiale verdeling, d.i. de binomiale verdeling waarvan μ_x gelijk wordt gesteld aan $\bar{x}$, het gemiddelde van de waargenomen verdeling. Daar $\mu_x = Pn$, stelt men $\bar{x} = Pn$, zodat $P = \bar{x}/n$ uit de best-passende verdeling geschat kan worden.

Voorbeeld

4.5. Men beschikt over een — als aselect te beschouwen — steekproef, bestaande uit 103 nesten, elk met 4 pasgeboren muizen. Als per nest de kans op een vrouwtjesgeboorte constant ($= P$) is volgt het aantal vrouwtjes per nest, $\underline{x}$, de binomiale verdeling

$$P(\underline{x} = x) = C_x^4 \cdot P^x Q^{4-x}.$$

De parameter P is onbekend, maar wordt geschat uit de waargenomen verdeling van het aantal vrouwtjes per nest (links in tabel 4.4). Voor deze verdeling is: $\bar{x} = 192/103 = 1{,}864$. Stellen wij $\mu_x = \bar{x}$ dan is $P \times 4 = 1{,}864$, dus $P = 1{,}864/4 = 0{,}466$. De best-passende binomiale verdeling is:

$$P(\underline{x} = x) = C_x^4 \cdot (0{,}466)^x\,(0{,}534)^{4-x}.$$

Deze is in tabel 4.4 berekend. Vermenigvuldigt men de kansen van deze verdeling met het totale aantal nesten, 103, dan verkrijgt men de verwachte frequenties, e, die met de waargenomen frequenties, f, kunnen worden vergeleken. Wij zullen later een toetsingstechniek behandelen, waar de overeenstemming tussen de waargenomen en de best-passende verdeling mee kan worden onderzocht, de zg. χ^2-toets voor aanpassing, zie 10.1. Zonder meer is hier echter wel in te zien dat deze overeenstemming zeer goed is.

Tabel 4.4. Muizennesten

Waargenomen verdeling			Best-passende binomiale verdeling	
Aantal vrouwtjes x	Aant. nesten f	fx	$C_x^4 \cdot (0{,}466)^x\,(0{,}534)^{4-x}$	$e = 103 \cdot P(\underline{x} = x)$
0	9	0	$1\,(0{,}534)^4\,(0{,}466)^0 = 0{,}0813$	8,4
1	30	30	$4\,(0{,}534)^3\,(0{,}466)^1 = 0{,}2838$	29,2
2	35	70	$6\,(0{,}534)^2\,(0{,}466)^2 = 0{,}3715$	38,3
3	24	72	$4\,(0{,}534)^1\,(0{,}466)^3 = 0{,}2162$	22,3
4	5	20	$1\,(0{,}534)^0\,(0{,}466)^4 = 0{,}0472$	4,8
Totaal	103	192	1,0000	103,0

4.3.4. BINOMIALE BENADERING VAN EEN HYPERGEOMETRISCHE VERDELING

Als men aselect *zonder* teruglegging een steekproef van n elementen trekt uit een dichotome populatie, waarvan de *omvang N aanzienlijk groter is dan n*, veranderen P en Q niet noemenswaard door het niet terugleggen der getrokken elementen. De hypergeometrische verdeling kan dan zeer nauwkeurig worden benaderd door de binomiale verdeling met $P = N_1/N$ en $Q = 1 - P$ [vergelijk ook (4.3) met (4.6): bij constante n en toenemende N nadert de term $(N - n)/(N - 1)$ in (4.3) tot 1].

Voorbeeld

4.6. Een populatie bestaat uit 120 jongens en 80 meisjes. Men trekt aselect zonder teruglegging een steekproef van 3 kinderen. De grootheid $\underline{x}$, het aantal jongens in de steekproef, volgt de hypergeometrische verdeling:

$$P(\underline{x} = x) = \frac{C_x^{120} \cdot C_{3-x}^{80}}{C_3^{200}}.$$

Daar $n = 3$ aanzienlijk kleiner is dan $N = 200$, kunnen wij verwachten dat deze verdeling bevredigend wordt benaderd door de binomiale verdeling:

$$P(\underline{x} = x) = C_x^3 \left(\frac{120}{200}\right)^x \left(\frac{80}{200}\right)^{3-x}.$$

Beide verdelingen zijn, in vier decimalen nauwkeurig, in onderstaande tabel gegeven. De benadering blijkt inderdaad zeer goed te voldoen.

	x	3	2	1	0
Exact: hypergeometrisch	$P(\underline{x} = x)$	0,2138	0,4349	0,2887	0,0626
Benaderd: binomiaal	$P(\underline{x} = x)$	0,2160	0,4320	0,2880	0,0640

4.3.5. OPGAVEN

4.6. Geef – zo mogelijk – voor elk der onderstaande steekproeffuncties:

$$P(\underline{x} = x),\ \mu_{\underline{x}} \text{ en } \sigma_{\underline{x}}^2.$$

a. Het aantal 'zessen' bij 50 worpen met een zuivere dobbelsteen.
b. Het aantal harten in een zonder teruglegging getrokken aselecte steekproef van 8 kaarten uit een spel van 52 kaarten.
c. Het aantal jongens in een gezin met drie kinderen, aselect getrokken uit een populatie waarin de kans op een jongensgeboorte 0,515 is.
d. Het aantal kinderen met ziekte A in een aselecte steekproef van 100 kinderen, getrokken uit een populatie waarin de kans op deze ziekte 0,08 is.
e. Het gemiddelde van de bloeddruk van een steekproef van 20 personen die aselect getrokken is uit een populatie van 130.000 personen.
f. Het aantal personen met een bloeddruk hoger dan 150 mm Hg in de onder *e* genoemde steekproef.

g. Het aantal defecte exemplaren in een steekproef van 10 exemplaren, die aselect zonder teruglegging wordt getrokken uit een partij van 20.000 stuks, waarin zich $\frac{1}{2}\%$ defecten bevindt.

4.7. Kunnen de volgende hypergeometrische verdelingen door binomiale worden benaderd?

a. $P(\underline{x} = x) = \dfrac{C_{x}^{100} \cdot C_{10-x}^{200}}{C_{10}^{300}}$ *b.* $P(\underline{x} = x) = \dfrac{C_{x}^{15} \cdot C_{11-x}^{20}}{C_{11}^{35}}$

4.8. Men werpt met 6 zuivere geldstukken. Geef de kansverdeling van $\underline{x}$, het aantal malen kruis bij één worp.

4.9. Men werpt 24 keer met twee zuivere dobbelstenen. Hoe groot is de kans dat minstens één keer dubbel zes verschijnt?

4.10. Voor een bepaalde populatie van zeer grote omvang is de kans, dat een persoon binnen een jaar sterft gelijk aan 0,01. Men trekt aselect een steekproef van 10 personen. Hoe groot is de kans, dat na 1 jaar het aantal sterfgevallen hierin bedraagt: *a.* Geen, *b.* Eén, *c.* Niet meer dan één, *d.* Meer dan één, *e.* Minstens één, *f.* Ten hoogste twee?

4.11. Van een binomiale verdeling is gegeven $\mu_x = 12{,}8$ en $\sigma_x = 3{,}2$. Geef (zonder deze kans uit te rekenen) $P(\underline{x} = 0)$.

4.12. Bereken met formule (4.10) de kans op 0, 1, . . ., 4 zessen bij het werpen met 4 zuivere dobbelstenen.

4.13. Bij cavia's heeft men bij 320 drielingworpen de onderstaande frequenties voor het aantal mannetjes (x) waargenomen. Pas bij deze verdeling een binomiale verdeling aan.

x	3	2	1	0
f	46	106	130	38

4.14. Een aselect getrokken steekproef vertoont onderstaande verdeling. Men vermoedt, dat deze steekproef uit een binomiaal verdeelde populatie afkomstig is. Acht u dit vermoeden juist?

x	0	1	2	3
f	71	232	605	92

4.15. Bepaal voor de binomiale verdeling met $n=20$ en $P=0{,}3$ (zie tabel 4.3): *a.* $P_L(5)$, *b.* $P_R(16)$, *c.* $P_R(13)$, *d.* $P_D(2)$, *e.* $P_D(13)$.

4.16. Bepaal voor de binomiale verdeling met $n = 20$ en $P = 0{,}5$ (tabel 4.3): *a.* $P_L(3)$, *b.* $P_R(17)$, *c.* $P_D(3)$, *d.* $P_D(17)$, *e.* $P_D(16)$, *f.* $P_L(6)$, *g.* $P_D(14)$, *h.* $P_R(15)$.

4.17. Bepaal vervolgens voor deze verdeling de waarde van $\underline{x}$, waarvoor geldt: *a.* $P_L(x) < 0{,}10$, *b.* $P_R(x) < 0{,}05$, *c.* $P_D(x) < 0{,}05$, *d.* $P_D(x) < 0{,}01$.

4.4.* Poisson-verdelingen

4.4.1. EEN POISSON-VERDELING ALS BENADERING VAN EEN BINOMIALE VERDELING

Formule (4.4) voor een binomiale kansverdeling kan ook als volgt worden geschreven (daar $\mu_x = \mu = Pn$, is $P = \mu/n$):

$$C_x^n \cdot P^x Q^{n-x} = \frac{n!}{x!\,(n-x)!} \cdot P^x Q^{n-x}$$

$$= \frac{n(n-1)\ldots(n-x+1)}{x!} \cdot \left(\frac{\mu}{n}\right)^x \cdot \left(1-\frac{\mu}{n}\right)^{n-x}$$

$$= \frac{n}{n} \cdot \frac{(n-1)}{n} \ldots \frac{(n-x+1)}{n} \cdot \frac{\mu^x}{x!}\left(1-\frac{\mu}{n}\right)^n \left(1-\frac{\mu}{n}\right)^{-x}.$$

Laat nu n tot oneindig en P tot nul naderen op zodanige wijze, dat hun product $Pn = \mu$ constant blijft. Als n nadert tot oneindig, nadert elk der termen $(n-1)/n$, $(n-2)/n$, ..., $(n-x+1)/n$ tot 1, zodat hun product 1 tot limiet heeft. Ook de term $(1-\mu/n)^{-x}$ nadert tot 1, daar μ/n tot nul nadert en $(1-\mu/n)$ dus 1 tot limiet heeft. Wij vermelden zonder bewijs dat de term $(1-\mu/n)^n$ nadert tot $e^{-\mu}$ [$e = 1/0! + 1/1! + 1/2! + \ldots = 2{,}7183\ldots$ is een bekend getal in de wiskunde: het grondtal der natuurlijke logarithmen]. Zodoende wordt de limiet van de gehele uitdrukking (4.4)

$$P(\underline{x} = x) = \frac{\mu^x e^{-\mu}}{x!}. \tag{4.11}$$

Men noemt deze kansverdeling een POISSON-*verdeling*. Zij kan als benadering van de binomiale verdeling met $\mu = Pn$ worden gebruikt, als n *groot* en P *klein* is en uit de wijze waarop zij is afgeleid blijkt reeds, dat de benadering beter zal zijn, naarmate n groter en P kleiner is. Een demonstratie hiervan is te vinden in tabel 4.5, waarin enkele binomiale verdelingen met $\mu = Pn = 3$ (en met verschillende waarden van P en n), zijn vergeleken met de POISSON-verdeling met $\mu = 3$. Men ziet dat reeds voor $P = 0{,}1$ en $n = 30$ een redelijke benadering verkregen wordt, dat voor $P = 0{,}02$ en $n = 150$ de benadering goed en voor $P = 0{,}01$ en $n = 300$ de benadering uitstekend is.

Tabel 4.5. Enkele binomiale verdelingen en de POISSON-verdeling met $\mu = Pn = 3$

x	Binomiale verdelingen				POISSON-verdeling
	$P = 0{,}1$ $n = 30$ $Pn = 3$	$P = 0{,}05$ $n = 60$ $Pn = 3$	$P = 0{,}02$ $n = 150$ $Pn = 3$	$P = 0{,}01$ $n = 300$ $Pn = 3$	$Pn = \mu = 3$
0	0,0424	0,0461	0,0483	0,0490	0,0498
1	0,1413	0,1455	0,1478	0,1486	0,1494
2	0,2276	0,2259	0,2248	0,2244	0,2240
3	0,2361	0,2298	0,2263	0,2252	0,2240
4	0,1771	0,1724	0,1697	0,1689	0,1680
5	0,1023	0,1016	0,1011	0,1010	0,1008
6	0,0474	0,0490	0,0499	0,0501	0,0504
7	0,0180	0,0199	0,0209	0,0213	0,0216
8	0,0058	0,0069	0,0076	0,0079	0,0081
9	0,0016	0,0021	0,0025	0,0026	0,0027
10	0,0004	0,0006	0,0007	0,0008	0,0008
11	0,0001	0,0001	0,0002	0,0002	0,0002
12	0,0000	0,0000	0,0000	0,0000	0,0001
enz.					

Uit het voorgaande blijkt dat het gemiddelde van een POISSON-verdeling $\mu = Pn$ is. De variantie is eveneens gelijk aan Pn; als $n \to \infty$ wordt P steeds kleiner, zodat Q nadert tot 1 en $PQn \to Pn$. Dus:

$$\mu = Pn = \sigma^2. \tag{4.12}$$

4.4.2. BEREKENING

De recursieformule (4.10) voor de verhouding van twee opvolgende kansen gaat voor een POISSON-verdeling over in

$$\frac{P(\underline{x} = x + 1)}{P(\underline{x} = x)} = \frac{\mu}{x + 1}, \tag{4.13}$$

of:

$$P(\underline{x} = x + 1) = P(\underline{x} = x) \cdot \frac{\mu}{x + 1}.$$

Als μ bekend is kan $P(\underline{x} = 0)$ — wij schrijven kortweg $P(0)$ — berekend worden met

$$P(0) = \frac{1}{e^{\mu}} = e^{-\mu}, \tag{4.14}$$

zodat bij benadering

$$\log P(0) = -0{,}4343\,\mu. \tag{4.15}$$

De overige kansen volgen door toepassing van (4.13):

$$P(1) = P(0) \cdot \frac{\mu}{1} = P(0) \cdot \mu, \quad P(2) = P(1) \cdot \frac{\mu}{2} = P(0) \cdot \frac{\mu^2}{2!},$$

dus in het algemeen uit

$$P(x) = P(0) \cdot \frac{\mu^x}{x!}. \tag{4.16}$$

In tabel 4.6 zijn voor een aantal waarden van μ de bijbehorende waarden van $e^{-\mu} = P(0)$ gegeven. Deze tabel kan worden gebruikt bij het maken van opgaven betreffende POISSON-verdelingen.

Tabel 4.6. Waarden van $e^{-\mu} = P(0)$

μ	$e^{-\mu}$	μ	$e^{-\mu}$	μ	$e^{-\mu}$	μ	$e^{-\mu}$	μ	$e^{-\mu}$
0,1	0,9048	0,6	0,5488	1,2	0,3012	2,5	0,08208	5,0	0,006738
0,2	0,8187	0,7	0,4966	1,4	0,2466	3,0	0,04979	6,0	0,002479
0,3	0,7408	0,8	0,4493	1,6	0,2019	3,5	0,03020	7,0	0,000912
0,4	0,6703	0,9	0,4066	1,8	0,1653	4,0	0,01832	8,0	0,000335
0,5	0,6065	1,0	0,3679	2,0	0,1353	4,5	0,01111	9,0	0,000123

In de tabellen van MOLINA (1942) zijn (cumulatieve) POISSON-verdelingen met μ tot 100 opgenomen. PEARSON & HARTLEY (1954) bevatten POISSON-verdelingen voor $\mu = 0{,}1(0{,}1)15{,}0$.

Voorbeeld

4.7. Benadering van de binomiale verdeling $P(\underline{x} = x) = C_n^x \cdot (0{,}1)^x\,(0{,}9)^{n-x}$ door de POISSON-verdeling: $P(\underline{x} = x) = (0{,}3)^x e^{-0{,}3}/x!$

Berekening: $\mu = Pn = (0{,}1)3 = 0{,}3$. Uit tabel 4.6 volgt $P(0) = 0{,}7408$. Uit (4.16) volgt: $P(1) = 0{,}7408(0{,}3) = 0{,}2222$; $P(2) = 0{,}7408(0{,}3)^2/2! = 0{,}0333$; $P(3) = 0{,}7408(0{,}3)^3/3! = 0{,}0033$.

Deze benaderende POISSON-verdeling is in onderstaande tabel met de binomiale vergeleken:

	x	0	1	2	3	Totaal
Binomiaal	$P(\underline{x} = x)$	0,729	0,243	0,027	0,001	1,000
POISSON	$P(\underline{x} = x)$	0,741	0,222	0,033	0,003	0,999

4.4.3. POISSON-VERDELINGEN ALS EXACTE KANSVERDELINGEN

Tot dusverre hebben wij POISSON-verdelingen behandeld als benadering voor een bepaald soort binomiale verdeling. Er zijn echter situaties,

waarvoor een POISSON-verdeling de *exacte* kansverdeling is, nl. als men te doen heeft met het voorkomen van onderling onafhankelijke, geïsoleerde gebeurtenissen in een continuum van tijd of ruimte. Het totale aantal mogelijke gebeurtenissen dient groot (feitelijk: oneindig groot) te zijn, maar de (constante) kans op het voorkomen van een individuele gebeurtenis moet klein zijn. Beschouwt men bijv. het aantal keren x dat een bepaalde gebeurtenis in een tijdsinterval van zekere lengte voorkomt, dan is het kenmerkende van de situatie dat men wel kan vaststellen hoe vaak de gebeurtenis heeft plaatsgevonden, maar dat het geen zin heeft te vragen naar het aantal malen dat de gebeurtenis niet voorkwam.

Voorbeelden

4.8. Het aantal leucocyten of erythrocyten in één vierkantje van een hemocytometer (zie N.T.v.G. 98 (1954), 3480—3485).

4.9. Het aantal sterfgevallen per dag tengevolge van een acute hartaandoening in een grote stad.

4.10. Het aantal auto's, dat een bepaald punt van een weg gedurende een bepaalde minuut van de dag passeert.

4.11. Het aantal aanvragen, dat een schakelbord in een telefooncentrale gedurende een bepaalde minuut van de dag te verwerken krijgt.

4.12. Het aantal garenbreuken per spindel (zie Sigma I, (1955), 78—81).

4.4.4. AANPASSING VAN EEN POISSON-VERDELING

Als men beschikt over de waargenomen frequentieverdeling van een grootheid en wil onderzoeken of deze een POISSON-verdeling volgt, kan de best-passende POISSON worden verkregen door $\bar{x}$ te berekenen en μ hieraan gelijk te stellen (bewezen kan worden, dat deze handelwijze te verkiezen is boven het schatten van μ uit s_x^2 of uit een combinatie van $\bar{x}$ en s_x^2).

Voorbeeld

4.13. Wij bepalen de best-passende POISSON-verdeling voor onderstaande verdeling van het aantal gistcellen, dat werd waargenomen in 400 vierkantjes van een hemocytometer (gegevens van STUDENT, Biometrika 1907):

x	0	1	2	3	4	5	6	7	8	Totaal
f	103	143	98	42	8	4	2	0	0	400
fx	0	143	196	126	32	20	12	0	0	$529 = \Sigma fx$

Berekeningen:

$\bar{x} = 529/400 = 1{,}3225$. Wij stellen dus: $\mu = 1{,}3225$. Volgens (4.15) is $\log P(0) = -0{,}4343\ (1{,}3225) = -0{,}57436 = 0{,}42564 - 1$. Hieruit volgt $P(0) = 0{,}2666$. Met (4.16) volgen hieruit de andere kansen, bijv.:
$P(1) = 0{,}2666\ (1{,}3225) = 0{,}3526$; $P(2) = 0{,}2666\ (1{,}3225)^2/2 = 0{,}2331$; $P(3) = 0{,}2666\ (1{,}3225)^3/6 = 0{,}1028$, enz.

(Merk op dat men in dit geval de best-passende binomiale verdeling niet kan bepalen, daar men dan $P = \bar{x}/n$ zou moeten stellen terwijl de waarde van n niet vaststaat.)

De best-passende POISSON-verdeling is:

x	0	1	2	3	4	5	6	7
$P(\underline{x} = x)$	0,2666	0,3526	0,2331	0,1028	0,0340	0,0090	0,0020	0,0004
$e = 400 \times P$	106,6	141,0	93,2	41,1	13,6	3,6	0,8	0,0

4.4.5. OPGAVEN

4.18. Van een POISSON-verdeling is gegeven:
a. $P(3) = P(4)$. Hoe groot is de kans, dat $\underline{x}$ de waarden 0 of 1 aanneemt?
b. $P(1) = 3P(4)$. Bepaal μ en $P(0)$.

4.19. Men werpt 64 keer met 5 geldstukken. Geef (bij benadering) de kans op 0, 1, 2, 3, . . ., 64 keer de uitkomst '5 kruis'.

4.20. Een massaproduct, dat 0,5% defecten vertoont, wordt verpakt in dozen van 100 stuks. Welk deel der dozen bevat 0 defecte exemplaren? Welk deel bevat er 2 of meer?

4.21. Een verzekeringsmaatschappij constateert, dat 0,004% van de bevolking per jaar sterft tengevolge van een bepaald ongeval. Hoe groot is de kans, dat deze maatschappij in één jaar tot uitkering zal moeten overgaan bij meer dan 3 van de 10.000 tegen het risico van dit ongeval gesloten verzekeringen?

4.22. Onderstaande tabel geeft het aantal sterfgevallen per week bij personen van 20 t/m 64 jaar aan kanker en andere maligne tumoren te 's-Gravenhage in de jaren 1920 t/m 1928. Bepaal de best-passende POISSON-verdeling.

x	0	1	2	3	4	5	6	7	8	9	10	11	12	Totaal
f	16	50	73	101	71	70	47	24	10	3	1	—	1	467

4.5. Normale verdelingen

Bepaalde binomiale verdelingen kunnen (zeer) goed worden benaderd door normale verdelingen. Normale verdelingen spelen een zeer belangrijke rol in de statistiek. Niet alleen worden zij voor het benaderen van vele andere kansverdelingen gebruikt, maar tevens berusten de zg. klassieke statistische toetsings- en schattingsmethoden op grootheden, waarvan de populatieverdeling normaal is. Wij wijden daarom het gehele volgende hoofdstuk aan normale verdelingen en daarbij komt tevens (in 5.4) de benadering van een binomiale door een normale verdeling aan de orde.

4.6. Samenvatting

In de voorgaande paragrafen van dit hoofdstuk bestudeerden wij de kansverdelingen van de grootheden $\underline{x}$, resp. $\underline{p} = \underline{x}/n$, dat zijn het aantal, resp. de fractie elementen met kenmerk A in een aselecte steekproef van n elementen uit een populatie, waarin de fractie elementen met kenmerk A gelijk is aan P.

In de praktijk trekt men gewoonlijk een steekproef zonder teruglegging. Wij hebben gezien dat de grootheden $\underline{x}$ en $\underline{p}$ dan een hypergeometrische verdeling volgens formule (4.1) volgen. De berekening van zo'n verdeling is voor steekproeven en populaties van grotere omvang zeer tijdrovend. Indien echter de steekproefomvang n klein is ten opzichte van de populatieomvang N (zodat tenminste geldt dat $N \geqslant 50n$) kan men een hypergeometrische verdeling goed benaderen met een binomiale verdeling, d.i. de kansverdeling van $\underline{x}$ en $\underline{p}$ die exact geldt wanneer met teruglegging getrokken wordt.

Binomiale verdelingen zijn eenvoudiger te berekenen dan hypergeometrische. Zij zijn voor vele waarden van $n \leqslant 100$ en een reeks waarden van P getabelleerd. Verder kan men indien n groot en P klein is een binomiale verdeling benaderen met een POISSON-verdeling. Tenslotte zullen wij in 5.4 zien dat in vele gevallen waarin de binomiale tabellen verstek laten gaan, nl. indien $n > 100$ is, een binomiale verdeling (met een minimum aan rekenwerk) benaderd kan worden door een normale verdeling.

De voorgaande procedure wordt in de mathematische statistiek herhaaldelijk gevolgd: men leidt eerst de exacte kansverdeling van een steekproeffunctie af en onderzoekt vervolgens of, en zo ja onder welke voorwaarden, deze door een bekende en reeds getabelleerde kansverdeling kan worden benaderd.

HOOFDSTUK 5

NORMALE VERDELINGEN

5.1. Een normale verdeling als limiet van een binomiale

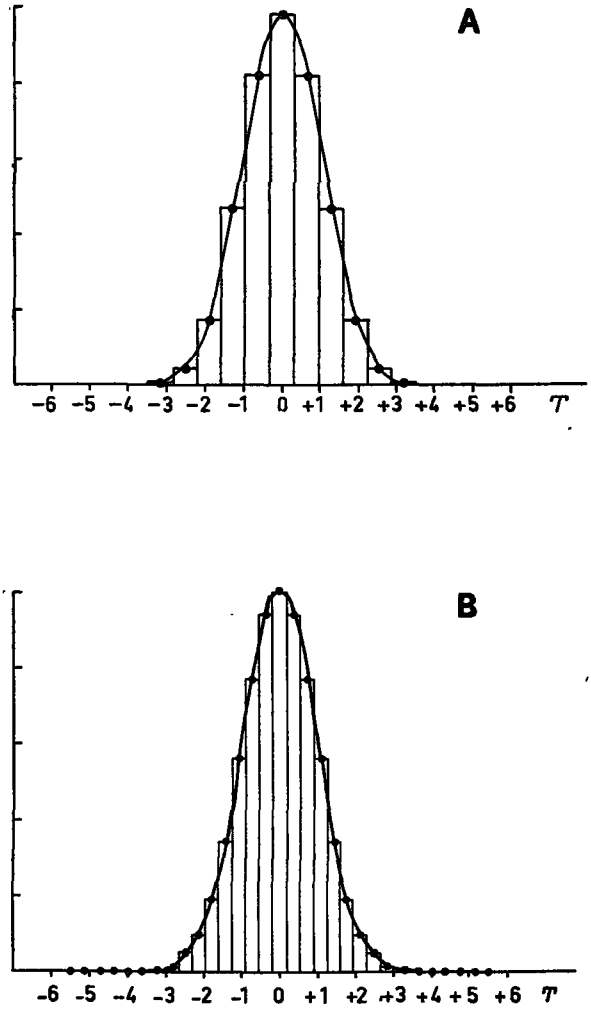
Beschouw een reeks binomiale verdelingen met $P = \frac{1}{2}$ en met in grootte toenemende waarden van n. Deze verdelingen zijn alle symmetrisch. Naarmate n toeneemt wordt het aantal mogelijke waarden van $\underline{x}$ groter; de som van de kansen van de $(n + 1)$ waarden van $\underline{x}$ is echter steeds gelijk aan 1, zodat de kansen op de afzonderlijke waarden van $\underline{x}$ geleidelijk afnemen (zie bijv. de verdelingen met $P = \frac{1}{2}$ en $n = 10$, resp. 20 in de tabellen 4.2 en 4.3). De steeds breder en vlakker wordende kansverdelingen zijn daardoor niet goed vergelijkbaar. Dit bezwaar kan worden ondervangen door elke verdeling te standaardiseren, d.w.z. door elke binomiaal verdeelde grootheid $\underline{x}$ met gemiddelde $Pn = \frac{1}{2}n$ en standaardafwijking $\sqrt{PQn} = \frac{1}{2}\sqrt{n}$ volgens (3.4) om te zetten in de gestandaardiseerde (binomiale) grootheid

$$\underline{T}_n = \frac{\underline{x} - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} \tag{5.1}$$

met gemiddelde 0 en standaardafwijking 1.

In figuur 5.1 zijn voor twee van deze gestandaardiseerde binomiale verdelingen met $P = \frac{1}{2}$ en $n = 10$, resp. 30, de kolommendiagrammen getekend. Per verdeling is daartoe van elke waarde van $\underline{x}$ met (5.1) de corresponderende waarde van $\underline{T}$ bepaald en vervolgens is op elk der $(n + 1)$ klassen met T_n als klassemidden een kolom opgericht, waarvan het oppervlak (= klassebreedte × hoogte) de grootte van de bij T_n behorende kans representeert. De som van de oppervlakken der $(n + 1)$ kolommen dient gelijk te zijn aan 1 en daar de klassebreedte van de twee verdelingen verschillend is (deze bedraagt nl. $1/\frac{1}{2}\sqrt{n}$) moet de hoogte van elke kolom gelijk zijn aan

$$Y_n = \frac{P(\underline{x} = x)}{\frac{1}{\frac{1}{2}\sqrt{n}}} = \frac{P(\underline{x} = x)\sqrt{n}}{2}. \tag{5.2}$$



Figuur 5.1. Twee gestandaardiseerde binomiale verdelingen met de best-passende normale verdeling, $P = Q = \frac{1}{2}$:

$$\text{A: } n = 10, T = \frac{x - \frac{1}{2}(10)}{\frac{1}{2}\sqrt{10}} = \frac{x - 5}{1{,}58}, \text{klassebreedte: } \frac{1}{\frac{1}{2}\sqrt{10}} = 0{,}633,$$

$$\text{B: } n = 30, T = \frac{x - \frac{1}{2}(30)}{\frac{1}{2}\sqrt{30}} = \frac{x - 15}{2{,}74}, \text{klassebreedte: } \frac{1}{\frac{1}{2}\sqrt{30}} = 0{,}365.$$

Laat men nu n steeds groter worden dan wordt het aantal kolommen ook steeds groter, maar de kolombreedte wordt voortdurend kleiner. Nadert n tot oneindig, dan nadert het aantal kolommen eveneens tot oneindig en de kolombreedte (= klassebreedte) nadert tot nul. De trappenlijn van het kolommendiagram gaat dan over in de vloeiende gebogen lijn, die in figuur 5.1 reeds is ingeschetst (en die bij $n = 30$ al aanzienlijk beter 'past' dan bij $n = 10$). Bewezen kan worden, dat deze curve een zeer bekende continue kansverdeling, een zg. *normale verdeling* (verdeling van DE MOIVRE, verdeling van GAUSS-LAPLACE) representeert. Bij $P = \frac{1}{2}$ ziet men dus, als n nadert tot oneindig, een normale verdeling verschijnen als limiet van de binomiale verdeling.

Voor $P \neq \frac{1}{2}$ is een binomiale verdeling asymmetrisch (scheef naar rechts voor $P < \frac{1}{2}$, scheef naar links voor $P > \frac{1}{2}$, zie figuur 4.1). Voor een bepaalde n is de scheefheid sterker, naarmate P meer van $\frac{1}{2}$ verschilt (vergelijk bijv. de verdelingen met $n = 10$ en verschillende P in tabel 4.2). Maar bij een bepaalde P ($\neq \frac{1}{2}$) neemt de scheefheid af, naarmate n groter wordt (vergelijk bijv. de binomiale verdelingen met $P = 0{,}3$ en $n = 10$, resp. 20, in de tabellen 4.2 en 4.3). Het blijkt dat ook voor $P \neq \frac{1}{2}$ een binomiale verdeling bij aangroeiende n tot een symmetrische normale verdeling nadert. Een duidelijk voorbeeld hiervan geeft figuur 5.2, waarin de binomiale verdelingen met $P = 0{,}37$ en $n = 5$, 10, 20, 40, 80 en 160 zijn getekend[1]. De scheefheid naar rechts, die bij kleine n zeer duidelijk is, verdwijnt vrijwel geheel naarmate n groter wordt. Het naderen van een (scheve) binomiale verdeling tot een (symmetrische) normale verdeling verloopt sneller, naarmate P minder van $\frac{1}{2}$ afwijkt. Is P bijv. zeer klein (0,001), dan is de benadering pas goed bij zeer grote n (ongeveer 10.000).

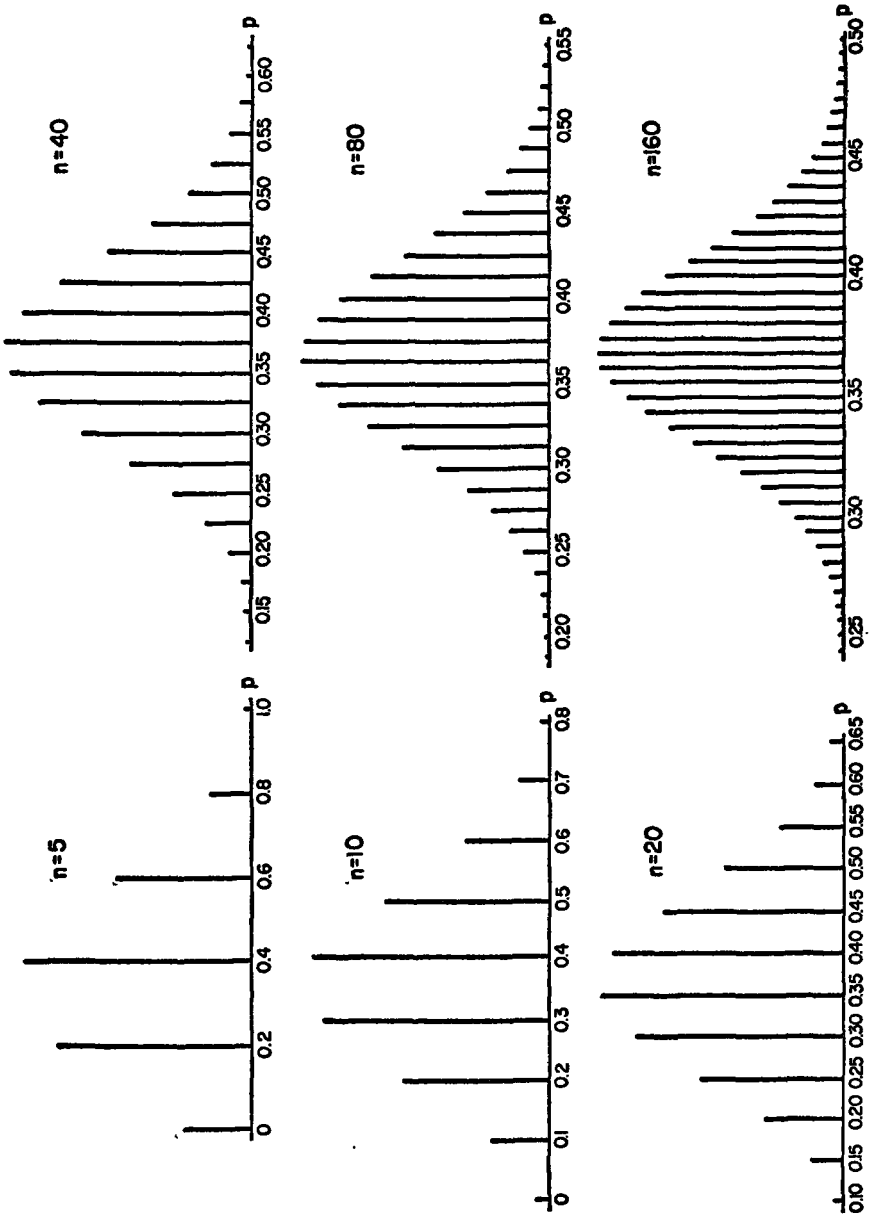
Wij hebben in het voorgaande de stelling, dat een binomiale verdeling nadert tot een normale verdeling als n nadert tot oneindig, slechts aannemelijk trachten te maken; een exact bewijs hiervan valt buiten het kader van dit boek[2]. Wij wijden nu eerst enige paragrafen aan normale verdelingen en bespreken daarna de wijze waarop — en de voorwaarden waaronder — een bepaalde binomiale kansverdeling door een normale verdeling kan worden benaderd.

5.2. Normale verdelingen en de standaardnormale verdeling

In de hoofdstukken 3 en 4 hebben wij uitsluitend te maken gehad met discreet verdeelde stochastische grootheden, die dus een bepaald aantal geïsoleerde waarden konden aannemen. De normale verdeling is

[1] Deze grafieken zijn — analoog aan de kolommendiagrammen in fig. 5.1 — zo geconstrueerd, dat na standaardisering de hoogte van de ordinaten op de klassemiddens gelijk is gesteld aan y_0 (volgens (5.2)). De plaats van de verdelingen op de X-as is zodanig, dat de gemiddelden van de onder elkaar staande verdelingen op dezelfde plaats liggen.

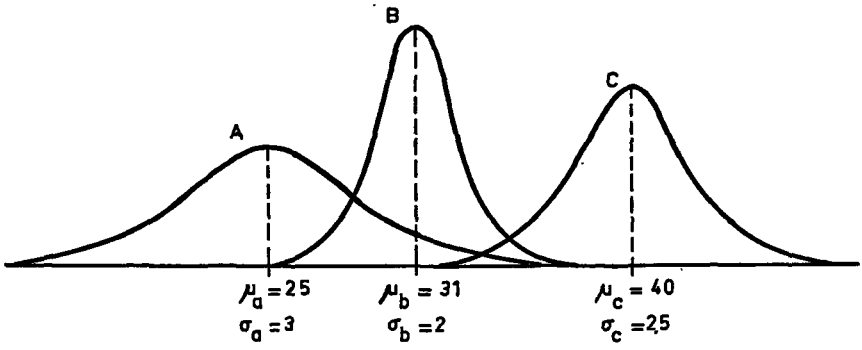
[2] Zie bijv. FREUDENTHAL (1957) of NEYMAN (1950).



Figuur 5.2. Grafieken van de binomiaal verdeelde grootheid $\underline{p} = \underline{x}/n$ voor $P = 0{,}37$ en verschillende in grootte toenemende n [uit WALLIS EN ROBERTS (1956)].

de kansverdeling van een *continu* verdeelde stochastische grootheid $\underline{x}$, die een oneindig groot aantal waarden van $-\infty$ tot $+\infty$ kan aannemen.

Een normale verdeling wordt gespecificeerd door een formule, waarin — behoudens een tweetal constanten — haar rekenkundig gemiddelde μ_x en haar spreiding σ_x voorkomen. De grafische voorstelling van een normale verdeling (zie fig. 5.4) toont de bekende klokvormige kromme, die symmetrisch is t.o.v. de ordinaat op μ_x (zodat gemiddelde, mediaan en modus samenvallen) en die aan beide zijden van μ_x steeds dichter tot de X-as nadert, zonder deze ooit te bereiken. Voor elk verschillend paar waarden van μ_x en σ_x verkrijgt men een andere normale verdeling. De *plaats* van de verdeling op de X-as wordt bepaald door de waarde van μ_x, haar *spreiding* (en vorm) door σ_x (zie figuur 5.3). Voor een normale verdeling met gemiddelde μ en spreiding σ gebruikt men de afkorting: $N(\mu, \sigma)$; verdeling A in figuur 5.3 kan dus worden aangeduid met $N(25, 3)$.



Figuur 5.3. Drie normale verdelingen met verschillende gemiddelden en verschillende spreidingen.

Een normaal verdeelde grootheid $\underline{x}$ met gemiddelde μ_x en spreiding σ_x kan door standaardisering (zie 3.3.2) worden omgezet in *de standaardnormaal* verdeelde grootheid $\underline{T}$ met gemiddelde 0 en spreiding 1

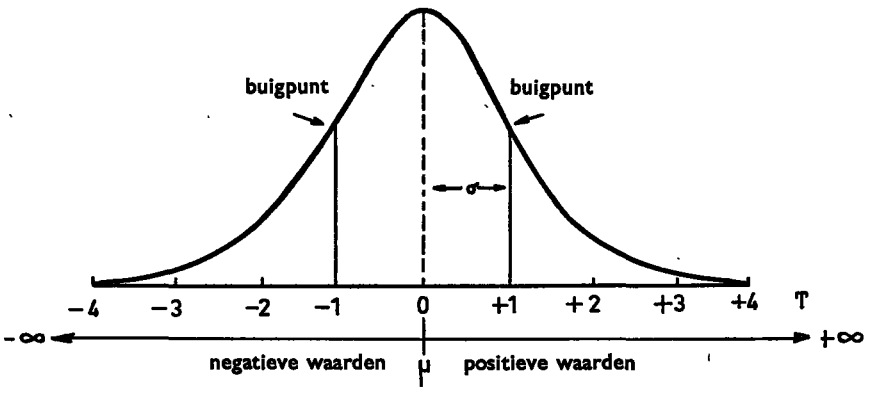
$$\underline{T} = \frac{\underline{x} - \mu_x}{\sigma_x} \tag{5.3}$$

zodat

$$\underline{x} = \mu_x + \underline{T}\sigma_x. \tag{5.4}$$

In plaats van het symbool T, afkomstig van Bok (1948), gebruikt men vaak u of z.

De standaardnormale verdeling is in figuur 5.4 getekend. Elke normale verdeling $N(\mu, \sigma)$ gaat door standaardisering volgens (5.3) in deze



Figuur 5.4. De standaardnormale verdeling: $\underline{T} = \frac{\underline{x} - \mu}{\sigma}$, $\mu_T = 0$, $\sigma_T = 1$.

standaardnormale verdeling $N(0,1)$ over. Uit (5.3) en (5.4) blijkt duidelijk, dat T_i de afstand is van x_i tot μ_x, uitgedrukt in de spreiding σ_x. Voor $x_i < \mu_x$ neemt T_i een *negatieve* waarde aan, voor $x_i > \mu_x$ een *positieve*. Beschouw bijv. de verdeling $N(25,5)$. Wordt deze in de verdeling $N(0,1)$ omgezet, dan vindt men negatieve waarden van $\underline{T}$ voor waarden van $\underline{x}$ kleiner dan $\mu_x = 25$ en positieve waarden van $\underline{T}$ voor waarden van $\underline{x}$ groter dan 25. Voor $x = 35$ volgt uit (5.3): $T = (35 - 25)/5 = +2$, d.w.z. deze waarde van $\underline{x}$ ligt tweemaal de spreiding boven haar gemiddelde. Beschouw vervolgens de verdeling $N(100, 12)$. Voor $x' = 124$ vindt men met (5.3) $T = (124 - 100)/12 = +2$. Relatief, d.i. rekening houdend met de grotere spreiding van deze tweede normale verdeling, ligt $x' = 124$ dus even ver boven haar gemiddelde als $x = 35$ in de eerstgenoemde normale verdeling.

5.3. De standaardnormale verdeling

5.3.1. CONTINUE KANSVERDELINGEN[1]

De standaardnormale verdeling is een continue kansverdeling. Continue kansverdelingen wijken in enkele belangrijke opzichten van discrete kansverdelingen af. Wij demonstreren dit met een voorbeeld.

Onderstel, dat men een cirkelvormige schaal heeft, die één eenheid lang is en dat in het middelpunt daarvan een zorgvuldig uitgebalanceerde wijzer is geplaatst. Als men deze wijzer laat draaien, heeft elk punt van de schaal dezelfde kans om door de punt van de wijzer te worden aangewezen. Hoe groot is nu de kans op een bepaalde uitkomst,

[1] Zie voor een uitvoeriger behandeling WILKS (1943), ADAMS (1955).

zeg $\frac{1}{4}$? Stel dat deze uitkomst de zeer kleine kans ξ (bijv. 0,00001) bezit. Om consequent te zijn dient men aan elke andere mogelijke uitkomst dezelfde kans toe te kennen — maar dan is de som van de kansen van alle uitkomsten groter dan 1. Om dit in te zien behoeft men slechts de (beperkte) reeks uitkomsten 1, $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{5}$, ... te beschouwen. Neemt men nu een willekeurig geheel getal $N > 1/\xi$ en neemt men de eerste N leden van de voorgaande reeks, dan is de som van hun kansen groter dan 1. Hieruit volgt dat men bij een continue kansverdeling de kans op een bepaalde uitkomst, zoals $\frac{1}{4}$, *nul* moet stellen.

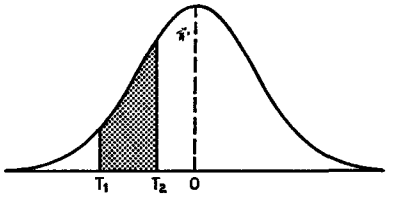
De niet-mathematisch geschoolde lezer komt dit wellicht vreemd voor, maar hij dient te bedenken, dat een continue kansverdeling een *mathematische conceptie* is, evenals een lijn zonder dikte, een punt zonder dimensie, een perfecte cirkel, e.d. In werkelijkheid kan men het wijzer-model niet verkrijgen, omdat slechts met beperkte nauwkeurigheid kan worden afgelezen zodat men dan de schaal zal onderverdelen in intervallen van gelijke lengte. De kans op elk interval zal, als alles zeer zorgvuldig geconstrueerd wordt, (ongeveer) gelijk zijn (denk bijv. aan een roulette).

In de statistiek levert het gebruik van continue kansverdelingen, zoals de standaardnormale, geen moeilijkheden op, omdat meestal naar de kans op een bepaald interval gevraagd wordt. Zo'n intervalkans kan bepaald worden door te integreren (d.i. op bepaalde wijze te sommeren), maar zelfs dit behoeft men niet zelf te doen. Van alle continue kansverdelingen, waar de statistiek gebruik van maakt, zijn reeds lang uitgebreide tabellen met intervalkansen beschikbaar.

Een praktische consequentie van het voorgaande is, dat men voor een continu verdeelde grootheid $\underline{x}$ geen onderscheid behoeft te maken tussen $P(\underline{x} < x_1)$ en $P(\underline{x} \leqslant x_1)$, of tussen $P(x_1 < \underline{x} < x_2)$ en $P(x_1 \leqslant \underline{x} \leqslant x_2)$, daar zowel $P(\underline{x} = x_1)$ als $P(\underline{x} = x_2)$ gelijk is aan nul.

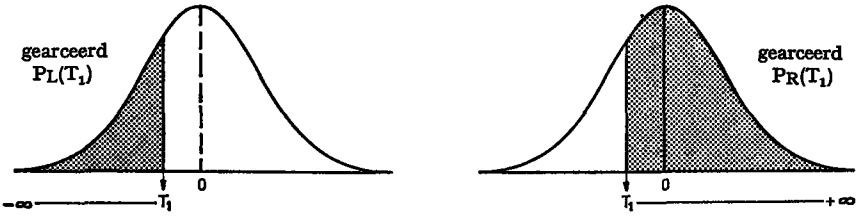
5.3.2. LINKSE, RECHTSE EN TWEEZIJDIGE OVERSCHRIJDINGSKANSEN

De standaardnormale verdeling is een kansverdeling, zodat het totale oppervlak onder de standaardnormale kromme gelijk is aan 1. Richt men in de punten T_1 en T_2 ordinaten op, dan snijden deze een deel van het oppervlak onder de kromme uit (zie figuur 5.5), dat gelijk is aan de kans dat de grootheid $\underline{T}$ een waarde tussen T_1 en T_2 aanneemt. Voor deze kans voeren wij de notatie $P(T_1, T_2)$ in.



Figuur 5.5. Standaardnormale verdeling: gearceerd is $P(T_1 < \underline{T} < T_2) = P(T_1, T_2)$.

Analoog aan 3.6 definiëren wij als *linkse (overschrijdings)kans* van de waarde T_1, $P_L(T_1)$, de kans die behoort bij het interval $(-\infty, T_1)$ en als *rechtse (overschrijdings)kans* van T_1, $P_R(T_1)$, de kans die behoort bij het interval $(T_1, +\infty)$ (zie figuur 5.6).



Figuur 5.6. Standaardnormale verdeling: linkse en rechtse kansen.

De *tweezijdige overschrijdingskans*, $P_D(T_i)$ van de waarde T_i is de kans, dat de grootheid $\underline{T}$ een waarde aanneemt die verder dan $\pm T_i$ van het centrum der verdeling, $\mu = 0$, verwijderd is. Wegens de symmetrie van de $N(0,1)$-verdeling geldt

$$\text{voor } T_i < 0\text{: } P_D(T_i) = 2P_L(T_i),$$
$$\text{voor } T_i > 0\text{: } P_D(T_i) = 2P_R(T_i).$$

Als men één der kansen $P_L(T_1)$, $P_R(T_1)$ of $P_D(T_1)$ kent, kan men de beide andere direct berekenen. Kent men de linkse kansen van de waarden T_1 en T_2, dan kan men de kans van het interval (T_1, T_2) vinden door te nemen (zie figuur 5.5)

$$P(T_1, T_2) = P_L(T_2) - P_L(T_1).$$

5.3.3. TABELLEN VAN DE STANDAARDNORMALE VERDELING

Men kan alle kansen betreffende een normale verdeling, die men in de praktijk nodig heeft, bepalen als men beschikt over een tabel die bijv. $P_L(T)$ geeft voor vele waarden van $\underline{T}$. Dergelijke tabellen zijn reeds lang geleden berekend en in allerlei graden van uitvoerigheid gepubliceerd. In tabel A (bijlage 1) is zo'n tabel met linkse kansen gegeven voor een groot aantal waarden van $\underline{T}$ [van —6,9 t/m +6,9, eerst met stappen van 0,01 (tabel A–1) en later met stappen van 0,1 (tabel A–2]. Uit tabel A kan men een groot aantal *linkse* kansen direct aflezen.

Voorbeeld (zie fig. 5.7)

5.1. LINKSE KANSEN

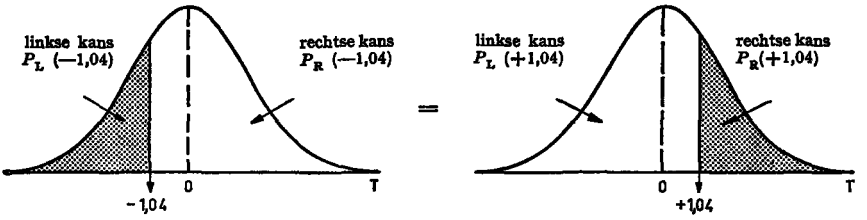
a. T is negatief: $P_L(-1{,}04) = ?$
Zoek in de eerste kolom (T) de waarde 1,0 op en ga op de met een *minteken* aangegeven regel naar rechts tot in de kolom 0,04. Daar staat 0,1492, zodat $P_L(-1{,}04) = 0{,}1492$.

b. T is positief: $P_L(1{,}04) = ?$
Als (*a*), maar op de met een *plusteken* aangeduide regel. Men vindt dan $P_L(1{,}04) = 0{,}8508$.

Tabel A kan echter ook worden gebruikt voor het bepalen van *rechtse* kansen.

Voorbeeld (zie eveneens fig. 5.7)

5.2. { RECHTSE { *c. T is negatief*: $P_R(-1{,}04) = P_L(+1{,}04)$, zie dus 1*b*.
{ KANSEN { *d. T is positief*: $P_R(+1{,}04) = P_L(-1{,}04)$, zie dus 1*a*.



Figuur 5.7. Standaardnormale verdeling: linkse en rechtse kansen.

Bij elke waarde van T kan men een kans vinden, die aangeeft welke fractie van het oppervlak onder de standaardnormale curve links van de ordinaat op deze waarde gelegen is. Men kan een *fractiel* van de standaardnormale verdeling aangeven, door deze fractie (= linkse kans) als index aan T toe te voegen (zie 2.4.3).

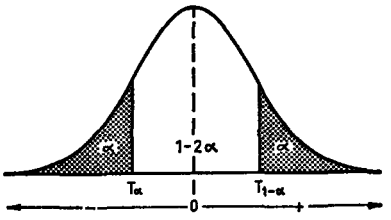
Voorbeelden

5.3. $T_{0,67}$ (= + 0,44) is het fractiel 0,67 (67e percentiel) van de $N(0,1)$ verdeling. Dus $P_L(T_{0,67}) = 0{,}67$, $P_R(T_{0,67}) = 1 - 0{,}67 = 0{,}33$.

5.4. Verifieer met tabel A dat: $T_{0,20} = -0{,}84$, $T_{0,50} = 0$, $T_{0,90} = +1{,}28$.

Een willekeurig fractiel wordt aangeduid met T_α. Uit figuur 5.8 blijkt dat:

$$T_\alpha = -T_{1-\alpha} \qquad -T_\alpha = T_{1-\alpha}$$
$$T_\alpha + T_{1-\alpha} = 0$$
$$P(T_\alpha, T_{1-\alpha}) = 1 - 2\alpha \qquad P_L(T_\alpha) = P_R(T_{1-\alpha}).$$



Figuur 5.8. Fractielen.

In tabel B (bijlage 1) kunnen van verschillende fractielen (= linkse kansen) de bijbehorende waarden van $\underline{T}$ worden afgelezen. Uit deze tabel blijkt o.m. dat:

$$T_{0,025} = -1{,}960 = -T_{0,975} \qquad T_{0,005} = -2{,}576 = -T_{0,995}$$
$$T_{0,0005} = -3{,}291 = -T_{0,9995}.$$

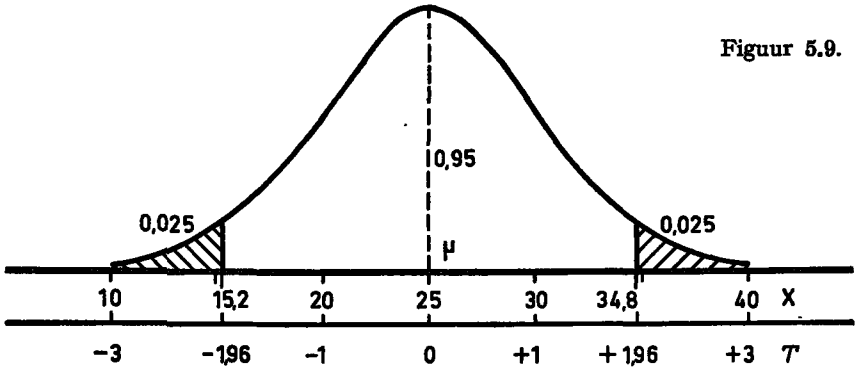
Hieruit volgt:

$$P(-1{,}960, +1{,}960) = 0{,}95$$
$$P(-2{,}576, +2{,}576) = 0{,}99$$
$$P(-3{,}291, +3{,}291) = 0{,}999.$$

En:

$$P_D(-1{,}960) = P_D(+1{,}960) = 1 - 0{,}95 = 0{,}05$$
$$P_D(-2{,}576) = P_D(+2{,}576) = 1 - 0{,}99 = 0{,}01$$
$$P_D(-3{,}291) = P_D(+3{,}291) = 1 - 0{,}999 = 0{,}001.$$

Hieruit volgt [zie ook (5.4)] dat voor een $N(\mu, \sigma)$-verdeling de kans op het interval $\mu \pm 2{,}576\sigma$ gelijk is aan 0,99 en dat de kans op het interval $\mu \pm 3{,}291\sigma$ gelijk is aan 0,999. Voor een $N(25,5)$ verdeling is de kans op het interval $25 \pm (1{,}96) \times 5$, d.i. het interval (15,2, 34,8), gelijk aan 0,95. De kans op een waarde buiten dit interval is $1 - 0{,}95 = 0{,}05$ (zie figuur 5.9).



Figuur 5.9.

Voorbeelden

5.5. Hoe groot is de kans dat de standaardnormaal verdeelde grootheid $\underline{T}$ een waarde aanneemt binnen het interval: *a.* (—2,52, +0,25), *b.* (+0,23, +1,54)? Deze vragen zijn op te lossen met tabel A:

a.	$P_L(+0{,}25)$	$= 0{,}5987$	*b.*	$P_L(+1{,}54)$	$= 0{,}9382$
	$P_L(-2{,}52)$	$= 0{,}0059$		$P_L(+0{,}23)$	$= 0{,}5910$
	$P(-2{,}52, +0{,}25)$	$= 0{,}5928$		$P(+0{,}23, +1{,}54)$	$= 0{,}3472$

5.6. Hoe groot is de kans dat $\underline{T}$ een waarde aanneemt tussen $T_{0,20}$ en $T_{0,70}$? De indices van deze fractielen geven linkse kansen aan, zodat de gevraagde kans gelijk is aan: $0{,}70 - 0{,}20 = 0{,}50$.

5.7. Verifieer met tabel B, dat $T_{0,05} = -1{,}645$, $T_{0,95} = +1{,}645$, $T_{0,10} = -1{,}282$, $T_{0,25} = -0{,}6745$, $T_{0,80} = 0{,}842$, $T_{0,9999} = 3{,}719$.

5.8. De grootheid $\underline{x}$ volgt een normale verdeling met $\mu_x = 30$ en $\sigma_x = 5$.
Bepaal de kans dat deze grootheid een waarde tussen 25 en 45 aanneemt.
Volgens (5.3) is:

$$\underline{T} = \frac{\underline{x} - 30}{5}, \text{ zodat } \begin{cases} T_2 = \dfrac{45 - 30}{5} = +3,\ P_L(+3) = 0{,}9987 \text{ (tabel A)} \\ T_1 = \dfrac{25 - 30}{5} = -1,\ P_L(-1) = 0{,}1587 \text{ (tabel A)} \end{cases}$$

$$\overline{P(25 < \underline{x} < 45) = 0{,}8400.}$$

5.3.4. OPGAVEN

(Teken bij deze opgaven steeds de (standaard) normale verdeling en arceer het oppervlak onder de kromme, waarop de opgave betrekking heeft.)

5.1. Bepaal voor de standaardnormaal verdeelde grootheid $\underline{T}$ de kansen:

a. $P_R(+2{,}05)$ *c.* $P_D(+0{,}44)$ *e.* $P_D(+3{,}00)$ *g.* $P(-1{,}54, +0{,}46)$
b. $P_L(+0{,}44)$ *d.* $P_D(-0{,}44)$ *f.* $P(0, +1{,}32)$ *h.* $P(-1{,}82, -0{,}82)$.

5.2. Als (5.1), van:

a. $P_L(+0{,}20)$ *c.* $P_D(-2{,}00)$ *e.* $P(-0{,}4, +0{,}4)$ *g.* $P(-\infty, 0)$
b. $P_R(+2{,}00)$ *d.* $P_D(+4{,}00)$ *f.* $P(+1{,}20, +1{,}95)$ *h.* $P(+0{,}5, +\infty)$.

5.3. Geef de waarde van:

a. $T_{0,98}$ *b.* $T_{0,998}$ *c.* $T_{0,28}$ *d.* $T_{0,001}$ *e.* $T_{0,025}$ *f.* $T_{0,95}$.

5.4. Geef de waarde van de volgende kansen, zonder een tabel te gebruiken:

a. $P_L(T_{0,12})$ *b.* $P_R(T_{0,87})$ *c.* $P_R(T_{0,01})$
d. $P(T_{0,12} < \underline{T} < T_{0,72})$ *e.* $P_D(T_{0,95})$ *f.* $P_L(T_{0,99})$
g. $P_R(T_{0,25})$ *h.* $P(T_{0,26} > \underline{T} > T_{0,16})$.

5.5. Bepaal de waarde van $\underline{T}$ voor elk der volgende vergelijkingen:

a. $P_R(T) = 0{,}16$ *b.* $P_L(T) = 0{,}04$ *c.* $P(-T, T) = 0{,}28$
d. $P_D(T) = 0{,}50$.

5.6. Een grootheid $\underline{x}$ volgt een normale verdeling, $\mu_x = 15$, $\sigma_x = 4$. Bepaal:

a. $P(\underline{x} < 21)$ *b.* $P(10 < \underline{x} < 20)$ *c.* $P(\underline{x} > 23)$.

5.7. Ontleend aan het examen 'Statistisch Analyst', 1953—54:

Wat is de kans, dat althans de grootste van drie waarden, gevonden bij drie onafhankelijke trekkingen uit een normaal verdeelde populatie (met gemiddelde 0 en standaarddeviatie 1), groter is dan twee?

5.8. Van een normaal verdeelde grootheid $\underline{x}$ is gegeven:

a. $\sigma_x = 40$, $P_D(\underline{x} = 110) = 0{,}8026$. Bepaal μ_x.
b. $\mu_x = 100$, $P_D(\underline{x} = 153) = 0{,}0340$. Bepaal σ_x en $P_L(\underline{x} = 60)$.

5.9. Een populatie van grote omvang bestaat uit objecten, waarvan de lengten bij benadering normaal verdeeld zijn. Als deze objecten in drie categorieën worden verdeeld, t.w.: I — korter dan 75 cm, II — 75 t/m 81,2 cm, en III — langer dan 81,2 cm, bevatten deze categorieën de volgende percentages der objecten: I — 58%, II — 38%, III — 4%. Bepaal het gemiddelde en de spreiding van de lengteverdeling.

5.4. Benadering van een binomiale verdeling door de best-passende normale verdeling

5.4.1. BINOMIALE VERDELINGEN MET $P = Q = \frac{1}{2}$

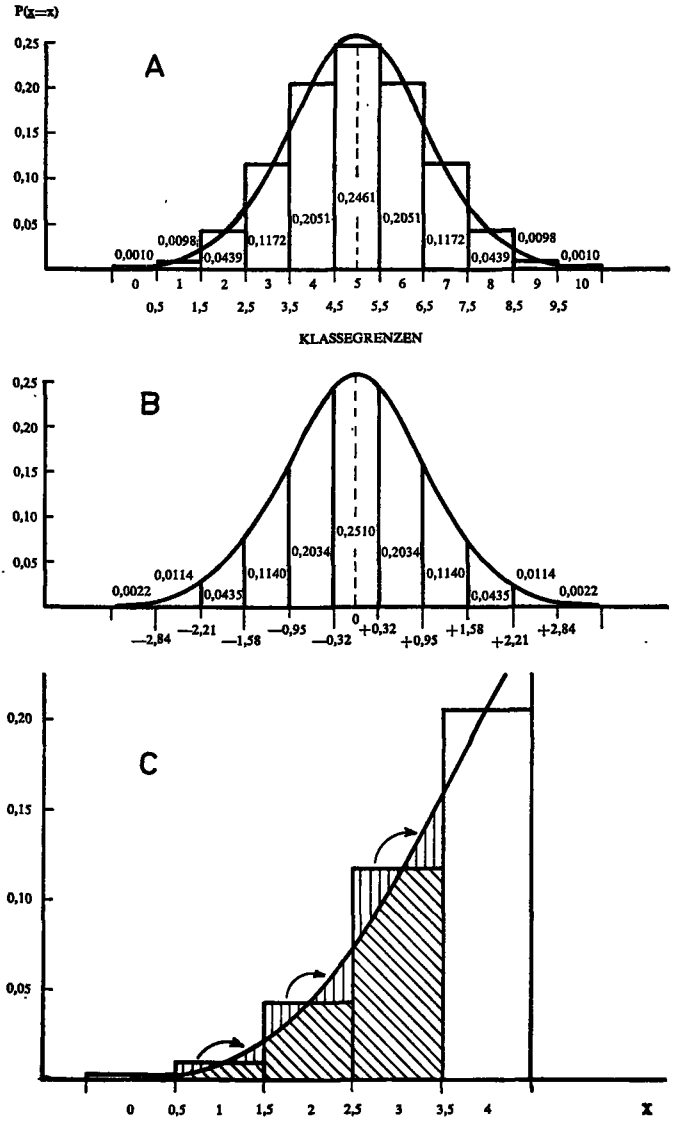
De symmetrische binomiale verdeling wordt in dit geval reeds bij kleine n goed benaderd door de best-passende normale verdeling, d.i. de normale verdeling met $\mu = Pn = \frac{1}{2}n$ en $\sigma = \sqrt{PQn} = \frac{1}{2}\sqrt{n}$. De grootheid

$$\underline{T} = \frac{\underline{x} - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} \tag{5.5}$$

is dan bij benadering standaardnormaal verdeeld, zodat voor het berekenen van binomiale kansen tabel A kan worden gebruikt.

Tabel 5.1. De binomiale verdeling $P(\underline{x} = x) = C_x^{10} \cdot 2^{-10}$ en de best-passende normale verdeling $N(5,\ 1{,}58)$

(1)	(2)	(3)	(4)	(5)	(6)
x	Bovengrens interval $g = x + \frac{1}{2}$	$T = \frac{g - 5}{1{,}58}$	$P_L(T) = P_L(x)$	$P(\underline{x} = x)$ Benaderd	$P(\underline{x} = x)$ Exact
	$+\infty$	——	1,0000		
10				0,0022	0,0010
	9,5	2,84	0,9978		
9				0,0114	0,0098
	8,5	2,21	0,9864		
8				0,0435	0,0439
	7,5	1,58	0,9429		
7				0,1140	0,1172
	6,5	0,95	0,8289		
6				0,2034	0,2051
	5,5	0,32	0,6255		
5				0,2510	0,2461
	4,5	—0,32	0,3745		
4				0,2034	0,2051
	3,5	—0,95	0,1711		
3				0,1140	0,1172
	2,5	—1,58	0,0571		
2				0,0435	0,0439
	1,5	—2,21	0,0136		
1				0,0114	0,0098
	0,5	—2,84	0,0022		
0				0,0022	0,0010
		——	0,0000		
				1,0000	1,0001



Figuur 5.10. Binomiale verdeling en best-passende normale verdeling.

A – Kolommendiagram van de binomiale verdeling met $P = \frac{1}{2}$ en $n = 10$.

B – Best-passende normale verdeling: $\mu = 5$, $\sigma = 1{,}58$.

C – Vergroting van het linkerdeel van figuur A. Deze figuur demonstreert dat de binomiale kans $P(x < 3)$ het best wordt benaderd door het oppervlak onder de best-passende normale kromme te nemen tot aan de bovengrens ($x = 3\frac{1}{2}$) van de klasse met $x = 3$ als midden.

Voorbeeld

5.9. In tabel 5.1 is de berekening opgenomen van de best-passende normale verdeling bij de binomiale verdeling:

$P(\underline{x} = x) = C_x^{10} \cdot 2^{-10}$, waarvoor: $\mu = 5$ en $\sigma = \frac{1}{2}\sqrt{10} = 1{,}58$.

Deze berekening verloopt als volgt. De kans $P(\underline{x} = x_i)$ wordt het nauwkeurigst benaderd door het oppervlak onder de best-passende normale verdeling tussen de ordinaten op $(x_i - \frac{1}{2})$ en $(x_i + \frac{1}{2})$, behalve voor de beide uiterste waarden van $\underline{x}$, 0 en 10. Hierbij wordt het oppervlak van $-\infty$ tot $\frac{1}{2}$, resp. van $n - \frac{1}{2} = 9\frac{1}{2}$ tot $+\infty$ genomen, omdat anders de som van de benaderende kansen niet gelijk aan 1 wordt (zie fig. 5.10).

De intervalgrenzen (kolom 2) worden omgezet in waarden van $\underline{T}$ (kolom 3) en hiervan worden de linkse kansen met tabel A bepaald (kolom 4). De kansen van de intervallen (dat zijn de benaderende kansen van de waarden van $\underline{x}$) kunnen door aftrekking worden verkregen (kolom 5). Als wij deze benaderende kansen met de exacte kansen (kolom 6) vergelijken, blijkt zelfs bij deze kleine n reeds een behoorlijke overeenstemming te bestaan (slechts in de staarten van de verdeling komen relatief grote verschillen voor).

Als men van een waarde x_1 van een binomiale verdeling met $P = \frac{1}{2}$ uitsluitend de *linkse kans* $P_L(x_1) = P(\underline{x} \leqslant x_1)$ behoeft te weten, kan deze bij benadering worden bepaald door

$$T_1 = \frac{x_1 + \frac{1}{2} - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} \tag{5.6}$$

te berekenen en de linkse kans van T_1 in tabel A op te zoeken.

De *rechtse kans* van een waarde x_2, $P_R(x_2) = P(\underline{x} \geqslant x_2)$, kan bij benadering worden bepaald door de rechtse kans van

$$T_2 = \frac{x_2 - \frac{1}{2} - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} \tag{5.7}$$

uit tabel A af te lezen.

De 'correctie' van $+\frac{1}{2}$ of $-\frac{1}{2}$, die wordt aangebracht om de discrete binomiale verdeling nauwkeuriger te benaderen met de continue normale verdeling (zie nogmaals figuur 5.10C), noemt men de *continuïteitscorrectie*.

Voorbeeld

5.10. Men verricht 30 worpen met een zuiver geldstuk. Hoe groot is de kans dat ten hoogste 8 maal de uitkomst 'kruis' verschijnt?

De binomiale verdeling, die benaderd wordt luidt: $P(\underline{x} = x) = C_x^{30} \cdot 2^{-30}$, zodat $\mu_x = \frac{1}{2}n = 15$ en $\sigma_x = \frac{1}{2}\sqrt{n} = 2{,}734$. Gevraagd wordt naar de kans $P(\underline{x} < 8) \approx P_L(8)$, zodat volgens (5.6):

$$T = \frac{8 + \frac{1}{2} - 15}{2{,}734} = -2{,}38.$$

In tabel A vinden wij: $P_L(-2{,}38) = 0{,}0087$. De gevraagde kans is dus bij benadering hieraan gelijk (deze benadering is zeer goed, daar de exacte kans 0,0081 blijkt te bedragen). Zonder continuïteitscorrectie vindt men:

$$T = (8 - 15)/2{,}734 = -2{,}56,$$

met een overschrijdingskans die veel te klein is, nl. 0,0052.

5.4.2. BINOMIALE VERDELINGEN MET $P \neq \frac{1}{2}$

Zoals wij in 5.1 gezien hebben, nadert ook in dit geval een binomiale verdeling bij toenemende n tot een normale verdeling en wel des te sneller, naarmate P dichter bij $\frac{1}{2}$ ligt. Voor enkele waarden van P vermelden wij onderstaand de kleinste waarden van n, waarbij de normale benadering zonder bezwaar kan worden toegepast (uit COCHRAN (1953); naar onze mening zijn deze n's wat aan de hoge kant):

P	0,5	0,4	0,3	0,2	0,1	0,05
n	30	50	80	200	600	1400

Linkse, resp. rechtse kansen kan men bij benadering bepalen via de standaardnormaal verdeelde grootheid

$$\underline{T} = \frac{\underline{x} - Pn}{\sqrt{PQn}} \tag{5.8}$$

waarbij de continuïteitscorrectie als in (5.6), resp. (5.7), kan worden aangebracht.

Voorbeeld

5.11. Van een binomiale verdeling is gegeven: $P = 0{,}4$, $n = 30$ (n is dus kleiner dan de 'norm' voor het toepassen van de normale benadering).

Men wil bepalen $P_R(16) = P(\underline{x} > 16)$. Uit (5.8) volgt:

$$T = \frac{16 - \frac{1}{2} - (0{,}4)\,(30)}{\sqrt{(0{,}4)\,(0{,}6)\,(30)}} = 1{,}305.$$

Uit tabel A blijkt, dat $P_R(1{,}305)$ ongeveer gelijk is aan 0,0960 [$P_R(1{,}30) = 0{,}0968$, $P_R(1{,}31) = 0{,}0951$, tussen deze waarden kan men zonder bezwaar lineair interpoleren]. Deze kans blijkt goed overeen te stemmen met de exact berekende kans, die 0,0971 bedraagt.

5.4.3. OPGAVEN

5.10. Bepaal bij de binomiale verdeling met $P = 0{,}4$ en $n = 20$ de bestpassende normale verdeling. Vergelijk deze met de exacte verdeling in tabel 4.3.

5.11. Men werpt 500 maal met een zuiver geldstuk. Hoe groot is de kans dat: *a.* Ten hoogste 220 keer 'munt' wordt geworpen? *b.* Ten hoogste 220 of ten minste 280 keer 'munt' wordt waargenomen?

5.12. Men trekt een aselecte steekproef van 100 elementen uit een populatie van grote omvang, waarvan 30% der elementen kenmerk A draagt. Hoe groot is bij benadering de kans, dat deze steekproef:

a. Minder dan 21 elementen met kenmerk A bevat? *b.* Meer dan 42 elementen met kenmerk A bevat? *c.* Ten minste 25 en ten hoogste 35 elementen met kenmerk A bevat?

5.5.* Aanpassing van een normale verdeling

5.5.1. BEREKENING VAN DE BEST-PASSENDE NORMALE VERDELING

Als men beschikt over een waargenomen frequentieverdeling van een grootheid $\underset{\sim}{x}$ en vermoedt dat deze (ongeveer) normaal verdeeld is, kan men de best-passende normale verdeling (met $\mu = \bar{x}$ en $\sigma = s_w$) berekenen en deze met de waargenomen verdeling vergelijken.

Voorbeeld

5.12. In tabel 5.2 is hier een voorbeeld van gegeven, dat voor zichzelf spreekt (de verdeling is ontleend aan voorbeeld 1.25; nu wordt echter ondersteld, dat de 190 uitkomsten een steekproef betreffen). De overeenstemming tussen de waargenomen en best-passende normale verdeling kan worden onderzocht met de χ^2-toets voor aanpassing in 10.1.

Tabel 5.2. Waargenomen verdeling en best-passende normale verdeling
Gesteld is: $\mu = \bar{x} = 13{,}40$, $\sigma = s = 0{,}95$

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Klassegrenzen	f	Bovengrens g	$T = \frac{g-13{,}4}{0{,}95}$	$P_L(T)$	$P(\underset{\sim}{x} = x)$	$190\,P$
15,75—16,25	2	$+\infty$	—	1,0000	0,0068	1,3
15,25—15,75	3	15,75	2,47	0,9932	0,0188	3,6
14,75—15,25	11	15,25	1,95	0,9744	0,0522	9,9
14,25—14,75	19	14,75	1,42	0,9222	0,1089	20,7
13,75—14,25	26	14,25	0,89	0,8133	0,1690	32,1
13,25—13,75	46	13,75	0,37	0,6443	0,2079	39,5
12,75—13,25	38	13,25	—0,16	0,4364	0,1881	35,7
12,25—12,75	24	12,75	—0,68	0,2483	0,1352	25,7
11,75—12,25	15	12,25	—1,21	0,1131	0,0722	13,7
11,25—11,75	4	11,75	—1,74	0,0409	0,0290	5,5
10,75—11,25	1	11,25	—2,26	0,0119	0,0093	1,8
10,25—10,75	1	10,75	—2,79	0,0026	0,0026	0,5
	190					190,0

5.5.2. NORMALE WAARSCHIJNLIJKHEIDSPAPIER

Dit speciale grafiekenpapier kan worden gebruikt als men snel wil nagaan, of een waargenomen verdeling ongeveer normaal verdeeld is. Blijkt dit het geval te zijn dan kan men — als het aantal waarnemingen niet te klein is — met behulp van dit papier ook het gemiddelde en de spreiding van de waargenomen verdeling schatten. Voor een uitvoerige beschrijving en gebruiksaanwijzing raadplege men VEEN (1956) of EXALTO (1950).

5.5.3. OPGAVEN

5.13. Bepaal bij onderstaande waargenomen frequentieverdeling de bestpassende normale verdeling. Gegeven: $\bar{x} = 44{,}11$, $s = 2{,}64$.

x	51	50	49	48	47	46	45	44	43	42	41	40	39	38	37	36	Totaal
f	1	8	4	19	33	31	49	47	51	28	26	18	7	4	0	1	327

5.14. Als opgave 5.13, voor de verdeling van de geboortegewichten in tabel 2.9; gegeven is $\bar{x} = 34{,}5$ en $s = 6{,}32$.

HOOFDSTUK 6

TOETSINGSTHEORIE

De statistische toetsingstheorie, die werd ontwikkeld door NEYMAN en PEARSON (1933), houdt zich bezig met het toetsen van onderstellingen (die *hypothesen* worden genoemd) omtrent niet geheel bekende populatieverdelingen, op basis van aselecte steekproeven. In 6.1 zullen wij de principes van deze theorie aan de hand van een eenvoudig beslissingsprobleem betreffende een dichotoom verdeelde populatie uiteenzetten. In de paragrafen 6.2 en 6.3 bespreken wij de praktische uitvoering van de zg. binomiale toetsen betreffende een dichotomie. Tenslotte volgt in 6.4 een korte beschouwing over de door WALD en zijn medewerkers gedurende de tweede wereldoorlog ontwikkelde techniek van sequente toetsing.

6.1. Grondslagen van de toetsingstheorie

6.1.1. PROBLEEMSTELLING [1]

Een arts A wordt door de Nederlandse regering uitgezonden naar Nieuw-Guinea, teneinde het voorkomen van framboesia in een bepaalde streek aldaar te bestuderen. In de ontvangen rapporten van gouvernementsambtenaren wordt vermeld, dat in deze vrij afgelegen streek ongeveer 50% van de bevolking aan manifeste huidafwijkingen tengevolge van framboesia lijdt. Een gouvernementsarts die slechts een klein deel van de streek heeft bezocht, spreekt van ten hoogste 20%.

Indien het percentage framboesialijders inderdaad ongeveer 50% is, wil de regering op korte termijn aan een uitgebreid saneringsprogramma voor deze streek preferentie geven boven andere projecten van openbare gezondheidszorg. Bedraagt het 20% of minder, dan komt zo'n programma voorlopig niet voor uitvoering in aanmerking.

Bij aankomst in Nieuw-Guinea besluit A een aselecte steekproef uit de bevolking van de streek te trekken, de daarin opgenomen personen te onderzoeken en zijn conclusie te baseren op het aantal framboesialijders, dat zich daarbij bevindt. Onderstel dat deze steekproef 20

[1] Het werkelijke probleem is, om het als voorbeeld te kunnen gebruiken, uiteraard sterk vereenvoudigd. In de eerste plaats kan men zich afvragen, hòe de aselecte steekproef tot stand kan komen. Verder zal men in werkelijkheid een veel grotere steekproef gebruiken: de kleine steekproefomvang is gekozen, om het probleem nog eenvoudig exact te kunnen oplossen.

personen telt en dat zich hieronder 8 lijders bevinden. Tot welke conclusie komt hij?

6.1.2. TERMEN EN BEGRIPPEN

Omtrent het voorkomen van framboesia bij de bevolking van de betrokken streek zijn twee hypothesen naar voren gebracht:

$H_0 : P = 0{,}5$ (50% van de bevolking lijdt aan framboesia),
$H_1 : P = 0{,}2$ (20% van de bevolking lijdt aan framboesia).

De indices 0 en 1 zijn aan deze hypothesen (symbool H) *niet* willekeurig toegekend. H_0 noemt men de *te toetsen hypothese* of *nulhypothese*, H_1 de *alternatieve hypothese* of *tegenhypothese*. Om redenen die wij later bespreken kiest men — zo mogelijk — als H_0 die hypothese, waarvan het ten onrechte verwerpen de ernstigste consequenties heeft. Hypothesen van dit type, waarin één waarde van P wordt gestipuleerd noemt men *enkelvoudige hypothesen*. Hypothesen van de vorm $P \leqslant 0{,}2$, $P \geqslant 0{,}5$, e.d., die samengesteld zijn uit een groot aantal waarden van P noemt men *samengestelde hypothesen*.

Een statistisch onderzoek naar de juistheid van een hypothese wordt een *toetsing* genoemd. Een methode waarmee een toetsing wordt uitgevoerd noemt men een *toets*. Elke toetsing wordt gebaseerd op een aselecte steekproef. De steekproeffunctie waarvan gebruik wordt gemaakt bij het uitvoeren van een toetsing noemt men de *toetsingsgrootheid*.

6.1.3. DE VERDELINGEN VAN DE TOETSINGSGROOTHEID ONDER H_0 EN H_1

Bij het framboesia-probleem kiezen wij als toetsingsgrootheid vanzelfsprekend het aantal lijders in de steekproef, x. Bij $n = 20$ kan de grootheid $\underset{\sim}{x}$ de waarden 0(1) 20 aannemen.

Als H_0 juist is (men zegt hiervoor kortweg: *onder* H_0) volgt de grootheid $\underset{\sim}{x}$ de binomiale verdeling (zie 4.3.1)

$$P(\underset{\sim}{x} = x \mid H_0) = \frac{20!}{x!(20-x)!} \cdot 2^{-20}. \tag{6.1}$$

Onder H_1 volgt de grootheid $\underset{\sim}{x}$ de binomiale verdeling

$$P(\underset{\sim}{x} = x \mid H_1) = \frac{20!}{x!(20-x)!} (0{,}2)^x (0{,}8)^{20-x}. \tag{6.2}$$

Deze twee binomiale verdelingen zijn in tabel 6.1 opgenomen.

Tabel 6.1. Kansverdelingen van de grootheid $\underline{x}$, het aantal framboesialijders in een aselecte steekproef van 20 personen*

	$H_0 : P = 0{,}5$	$H_1 : P = 0{,}2$
x	$P(\underline{x} = x \mid H_0)$	$P(\underline{x} = x \mid H_1)$
20	0,0000	0,0000
19	0,0000	0,0000
18	0,0002	0,0000
17	0,0011	0,0000
16	0,0046	0,0000
15	0,0148	0,0000
14	0,0370	0,0000
13	0,0739	0,0000
12	0,1201	0,0001
11	0,1602	0,0005
10	0,1762	0,0020
9	0,1602	0,0074
8	0,1201	0,0222
7	0,0739	0,0545
6	0,0370	0,1091
5	0,0148	0,1746
4	0,0046	0,2182
3	0,0011	0,2054
2	0,0002	0,1369
1	0,0000	0,0576
0	0,0000	0,0115

* Kansen in vier decimalen nauwkeurig.

6.1.4. HET OPSTELLEN VAN EEN BESLISSINGSSCHEMA

Kritieke zone. Op grond van de kansverdelingen in tabel 6.1 kan men nu overgaan tot het opstellen van een beslissingsschema, waarin voor iedere waarde van $\underline{x}$ wordt aangegeven, welke beslissing wordt genomen. Daartoe dient men te kiezen:

1. Een verzameling van waarden van $\underline{x}$, waarvoor H_0 wordt verworpen ten gunste van H_1. Men noemt deze verzameling de *kritieke zone* van de toets en gebruikt hiervoor het symbool Z.
2. De complementaire verzameling van waarden, waarvoor H_0 niet wordt verworpen.

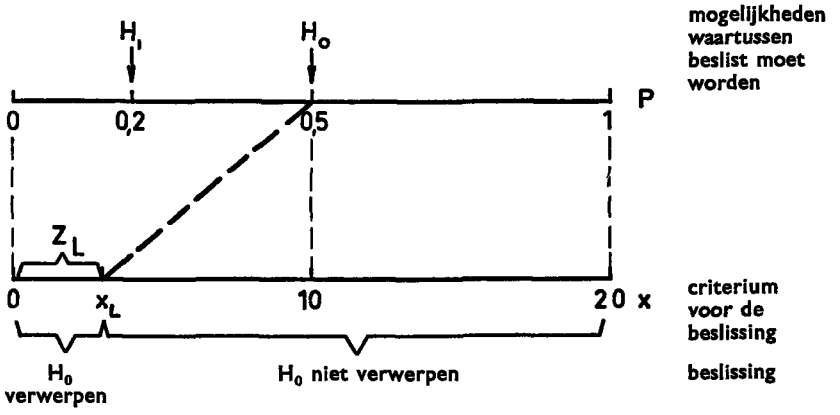
Iedere waarde van $\underline{x}$ kan zowel onder H_0 als onder H_1 voorkomen, maar in het algemeen zullen de kansen van een bepaalde waarde onder H_0 en onder H_1 verschillen. Zo zien wij bijv. in tabel 6.1 dat:

$$P(\underline{x} = 5 \mid H_0) = 0{,}0148 \text{ en } P(5 \mid H_1) = 0{,}1746,$$
$$P(\underline{x} = 11 \mid H_0) = 0{,}1602 \text{ en } P(11 \mid H_1) = 0{,}0005.$$

Dit feit wijst de richting die men bij het opstellen van een beslissings-

schema in principe kan inslaan. Waarden van $\underline{x}$ met een kleine kans onder H_0 en een grote kans onder H_1 plaatst men in de kritieke zone (zij leiden tot het verwerpen van H_0). Waarden van $\underline{x}$ met een grote kans onder H_0 en een kleine kans onder H_1 vallen buiten deze zone en leiden tot het niet-verwerpen van H_0).

Plaats van de kritieke zone. Bij het framboesia-probleem is $P_1 < P_0$ (onder H_1 is de fractie lijders in de populatie 0,2, onder H_0 0,5). Men zal daarom alleen geneigd zijn tot het verwerpen van H_0 ten gunste van H_1, indien het in de steekproef waargenomen aantal lijders *laag* is, dus voor $x = 0, 1, 2, \ldots$ tot een zekere waarde. De kritieke zone wordt in zo'n geval dus uitsluitend uit lage (d.i. grafisch gezien: links gelegen) waarden van $\underline{x}$ opgebouwd. Men spreekt dan van een éénzijdige *linker kritieke zone* en van *links éénzijdige* toetsing. De bovengrens van deze linkerzone, x_L, noemt men de *kritieke* waarde van deze zone (zie figuur 6.1).



Figuur 6.1. Beslissingsschema bij links éénzijdige toetsing: $H_0 : P = 0{,}5$, $H_1 : P = 0{,}2$, $n = 20$ (naar HEMELRIJK & VAN ELTEREN (1953)).

Indien $P_1 > P_0$ is (bijv. als de te toetsen hypothese luidt $H_0 : P = 0{,}4$ met als alternatief $H_1 : P = 0{,}8$) zal men H_0 uitsluitend willen verwerpen voor hoge waarden van $\underline{x}$, dus voor $x = n, n - 1, \ldots$ tot een zekere waarde x_R. Nu wordt een *rechter kritieke zone* gekozen en *rechts éénzijdig* getoetst.

Als de te toetsen hypothese luidt $H_0 : P = P_0$ met als alternatief $H_1 : P \neq P_0$ (bijv. $H_0 : P = 0{,}5$, $H_1 : P \neq 0{,}5$) zal men tot verwerpen van H_0 overgaan voor lage èn voor hoge waarden van $\underline{x}$. De kritieke zone bestaat dan uit een linker èn een rechter deel: $Z = Z_L + Z_R$ en men voert een *tweezijdige* toetsing uit.

Fouten bij de toetsing van hypothesen. Bij het toetsen van de hypothese H_0 met als alternatief de hypothese H_1 kunnen de volgende situaties optreden die in tabel 6.2 zijn weergegeven.

Tabel 6.2. Vier situaties bij het nemen van een beslissing

De beslissing luidt:	*Juist* is de hypothese H_0	*Juist* is de hypothese H_1
H_0 niet verwerpen	juiste beslissing	Fout II
H_0 verwerpen	Fout I	juiste beslissing

In twee gevallen komt men tot een juiste uitspraak, in de twee andere tot een foutieve. De consequenties van de twee soorten fouten zijn echter in het algemeen niet gelijkwaardig.

Voorbeelden

6.1. Iemand staat terecht wegens moord. De rechtbank doet uitspraak; hierbij kunnen de volgende situaties optreden:

De uitspraak van de rechtbank is:	De beklaagde is Onschuldig	De beklaagde is Schuldig
Onschuldig	Juist	Fout II
Schuldig	Fout I	Juist

Welke van deze fouten acht ons rechtssysteem het ernstigst en waarom?

6.2. Beschouw het framboesia-probleem. Hierbij hebben wij als H_0 gekozen: $P = 0{,}5$. Het maken van fout I houdt hier in, dat ten onrechte wordt geconcludeerd dat (ongeveer) 20% van de bevolking aan framboesia lijdt. Als A dit rapporteert, zal het saneringsprogramma althans voorlopig achterwege blijven. De werkelijk ernstige framboesia in de streek wordt dan niet bestreden en de betrokken bevolking ondervindt hier de nadelige gevolgen van.

Bij het maken van fout II wordt ten onrechte geconcludeerd dat (ongeveer) 50% van de bevolking aan framboesia lijdt. Het rapport van A leidt nu tot een te uitgebreide hulpactie, die weliswaar voor de betrokken bevolking gunstig is, maar waardoor wellicht andere projecten onuitgevoerd blijven. Dit voorbeeld toont wel aan dat het in de praktijk vaak niet eenvoudig is om uit te maken aan welke fout de ernstigste consequenties zijn verbonden.

Fout I noemt men een *fout van de eerste soort.* Deze wordt gemaakt als de te toetsen hypothese H_0 ten onrechte wordt verworpen. De grootte van de kans op deze fout, behorend bij een bepaalde kritieke zone, wordt aangeduid met α_Z.

Fout II noemt men een *fout van de tweede soort*. Deze fout wordt gemaakt als de hypothese H_0 niet wordt verworpen, terwijl H_1 juist is. De grootte van de kans op deze fout, behorend bij een bepaalde kritieke zone, wordt aangeduid met β_Z. De kans dat H_0 wordt verworpen terwijl H_1 juist is bedraagt $1 - \beta_Z$ en wordt het *onderscheidingsvermogen* van de toets genoemd.

Deze drie begrippen worden in het volgende punt aan de hand van het framboesia-voorbeeld toegelicht.

Omvang van de kritieke zone. Wij hebben reeds gezien dat de *plaats* van de kritieke zone wordt bepaald door de tegenhypothese. Om nog even te recapituleren: toetst men de hypothese $H_0 : P = P_0$, dan kiest men:

bij H_1 $\begin{cases} P < P_0 \text{ — een linker kritieke zone } (x = 0, 1, 2, \ldots, x_L), \\ P > P_0 \text{ — een rechter kritieke zone } (x = x_R, x_R + 1, \ldots n), \\ P \neq P_0 \text{ — een tweezijdige kritieke zone, bestaande uit een linker- en een rechterdeel.} \end{cases}$

Bij het framboesia-probleem dient een linker kritieke zone te worden gekozen. Om te kunnen uitmaken, welke waarden van $\underset{\sim}{x}$ in deze zone dienen te worden geplaatst, hebben wij in tabel 6.4 de waarden van α_Z, β_Z en $1 - \beta_Z$ opgenomen, die bij verschillende linker kritieke zones behoren. Tabel 6.3 laat voor de kritieke zone $x = 0(1)6$ zien hoe wij aan deze waarden gekomen zijn.

Uit tabel 6.3 blijkt dat:

α_Z = de kans op een fout van de eerste soort = de kans, dat de grootheid $\underset{\sim}{x}$ een tot Z_L behorende waarde aanneemt indien H_0 juist is.

β_Z = de kans op een fout van de tweede soort = de kans, dat de grootheid $\underset{\sim}{x}$ een niet tot Z_L behorende waarde aanneemt indien H_1 juist is.

$1 - \beta_Z$ = het onderscheidingsvermogen = de kans, dat de grootheid $\underset{\sim}{x}$ een tot Z_L behorende waarde aanneemt indien H_1 juist is.

Tabel 6.4 toont aan, dat bij een gegeven steekproefomvang (in dit geval $n = 20$) β_Z toeneemt naarmate α_Z kleiner wordt, en omgekeerd. Men kan dus de kritieke zone wel zodanig kiezen, dat α_Z zeer klein is, maar dan doet zich het onbevredigende feit voor dat β_Z groot is, zodat de ontworpen toets een te gering onderscheidingsvermogen bezit. D.w.z. men heeft slechts een kleine kans om de alternatieve hypothese H_1, indien deze juist is, als zodanig te erkennen. Bij de keuze van de omvang van de kritieke zone zal men dus t.a.v. de waarden van

Tabel 6.3. Bepaling van α_Z en β_Z voor de kritieke zone $x = 0$ (1) 6.

Kansverdelingen uit tabel 6.1

$H_0 : P = 0{,}5,\ H_1 : P = 0{,}2,\ n = 20$

	x	$P(\underline{x} = x \mid H_0)$		$P(\underline{x} = x \mid H_1)$	
	20	0,0000		0,0000	
	19	0,0000		0,0000	
	18	0,0002		0,0000	
	17	0,0011		0,0000	
	16	0,0046		0,0000	
	15	0,0148		0,0000	
	14	0,0370		0,0000	$\beta_Z = 0{,}0867$ Fout II
	13	0,0739		0,0000	
	12	0,1201		0,0001	
	11	0,1602		0,0005	
	10	0,1762		0,0020	
	9	0,1602		0,0074	
	8	0,1201		0,0222	
	7	0,0739	0,13	0,0545	0,96
	$x_L = 6$	0,0370		0,1091	
	5	0,0148		0,1746	
Linker kritieke zone	4	0,0046	$\alpha_Z = 0{,}0577$ Fout I	0,2182	$1 - \beta_Z = 0{,}9133$ Onderscheidings-vermogen
	3	0,0011		0,2054	
	2	0,0002		0,1369	
	1	0,0000		0,0576	
	0	0,0000		0,0115	
	Totaal	1,0000		1,0000	

Tabel 6.4. Waarden van α_Z, β_Z en $1 - \beta_Z$ voor verschillende linker kritieke zones

$H_0 : P = 0{,}5 \quad H_1 : P = 0{,}2 \quad n = 20$

Z_L bestaat uit	Fout I $\alpha_Z = P(x \text{ in } Z_L \mid H_0)$	Fout II $\beta_Z = P(x \text{ buiten } Z_L \mid H_1)$	Ondersch. vermogen $1 - \beta_Z$
$x = 0$	0,0000	0,9885	0,0115
$= 0$ en 1	0,0000	0,9309	0,0691
$= 0, 1, 2$	0,0002	0,7940	0,2060
$= 0$ t/m 3	0,0013	0,5886	0,4114
$= 0$ t/m 4	0,0059	0,3704	0,6296
$= 0$ t/m 5	0,0207	0,1958	0,8042
$= 0$ t/m 6	0,0577	0,0867	0,9133
$= 0$ t/m 7	0,1316	0,0322	0,9678
$= 0$ t/m 8	0,2517	0,0100	0,9900
$= 0$ t/m 9	0,4119	0,0026	0,9974

α_Z en β_Z tot een compromis moeten komen. Bij het onderhavige probleem kan dit worden bereikt door de in tabel 6.3 weergegeven zone te kiezen. Hiervoor is $\alpha_Z = 0{,}0577$ en $\beta_Z = 0{,}0867$. De kans op het ten

onrechte verwerpen van H_0 (de fout met de ernstigste consequenties) is dan wat kleiner dan de kans op het ten onrechte verwerpen van H_1, terwijl beide kansen — gezien de betrekkelijk kleine steekproefomvang — nog redelijk klein zijn. Wil men de eerstgenoemde kans nog verkleinen, dan kan men de zone $x = 0$ t/m 5 kiezen met $\alpha_Z = 0{,}0207$, maar dan is $\beta_Z = 0{,}1958$, dus aanzienlijk groter.

Het beslissingsschema. Voor het framboesia-probleem komen wij op grond van de voorgaande overwegingen tot het volgende beslissingsschema (zie figuur 6.2):

1. Verwerp H_0 indien de waargenomen waarde van $\underset{\sim}{x}$ (x_0 = het aantal lijders in de steekproef) in de kritieke zone $x = 0(1)6$ ligt. Indien men besluit tot het verwerpen van H_0, bedraagt de kans dat dit ten onrechte geschiedt ten hoogste 0,0577.

2. Verwerp H_0 niet indien x_0 buiten de kritieke zone ligt. Het niet verwerpen van H_0 houdt het verwerpen van H_1 in. Komt men tot deze uitspraak, dan is de kans dat zij fout is ten hoogste gelijk aan 0,0867.

Het waargenomen aantal framboesialijders bedraagt $x_0 = 8$ (zie 6.1.1). Daar deze waarde buiten de kritieke zone ligt, wordt H_0 niet verworpen. Merk op dat wij niet zeggen: H_0 wordt aanvaard. Hoewel deze formulering veel gebruikt wordt, is zij onjuist. In het algemeen zal immers $H_0 : P = 0{,}5$ niet de enige hypothese zijn, die niet verworpen kan worden (zo zal men bij het gegeven voorbeeld ook hypothesen als $P = 0{,}51$ of $P = 0{,}48$ niet kunnen verwerpen). Men kan daarom niet-verwerpen niet gelijkstellen met aanvaarden.

6.1.5. DE INVLOED VAN DE STEEKPROEFOMVANG OP DE FOUTEN-KANSEN

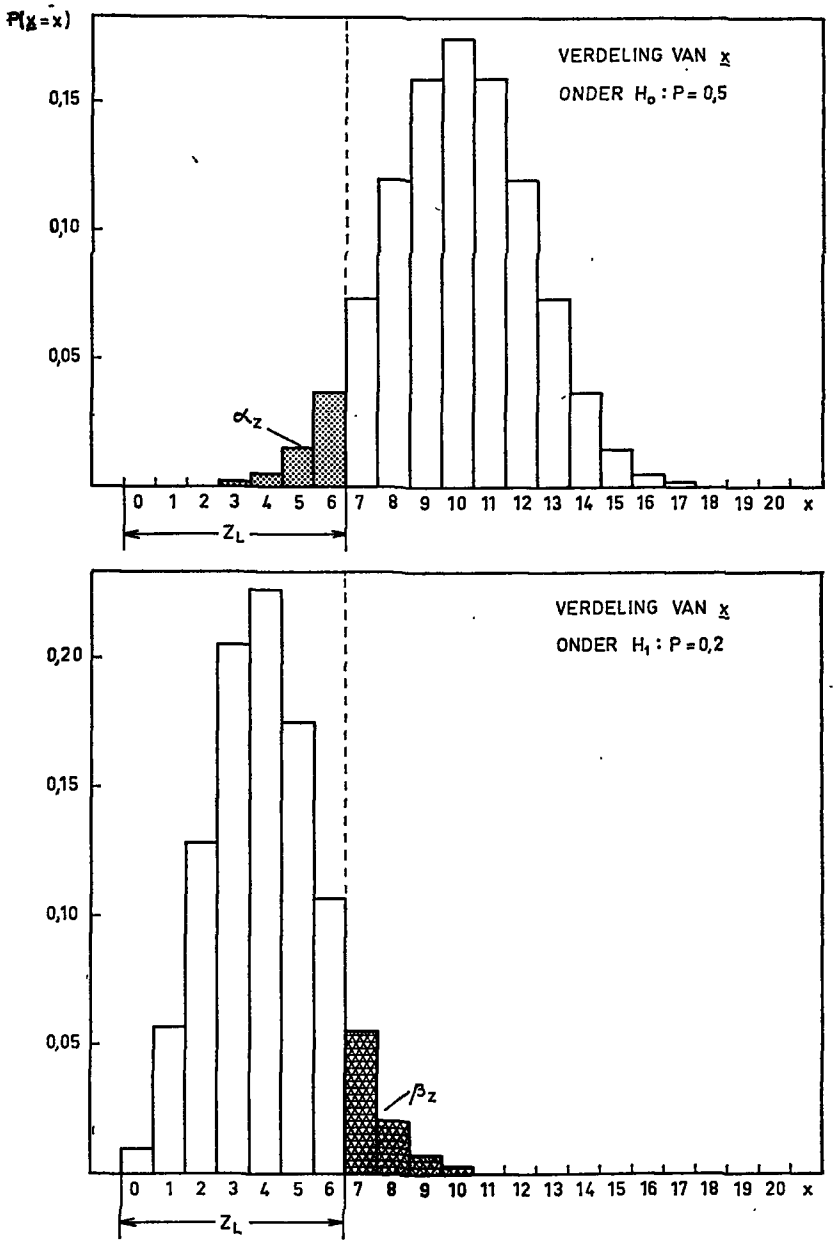
Veronderstel dat bij het framboesia-probleem de steekproef tot 30 personen wordt vergroot. De grootheid $\underset{\sim}{x}$ volgt dan één der volgende kansverdelingen:

$$\text{onder } H_0 : P(\underset{\sim}{x} = x \mid H_0) = \frac{30!}{x!(30-x)!} \cdot 2^{-30},$$

$$\text{onder } H_1 : P(\underset{\sim}{x} = x \mid H_1) = \frac{30!}{x!(30-x)!} (0{,}2)^x (0{,}8)^{30-x}.$$

Deze kansverdelingen zijn in tabel 6.5 opgenomen. In tabel 6.6 hebben wij voor verschillende linker kritieke zones de waarden van α_Z, β_Z en $1 - \beta_Z$ vermeld.

Vergelijkt men tabel 6.6 ($n = 30$) met tabel 6.4 ($n = 20$), dan blijkt, dat men door het vergroten van de steekproefomvang kan bereiken,



Figuur 6.2. Kansverdelingen van $\bar{x}$ onder H_0 en H_1 voor $n = 20$.

Tabel 6.5. Verdelingen van de grootheid $\underline{x}$ voor $n = 30$
$H_0 : P = 0{,}5 \quad H_1 : P = 0{,}2$

x	$P(\underline{x} = x \mid H_0)$		$P(\underline{x} = x \mid H_1)$	
30	0,00000		0,00000	
29	0,00000		0,00000	
28	0,00000		0,00000	
27	0,00000		0,00000	
26	0,00003		0,00000	
25	0,00013		0,00000	
24	0,00055		0,00000	
23	0,00190		0,00000	
22	0,00545		0,00000	
21	0,01332		0,00000	$\beta_Z = 0{,}02561$
20	0,02798		0,00000	Fout II
19	0,05088		0,00000	
18	0,08055		0,00000	
17	0,11154		0,00001	
16	0,13544		0,00004	
15	0,14446		0,00018	
14	0,13544		0,00067	
13	0,11154		0,00221	
12	0,08055		0,00638	
11	0,05088		0,01612	
10	0,02798		0,03547	
9	0,01332		0,06756	
8	0,00545		0,11056	
7	0,00190		0,15382	
6	0,00055		0,17946	
Z_L 5	0,00013	$\alpha_Z = 0{,}04936$	0,17228	$1 - \beta_Z = 0{,}97438$
4	0,00003	Fout I	0,13252	Onderscheidings-
3	0,00000		0,07853	vermogen
2	0,00000		0,03366	
1	0,00000		0,00928	
0	0,00000		0,00124	
Totaal	1,00000		0,99999	

Tabel 6.6. Kansen op een fout van de eerste en van de tweede soort en het onderscheidingsvermogen voor verschillende Z_L (ontleend aan tabel 6.5)
$H_0 : P = 0{,}5 \quad H_1 : P = 0{,}2 \quad n = 30$

Z_L bestaat uit	Fout I α_Z	Fout II β_Z	Ondersch. verm. $1 - \beta_Z$
$x =$ 0 t/m 5	0,00016	0,57249	0,42751
0 t/m 6	0,00071	0,39303	0,60697
0 t/m 7	0,00261	0,23921	0,76079
0 t/m 8	0,00806	0,12865	0,87135
0 t/m 9	0,02138	0,06109	0,93891
0 t/m 10	0,04936	0,02562	0,97438
0 t/m 11	0,10024	0,00950	0,99050
0 t/m 12	0,18079	0,00312	0,99688

dat aan een bepaalde waarde van α_Z een veel kleinere β_Z verbonden is. Bijvoorbeeld:

$$n = 20,\ Z_L\colon x = 0 \text{ t/m } 5,\ \alpha_Z = 0{,}0207,\ \beta_Z = 0{,}1958,$$
$$n = 30,\ Z_L\colon x = 0 \text{ t/m } 9,\ \alpha_Z = 0{,}0214,\ \beta_Z = 0{,}0611.$$

(Vergroot men de steekproefomvang tot $n = 50$, dan kan men bij de keuze van de zone Z_L: $x = 0$ t/m 14 bijv. werken met $\alpha_Z = 0{,}0130$ en $\beta_Z = 0{,}0607$).

Bij een toetsingsprobleem met twee enkelvoudige hypothesen, zoals het probleem dat wij tot dusverre hebben besproken, zou men ook als volgt tewerk kunnen gaan:

a. Stel vast, welke foutenkansen (α èn β) men aanvaardbaar acht.

b. Bepaal op grond van P_0, P_1 en de voor α en β gekozen waarde, de steekproefomvang waarbij de kritieke zone zodanig kan worden gekozen, dat α_Z en β_Z ten hoogste gelijk zijn aan de gekozen α en β.

Voorbeeld

6.3. Men wil beslissen tussen de mogelijkheden $P_0 = 0{,}9$ en $P_1 = 0{,}4$ met $\alpha = \beta = 0{,}05$. Met behulp van tabellen van de binomiale verdeling kan men vaststellen, dat bij $n = 25$ de linker kritieke zone $x = 0$ t/m 19 oplevert: $\alpha_Z = 0{,}0334$ en $\beta_Z = 0{,}0294$.

Deze ideale toetsingsprocedure, waarbij men de kans op beide fouten a priori kan stipuleren, kan in de praktijk vaak niet worden toegepast, omdat de alternatieve hypothese samengesteld, d.w.z. van het type $P_1 < P_0$, $P_1 > P_0$ of $P_1 \neq P_0$ is. In 6.2 bespreken wij hoe men deze toetsing kan uitvoeren.

6.1.6. HET VASTSTELLEN VAN DE STEEKPROEFOMVANG BIJ GEGEVEN WAARDEN VAN P_0, P_1, α EN β

Als men de in 6.1.5 genoemde toetsingsprocedure wil toepassen en dus de hypothese $P = P_0$ wil toetsen (met een kans op een fout van de eerste soort α) met als tegenhypothese $P = P_1$ (met een kans op een fout van de tweede soort β), dan kan men de vereiste steekproefomvang eenvoudig bij benadering bepalen met formule (6.4). Deze formule berust op de volgende gedachtengang:

Onderstel dat $P_1 < P_0$. De toetsing is dan links éénzijdig. Indien men de waarde van de toetsingsgrootheid $\underline{x}$, die bij de steekproefomvang n de bovengrens vormt van de linker kritieke zone met $\alpha_Z \leqslant \alpha$ en $\beta_Z \leqslant \beta$, x_n noemt geldt

$$P(\underline{x} \leqslant x_n \mid n, H_0) \leqslant \alpha \text{ en } P(\underline{x} \geqslant x_n \mid n, H_1) \leqslant \beta.$$

Wij veronderstellen nu, dat de kansverdeling van $\underline{x}$ zowel onder H_0 als onder H_1 bevredigend kan worden benaderd door de best-passende

normale verdeling. Dan geldt als men de continuïteitscorrectie verwaarloost (zie figuur 6.3)

$$T_\alpha = -T_{1-\alpha} = \frac{x_n - P_0 n}{\sqrt{P_0 Q_0 n}} \text{ en } T_{1-\beta} = \frac{x_n - P_1 n}{\sqrt{P_1 Q_1 n}},$$

zodat

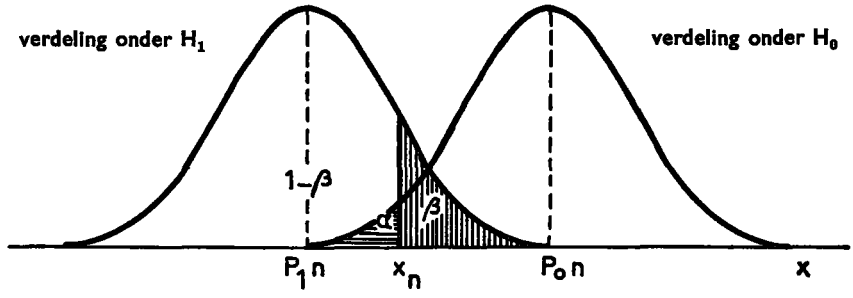
(6.3.1) $$x_n - P_0 n = -T_{1-\alpha}\sqrt{P_0 Q_0 n}$$

(6.3.2) $$\frac{x_n - P_1 n = T_{1-\beta}\sqrt{P_1 Q_1 n}}{n(P_0 - P_1) = T_{1-\alpha}\sqrt{P_0 Q_0 n} + T_{1-\beta}\sqrt{P_1 Q_1 n}}$$

$$n(P_0 - P_1) = \sqrt{n}[T_{1-\alpha}\sqrt{P_0 Q_0} + T_{1-\beta}\sqrt{P_1 Q_1}]$$

(6.4) $$n = \left[\frac{T_{1-\alpha}\sqrt{P_0 Q_0} + T_{1-\beta}\sqrt{P_1 Q_1}}{P_0 - P_1}\right]^2.$$

Voor $P_1 > P_0$ wordt de noemer van (6.4): $P_1 - P_0$. De teller verandert niet.



Figuur 6.3. Best-passende normale verdelingen onder H_0 en H_1 ($P_0 > P_1$).

Voorbeeld

6.4. Stel dat: $H_0 : P = 0{,}6$, met $\alpha = 0{,}01$, zodat $T_{1-\alpha} = T_{0,99} = 2{,}33$;
$H_1 : P = 0{,}5$, met $\beta = 0{,}05$, zodat $T_{1-\beta} = T_{0,95} = 1{,}65$.
Volgens (6.4) geldt dan bij benadering

$$n = \left[\frac{2{,}33\sqrt{0{,}6 \times 0{,}4} + 1{,}65\sqrt{0{,}5 \times 0{,}5}}{0{,}6 - 0{,}5}\right]^2 = (19{,}66)^2 = 386{,}5 = 387.$$

Uit (6.3.2) volgt:

$$x_n = P_1 n + T_{1-\beta}\sqrt{P_1 Q_1 n} = 193{,}5 + 1{,}65\sqrt{0{,}5 \times 0{,}5 \times 387} = 209{,}7 = 209.$$

(Waarom wordt hier naar beneden afgerond?)
Men kan de toetsing dus als volgt uitvoeren: Trek aselect een steekproef van 387 exemplaren. Verwerp H_0 voor $x_0 < 209$. Verwerp H_0 niet voor $x_0 > 209$.

Formule (6.4) geeft aanleiding tot enkele opmerkingen:

1. De kans dat bij een toetsing een fout (van de eerste òf van de tweede soort) wordt gemaakt, is altijd groter dan 0. Men kan deze kansen echter bij gegeven P_0 en P_1 willekeurig klein kiezen. Naarmate men α en/of β kleiner maakt, wordt n groter.

2. Als men α en β heeft gekozen kan men altijd een zodanige steekproefomvang kiezen, dat men een beslissing kan nemen tussen twee waarden P_0 en P_1, waarbij de kans op het ten onrechte verwerpen van H_0 ten hoogste gelijk is aan α, resp. de kans op het ten onrechte verwerpen van H_1 ten hoogste gelijk is aan β. De steekproefomvang kan dan, zoals uit tabel 6.7 blijkt, *zo* groot worden dat het onderzoek praktisch onmogelijk wordt.

Tabel 6.7. Steekproefomvang, nodig om bij $\alpha = \beta = 0{,}05$ te beslissen tussen $P_0 = 0{,}5$ en verschillende P_1

P_1	n	P_1	n
0,2	30	0,45	1084
0,3	63	0,475	3964
0,4	266	0,49	27225

Wanneer men α en β zeer klein gekozen heeft kan men in de eerste plaats overwegen of voor α en/of β grotere waarden toelaatbaar zijn. Wil men aan deze keuze niet tornen, dan kan men zich afvragen of een groter verschil tussen P_0 en P_1 kan worden toegelaten, dan wel of het verschil dat men wil constateren praktisch van voldoende betekenis is om het aan te tonen. Onderstel bijv. dat bij het framboesiavoorbeeld de beide hypothesen luiden $H_0 : P = 0{,}5$ en $H_1 : P = 0{,}45$. Beide fracties wijzen op een dusdanige ernst van de aandoening, dat een uitgebreid saneringsprogramma noodzakelijk is en dat het weinig uitmaakt, welke van deze twee waarden juist is. Pas indien beide hypothetische waarden van P meer verschillen wordt het van belang, ze van elkaar te onderscheiden. *Hoeveel meer* is een (moeilijke) vraag, waarvan de beantwoording niet meer op statistisch terrein ligt.

Te veel onderzoeken worden nog uitgevoerd met een min of meer lukraak gekozen steekproefomvang. Soms is de steekproef niet groot genoeg om tot een beslissing te kunnen komen. In andere gevallen is de steekproef te groot, d.w.z. men zou reeds bij een kleinere omvang (en dus efficiënter) een beslissing bereikt kunnen hebben. Uit het voorgaande blijkt hoe belangrijk het is, dat men zich bij de *opzet van een onderzoek* rekenschap geeft van de factoren die bij de *statistische analyse* van de uitkomsten van het onderzoek een rol spelen.

6.1.7. OPGAVEN

6.1. Op basis van een aselecte steekproef van 20 elementen toetst men de hypothese $H_0 : P = 0{,}6$ met als alternatief $H_1 : P = 0{,}9$. Stel een beslissingsschema op waarbij $\alpha_Z + \beta_Z$ minimaal is (hierbij kan gebruik worden gemaakt van tabel 4.3).
a. Wat is de kritieke zone en hoe groot zijn α_Z, β_Z en $1 - \beta_Z$?
b. De getrokken steekproef levert $x_0 = 16$ op. Tot welke uitspraak leidt het schema?

6.2. Als opgave 6.1, maar met de gegevens: $P_0 = 0{,}7$, $P_1 = 0{,}3$, $n = 20$, $x_0 = 8$.

6.3. Men wil de hypothese $P > 0{,}6$ toetsen met als tegenhypothese $P < 0{,}4$. Men kan de toetsing uitvoeren met als $H_0 : P = 0{,}6$ en als $H_1 : P = 0{,}4$; waarom? Bepaal bij benadering de steekproefomvang die nodig is om te toetsen met onderstaande waarden van α en β:
a. $\alpha = \beta = 0{,}05$ *b.* $\alpha = 0{,}01$, $\beta = 0{,}05$ *c.* $\alpha = \beta = 0{,}01$.
Geef voor elk van deze toetsingen de kritieke waarde.

6.2. Toetsing van de hypothese $P = 0{,}5$

6.2.1. ALGEMENE GANG VAN ZAKEN

In de praktijk doet zich vaak de situatie voor, dat de hypothese $P = 0{,}5$ wordt getoetst met één der volgende *samengestelde* alternatieve hypothesen:

1. $P < 0{,}5$ (links éénzijdige toetsing),
2. $P > 0{,}5$ (rechts éénzijdige toetsing),
3. $P \neq 0{,}5$ (tweezijdige toetsing; zowel $P < 0{,}5$ als $P > 0{,}5$ wordt mogelijk geacht).

Het is duidelijk dat de in 6.1 beschreven procedure niet zonder meer kan worden toegepast, daar geen specifieke waarde P_1 is gegeven zodat de kansverdeling van $\underline{x}$ onder H_1 niet ondubbelzinnig kan worden bepaald. De toetsing verloopt in dit geval als volgt:

a. Kies een kleine fractie α. Men noemt deze fractie de *onbetrouwbaarheidsdrempel* van de toets (Eng.: level of significance). De waarde van α representeert de (maximale) kans op een fout van de eerste soort (d.w.z. het ten onrechte verwerpen van H_0) die men aanvaardbaar acht. Vaak gebruikte drempelwaarden zijn 0,05, 0,01 en 0,001.

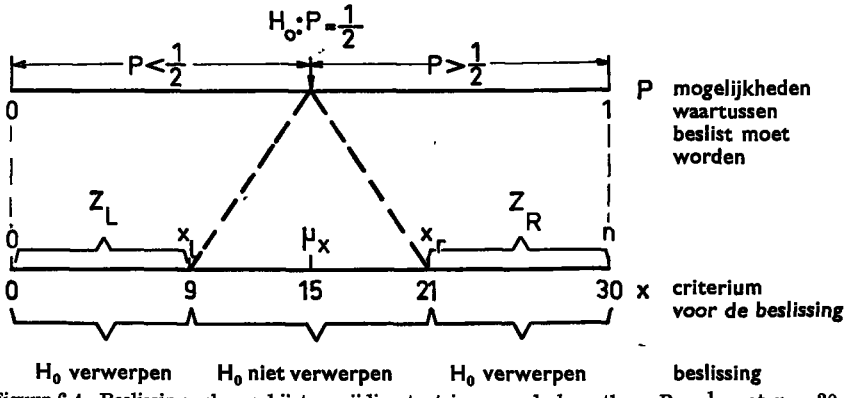
Men dient zich bij de keuze van α te realiseren, dat men deze drempelwaarde bij een bepaalde steekproefomvang slechts kan verkleinen ten koste van een vergroting van de kans op een fout van de tweede soort (het ten onrechte verwerpen van H_1). Overigens is deze keuze afhankelijk van de ernst van de consequenties, die het ten onrechte verwerpen van H_0 met zich meebrengt; zie 6.1.4.

b. Kies de (één- of tweezijdige) *kritieke zone* zo, dat onder H_0 de kans dat de toetsingsgrootheid $\underline{x}$ een in Z gelegen waarde aanneemt, ten hoogste gelijk is aan α, d.w.z. $\alpha_Z \leqslant \alpha$.

Links éénzijdige toetsing: Kies x_L zodanig dat $P(\underline{x} = 0) + P(\underline{x} = 1) + \ldots + P(\underline{x} = x_L) = P_L(x_L) \leqslant \alpha$ is.

Rechts éénzijdige toetsing: Kies x_R zodanig dat $P(\underline{x} = x_R) + P(\underline{x} = x_R + 1) + \ldots + P(\underline{x} = n) = P_R(x_R) \leqslant \alpha$ is.

Tweezijdige toetsing: De verdeling van $\underline{x}$ onder H_0 is symmetrisch, daar $P_0 = 0{,}5$. Men kiest nu $Z = Z_L + Z_R$ zodanig dat de kritieke waarden x_L en x_R van de sub-zones symmetrisch t.o.v. μ_x liggen, zodat: $x_R = n - x_L$ en: $P_L(x_L) + P_R(x_R) = 2P_L(x_L) = 2P_R(x_R) \leqslant P_D(x_L) = P_D(x_R)$ ten hoogste gelijk is aan α (zie figuur 6.4).



Figuur 6.4. Beslissingsschema bij tweezijdige toetsing van de hypothese $P = \frac{1}{2}$, met $n = 30$. Z bestaat uit twee even grote delen, zodanig dat $P_L(x_L) = P_R(x_R)$ ten hoogste gelijk is aan $\frac{1}{2}\alpha$.

c. Het *beslissingsschema* wordt dan:

1. Verwerp H_0 ten gunste van H_1, indien de waargenomen waarde van $\underline{x}$, x_0, in Z ligt. Men zegt in dit geval wel: de uitkomst is *significant* bij een drempelwaarde α (bij een $100\alpha\%$ drempel).
2. Verwerp H_0 niet, als x_0 niet in Z ligt.

Tabel 6.8. Kritieke zones corresponderend met de verschillende alternatieve hypothesen bij toetsing van $H_0 : P = 0{,}5$

Alternatief	Kritieke zone	Samenstelling van de zone
$P_1 < 0{,}5$	Links éénzijdig	$P(0) + P(1) + \ldots + P(x_L)$ ten hoogste gelijk aan α
$P_1 > 0{,}5$	Rechts éénzijdig	$P(n) + P(n-1) + \ldots + P(x_R)$ ten hoogste gelijk aan α
$P_1 \neq 0{,}5$	Tweezijdig	$P(0) + P(1) + \ldots + P(x_L)$ ten hoogste gelijk aan $\frac{\alpha}{2}$ $P(n) + P(n-1) + \ldots + P(x_R)$ ten hoogste gelijk aan $\frac{\alpha}{2}$

In tabel 6.8 hebben wij de keuze en samenstelling van Z bij één- en tweezijdige toetsing nogmaals samengevat.

Voorbeeld

6.5. Op basis van een 'steekproef' van 30 worpen wil men de hypothese toetsen, dat een geldstuk zuiver is. Men kiest $\alpha = 0{,}05$. De proef levert x_0 (aantal keren kruis) = 8 op. Tot welke uitspraak komt men?

In dit geval is $H_0 : P(\text{kruis}) = P = 0{,}5$ en $H_0 : P \neq 0{,}5$, daar men a priori niet weet in welke richting een eventuele onzuiverheid ligt dient tweezijdig te worden getoetst. Het beslissingsschema kan worden opgesteld met behulp van tabel 6.5, waarin de verdeling van x onder H_0 te vinden is. Passen wij het keuze-voorschrift toe dan blijken de (symmetrisch t.o.v. $\mu_x = Pn = 15$ gelegen) kritieke waarden te zijn (zie figuur 6.4):

$$\left.\begin{array}{l} x_L = 9,\ Z_L \text{ bestaat uit: } x = 0, 1, 2, \ldots, 9 \\ x_R = 21,\ Z_R \text{ bestaat uit: } x = 21, 22, 23, \ldots, 30 \end{array}\right\} Z = Z_L + Z_R.$$

De kans, dat x een waarde in deze zone aanneemt bedraagt:
$2(0{,}00000 + 0{,}00000 + \ldots + 0{,}00545 + 0{,}01332) = 2(0{,}02138) = 0{,}04276$.
Voegt men links en rechts resp. de waarden 10 en 20 toe, dan verkrijgt men een zone waarvan de onbetrouwbaarheid gelijk is aan:
$2(0{,}02138 + 0{,}02798) = 2(0{,}04936) = 0{,}09872$, zodat de gekozen drempelwaarde ver overschreden zou worden. Het beslissingsschema wordt nu:

1. Verwerp H_0 ten gunste van $P < 0{,}5$, als x_0 in Z_L (0 t/m 9) ligt.
2. Verwerp H_0 ten gunste van $P > 0{,}5$, als x_0 in Z_R (21 t/m 30) ligt.
3. Verwerp H_0 niet, als x_0 niet in Z ligt.

De waargenomen $x_0 = 8$ ligt in Z_L, zodat de hypothese $P = 0{,}5$ bij een 5% drempelwaarde verworpen wordt ten gunste van $P < 0{,}5$.

6.2.2. ÉÉNZIJDIGE OF TWEEZIJDIGE TOETSING

Bij bepaalde n en bepaalde drempelwaarde α is de linker éénzijdige zone groter dan het linkerdeel van de tweezijdige zone, dat bij deze n en α behoort. Men zal dus bij links éénzijdige toetsing een hogere waarde x_L verkrijgen dan bij tweezijdige toetsing. Dientengevolge bezit de links éénzijdige toets voor $P_1 < 0{,}5$ een kleinere kans op een fout van de tweede soort (= een groter onderscheidingsvermogen of een grotere *scherpte*) dan de tweezijdige toets. Vòòr het uitvoeren van een toetsing dient men zich daarom af te vragen of de alternatieve hypothese één- of tweezijdig is. Als men *weet, dat* $P > 0{,}5$ *niet kan voorkomen*, is links éénzijdige toetsing vanzelfsprekend. Als echter *tevoren vaststaat dat het effect van de uitspraken* $P > 0{,}5$ *en* $P = 0{,}5$ *hetzelfde is*, kan eveneens links éénzijdig worden getoetst. Uiteraard geldt m.m. hetzelfde voor de keuze tussen rechts éénzijdige en tweezijdige toetsing.

Voorbeeld

6.6. Beschouw nogmaals voorbeeld 6.5. Onderstel dat men weet dat getracht is het geldstuk zodanig te vervalsen, dat de uitkomst kruis te weinig optreedt. Indien deze vervalsing effect heeft, verwacht men dat $P(\text{kruis}) = P$ kleiner

dan 0,5 zal zijn. Nu zou men de hypothese $P = 0{,}5$ kunnen toetsen met als alternatief $P < 0{,}5$, zodat een linker éénzijdige kritieke zone wordt gekozen. Deze blijkt te bestaan uit de waarden 0 t/m 10, met $P_L(10) = 0{,}04936$ ($< 0{,}05$); zie tabel 6.6.

Deze zone is, bij dezelfde drempelwaarde 0,05, groter dan het linker deel van de zone die wij bij tweezijdige toetsing verkregen en bezit daardoor een kleinere kans op een fout van de tweede soort. Dit blijkt uit onderstaande tabel, waarin deze kans voor beide zones en enkele alternatieve waarden van P is opgenomen.

Kans op een fout van de tweede soort

$H_0 : P = 0{,}5 \quad n = 30$

	Waarde onder $H_1 : P =$		
	0,4	0,3	0,2
Z_L: 0 t/m 9	0,24464	0,04798	0,00259
Z_L: 0 t/m 10	0,12750	0,01716	0,00056

Tenslotte beschouwen wij nog het geval, waarin de hypothese $P_0 = 0{,}5$ wordt getoetst met bijv. als alternatief $P_1 \leqslant 0{,}2$. Hoewel de alternatieve hypothese samengesteld is, kan men toch de toetsing baseren op de enkelvoudige waarden $P_0 = 0{,}5$ en $P_1 = 0{,}2$. Toetst men nl. met een drempelwaarde α, dan bezit de bij de toets behorende critieke zone voor elke waarde van $P < 0{,}2$ een groter onderscheidingsvermogen dan voor $P = 0{,}2$.

6.2.3. DE WERKINGSKROMME

Als de alternatieve hypothese samengesteld is kan men een grafiek vervaardigen, waaruit bij een bepaalde steekproefomvang, resp. voor steekproeven van verschillende omvang, het onderscheidingsvermogen van de toets (en de complementaire kans op een fout van de tweede soort) bij verschillende drempelwaarden kan worden afgelezen.

Onderstel dat men de hypothese $P = 0{,}5$ toetst met als alternatieve hypothese $P > 0{,}5$ en het onderscheidingsvermogen bij $n = 10$ wil bestuderen van de rechter kritieke zones: *a*) $x = 9$, 10 en *b*) $x = 8$, 9, 10. In tabel 6.9 zijn voor $n = 10$ en een reeks met 0,05 opklimmende waarden van P de (binomiale) kansverdelingen van de toetsingsgrootheid x opgenomen.

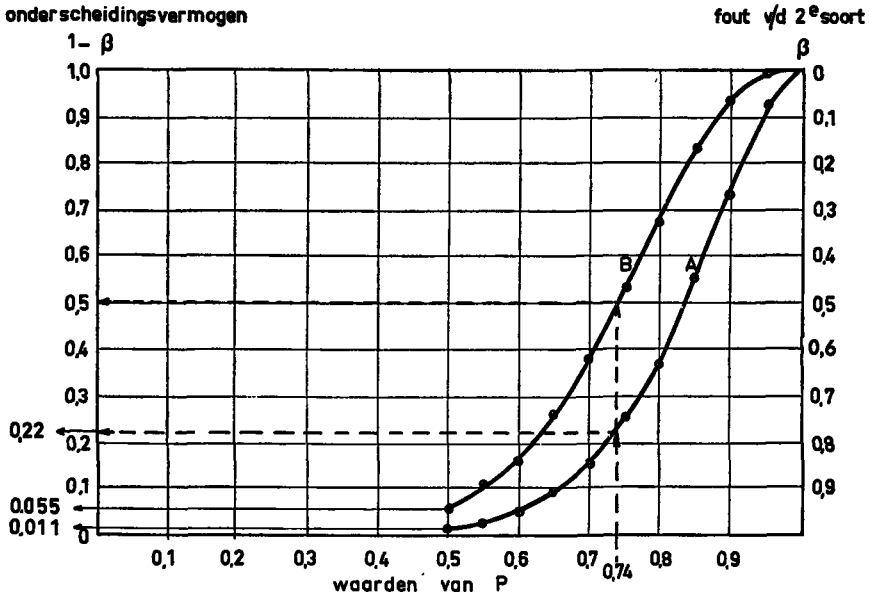
Voor zone A ($x = 9$, 10) lezen wij op de regel $P = 0{,}50$ als kans op een fout van de eerste soort $\alpha_Z = P(9) + P(10) = 0{,}010 + 0{,}001 = 0{,}011$ af. Voor de in de tabel voorkomende waarden van P groter dan 0,50 kan het onderscheidingsvermogen van zone A worden bepaald door de bij zo'n waarde behorende kansen van $x = 9$ en $x = 10$ te sommeren. De uitkomsten hiervan zijn rechts in tabel 6.9 opgenomen. Vervolgens is deze procedure herhaald voor de zone B ($x = 8$, 9, 10), met $\alpha_Z = 0{,}055$.

Tabel 6.9. Kansverdelingen van x voor $n = 10$ en $P = 0{,}05(0{,}05)0{,}95$

P	x											Som der kansen	Onderscheidingsvermogen	
	0	1	2	3	4	5	6	7	8	9	10		$ZR =$ 9 en 10	$ZR =$ 8 t/m 10
0,95	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,010	0,075	0,315	0,599	1,000	0,914	0,989
0,90	0,000	0,000	0,000	0,000	0,000	0,002	0,011	0,057	0,194	0,387	0,349	1,000	0,736	0,930
0,85	0,000	0,000	0,000	0,000	0,001	0,008	0,040	0,130	0,276	0,347	0,197	0,999	0,544	0,820
0,80	0,000	0,000	0,000	0,001	0,006	0,026	0,088	0,201	0,302	0,268	0,107	0,999	0,375	0,677
0,75	0,000	0,000	0,000	0,003	0,016	0,058	0,146	0,250	0,282	0,188	0,056	0,999	0,244	0,526
0,70	0,000	0,000	0,001	0,009	0,037	0,103	0,200	0,267	0,233	0,121	0,028	0,999	0,149	0,382
0,65	0,000	0,001	0,004	0,021	0,069	0,154	0,238	0,252	0,176	0,072	0,013	1,000	0,085	0,261
0,60	0,000	0,002	0,011	0,042	0,111	0,201	0,251	0,215	0,121	0,040	0,006	0,999	0,046	0,167
0,55	0,000	0,004	0,023	0,075	0,160	0,234	0,238	0,166	0,076	0,021	0,003	1,000	0,024	0,100
0,50	0,001	0,010	0,044	0,117	0,205	0,246	0,205	0,117	0,044	0,010	0,001	1,000	0,011	0,055
0,45	0,003	0,021	0,076	0,166	0,238	0,234	0,160	0,075	0,023	0,004	0,000	1,000		
0,40	0,006	0,040	0,121	0,215	0,251	0,201	0,111	0,042	0,011	0,002	0,000	0,999		
0,35	0,013	0,072	0,176	0,252	0,238	0,154	0,069	0,021	0,004	0,001	0,000	1,000		
0,30	0,028	0,121	0,233	0,267	0,200	0,103	0,037	0,009	0,001	0,000	0,000	0,999		
0,25	0,056	0,188	0,282	0,250	0,146	0,058	0,016	0,003	0,000	0,000	0,000	0,999		
0,20	0,107	0,268	0,302	0,201	0,088	0,026	0,006	0,001	0,000	0,000	0,000	0,999		
0,15	0,197	0,347	0,276	0,130	0,040	0,008	0,001	0,000	0,000	0,000	0,000	0,999		
0,10	0,349	0,387	0,194	0,057	0,011	0,002	0,000	0,000	0,000	0,000	0,000	1,000		
0,05	0,599	0,315	0,075	0,010	0,001	0,000	0,000	0,000	0,000	0,000	0,000	1,000		

De kansen zijn afgerond op drie decimalen. Tengevolge van afrondingsfoutjes is de som per verdeling (regel) niet steeds gelijk aan 1.

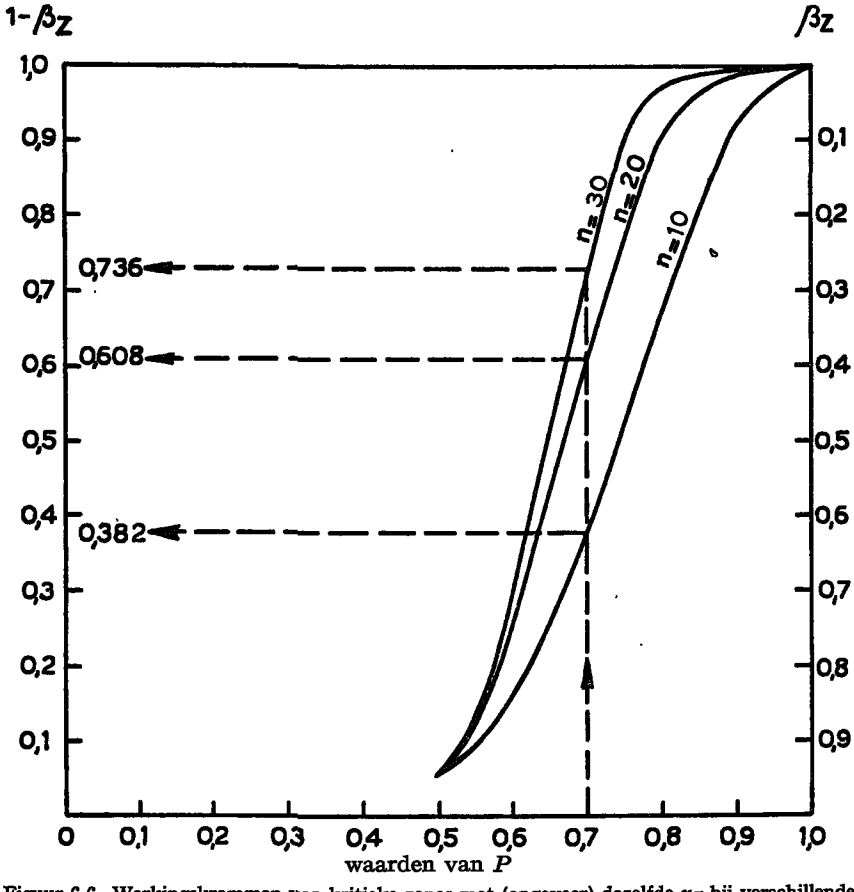
Uit de aldus verkregen gegevens kan een grafiek worden samengesteld, waarin langs de X-as de waarden van P en langs de Y-as het onderscheidingsvermogen (en/of de kans op een fout van de tweede soort) wordt afgezet: zie figuur 6.5. Voor elke zone kan in deze grafiek



Figuur 6.5. Werkingskrommen van twee rechter kritieke zones bij $n = 10$, A: $Z_R = 9$ en 10, B: $Z_R = 8$ t/m 10.

een reeks punten (P, $1 - \beta$) worden geplaatst. Als men deze punten door een vloeiende lijn verbindt verkrijgt men de zg. *werkingskromme* (Eng.: operating-characteristic curve, afgekort OC curve).

Uit de kromme kan men voor iedere alternatieve waarde van P bij benadering het onderscheidingsvermogen van de corresponderende zone, $1 - \beta_Z$ (of de kans op een fout van de tweede soort, β_Z) aflezen. Voor het alternatief $P_1 = 0{,}74$ vindt men bijv. (zie stippellijnen) voor zone A: $1 - \beta_Z = 0{,}22$ en voor zone B: $1 - \beta_Z = 0{,}50$. Het onderscheidingsvermogen van de grotere zone (B) blijkt in dit geval meer dan tweemaal zo groot te zijn als dat van de kleine (A), maar aan het gebruik van B is dan ook een onbetrouwbaarheid (kans op fout van de eerste soort) verbonden, die vijfmaal zo groot is als bij A.



Figuur 6.6. Werkingskrommen van kritieke zones met (ongeveer) dezelfde α_Z bij verschillende steekproefomvang en toetsing van $H_0 : P = \frac{1}{2}$:
$n = 10$, Z_R: 8 t/m 10, $\alpha_Z = 0{,}055$,
$n = 20$, Z_R: 14 t/m 20, $\alpha_Z = 0{,}058$,
$n = 30$, Z_R: 20 t/m 30, $\alpha_Z = 0{,}049$.

Men kan ook een grafiek samenstellen, waarin de werkingskrommen bij (ongeveer) dezelfde drempelwaarde voor steekproeven van verschillende omvang worden getekend. Hiermee kan men het onderscheidingsvermogen van de toets bij toenemende steekproefomvang bestuderen. In figuur 6.6 zijn als voorbeeld de werkingskrommen bij toetsing van de hypothese $P = 0{,}5$ voor rechter kritieke zones met (ongeveer) dezelfde α_Z voor drie steekproefgrootten getekend.

Bij tweezijdige toetsing krijgt men werkingskrommen die symmetrisch verlopen t.o.v. $P = 0{,}5$. Men tekene bijv. zelf met behulp van tabel 6.9 in figuur 6.5 de kromme van de tweezijdige zone bij $n = 10$ en met $x_L = 2$ en $x_R = 8$.

6.2.4. DE UITVOERING VAN DE TOETSING VIA OVERSCHRIJDINGSKANSEN

Tweezijdige toetsing. In 6.2.1 hebben wij gezien dat de kritieke waarden x_L en x_R bij tweezijdige toetsing zo worden gekozen dat

$$P_D(x_L) = P_D(x_R) = 2P_L(x_L) = 2P_R(x_R) = \alpha_Z \text{ ,}$$

zo groot mogelijk, maar ten hoogste gelijk aan de gekozen drempelwaarde α is. Hieruit volgt dat de tweezijdige overschrijdingskans van verder van $\mu_x = \frac{1}{2}n$ verwijderde waarden van x kleiner dan α is. De praktische consequentie hiervan is, dat men de kritieke zone niet volledig behoeft te bepalen, maar dat H_0 *bij tweezijdige toetsing kan worden verworpen, als* $P_D(x_0)$ *gelijk aan of kleiner dan* α *is.* Naarmate $P_D(x_0)$ kleiner is, vormt x_0 een sterkere aanwijzing tegen H_0 (men kan immers H_0 verwerpen bij *iedere* onbetrouwbaarheidsdrempel die ten minste gelijk is aan $P_D(x_0)$). Men geeft daarom bij de toepassing van een toets meestal niet alleen de conclusie, maar vermeldt ook $P_D(x_0)$ of de grenzen waartussen deze kans ligt.

Eénzijdige toetsing. Bij links éénzijdige toetsing kan $H_0 : P = 0{,}5$ bij een drempelwaarde α worden verworpen, indien $P_L(x_0) \leqslant \alpha$ is.

Bij rechts éénzijdige toetsing kan H_0 bij een drempelwaarde α worden verworpen, indien $P_R(x_0) \leqslant \alpha$ is.

6.2.5. BEPALING VAN OVERSCHRIJDINGSKANSEN

De overschrijdingskans van een waargenomen x_0 kan als volgt worden bepaald:

1. *Exact*: Met behulp van de in 4.3.2 genoemde tabellen.

2. *Normale benadering*: Voor $n \geqslant 30$ kan een binomiale verdeling met $P = 0{,}5$ worden benaderd door de best-passende normale verdeling (zie 5.4). De *tweezijdige overschrijdingskans* van x_0 kan bij benadering worden gevonden door te berekenen:

(6.5) $$T_0 = \frac{|\, x_0 - 0{,}5n \,| - 0{,}5}{0{,}5\sqrt{n}}$$

en de tweezijdige overschrijdingskans van T_0 in tabel A op te zoeken.

Bij *links éénzijdige* toetsing komt verwerping van H_0 uitsluitend in aanmerking, als $x_0 < 0{,}5n$ is (bij $x_0 \geqslant 0{,}5n$ kan men direct tot niet verwerpen besluiten). Men bepaalt dan de linkse overschrijdingskans van

(6.6.1) $$T_0 = \frac{x_0 - 0{,}5n + 0{,}5}{0{,}5\sqrt{n}}.$$

Bij *rechts éénzijdige* toetsing besluit men direct tot het niet verwerpen van H_0 indien $x_0 \leqslant 0{,}5n$ is. Voor $x_0 > 0{,}5n$ bepaalt men de rechtse kans van

(6.6.2) $$T_0 = \frac{x_0 - 0{,}5n - 0{,}5}{0{,}5\sqrt{n}}.$$

Als men bij gebruik van de normale benadering uitsluitend wil vaststellen of de berekende T_0 significant is bij een drempelwaarde 0,05, 0,01 of 0,001, kan men gebruik maken van tabel 6.10.

Tabel 6.10. Waarden van T, waarbij H_0 wordt verworpen bij tweezijdige en éénzijdige toetsing

Drempel α	Tweezijdig (6.5)	Links éénzijdig (6.6.1)	Rechts éénzijdig (6.6.2)
0,05	$1{,}96 < T < 2{,}58$	$-1{,}65 > T > -2{,}33$	$1{,}65 < T < 2{,}33$
0,01	$2{,}58 < T < 3{,}29$	$-2{,}33 > T > -3{,}09$	$2{,}33 < T < 3{,}09$
0,001	$T > 3{,}29$	$T < -3{,}09$	$T > 3{,}09$

3. *Met tabel I* (in bijlage 1): Voor $n \geqslant 100$ kunnen de grenzen, waartussen de tweezijdige overschrijdingskans van x_0 ligt direct uit deze tabel worden afgelezen. Tabel I bevat voor de drempel waarden 0,01, 0,02, 0,05 en 0,10 bij *tweezijdige toetsing* de linker kritieke waarden x_L van x. De hiermee corresponderende rechter kritieke waarden volgen uit: $x_R = n - x_L$.

Voor $n = 20$ vindt men bijv.:

Tabel I

$n = 20$	α			
	0,01	0,02	0,05	0,10
x_L	3	4	5	5
zodat: x_R	17	16	15	15

Hieruit volgt:

$0{,}02 < P_D(5) \leqslant 0{,}05$ (tabel 6.1 geeft de exacte kans: 0,0414)

$0{,}01 < P_D(16) \leqslant 0{,}02$ (tabel 6.1 geeft de exacte kans: 0,0118).

Tabel I kan voor *éénzijdige toetsing* worden gebruikt, als men de *drempelwaarden halveert*. Voor $n = 20$ leest men bijv. af:

$$P_L(4) \leqslant 0{,}01, \quad P_R(16) \leqslant 0{,}01, \quad P_L(3) \leqslant 0{,}005.$$

Voorbeelden

6.7. Men toetst de hypothese $P = 0{,}5$ (alternatieve hypothese: $P \neq 0{,}5$). Een aselecte steekproef van 50 elementen levert $x_0 = 15$ op. Is deze uitkomst significant bij een 5% drempel? (D.w.z. kan men H_0 verwerpen, als $\alpha = 0{,}05$ gekozen wordt?)

Er dient tweezijdig te worden getoetst. Ter illustratie wordt de tweezijdige overschrijdingskans van x_0 op drie manieren bepaald:

1. Exact. Met een tabel van de binomiale verdeling met $P = 0{,}5$ en $n = 50$ vindt men: $P_D(15) = 0{,}006598$.
2. Normale benadering. Toepassing van (6.5) levert op:

$$T_0 = \frac{|\,15 - 25\,| - 0{,}5}{0{,}5\sqrt{50}} = \frac{9{,}5}{3{,}536} = 2{,}69$$

$$P_D(2{,}69) = 2(0{,}00357) = 0{,}00714.$$

3. Tabel I. Op regel $n = 50$ en in kolom 0,05 staat: $x_L = 17$, zodat $P_D(17) < 0{,}05$. In kolom 0,01 staat: $x_L = 15$, zodat $P_D(15) < 0{,}01$.
   De uitspraak is dus: H_0 wordt verworpen ten gunste van $P < 0{,}5$.

6.8. Men toetst $H_0 : P = 0{,}5$ met als alternatief $H_0 : P > 0{,}5$ en kiest $\alpha = 0{,}001$. Een aselecte steekproef van 100 elementen levert $x_0 = 68$ op. Tot welke conclusie komt men?

Er wordt rechts éénzijdig getoetst. Uit tabel I blijkt dat $P_L(32) = P_R(100 - 32) = P_R(68) < 0{,}005$ is. Bij de gekozen onbetrouwbaarheidsdrempel (0,001) kan men op grond van deze tabel niet tot een uitspraak komen. Met de normale benadering vindt men via (6.6.2):

$$T_0 = \frac{68 - 50 - 0{,}5}{0{,}5\sqrt{100}} = \frac{17{,}5}{5} = 3{,}5.$$

Deze uitkomst overschrijdt de kritieke waarde ($T = 3{,}09$, tabel 6.10), zodat H_0 wordt verworpen.

6.2.6. OPGAVEN

6.4. Toets voor onderstaande aselecte steekproeven de hypothese $P = 0{,}5$. Geef uw beslissing en de overschrijdingskans van x_0 (bij gebruik van tabel I: de grenzen, waartussen deze kans ligt; bij gebruik van de normale benadering: de benaderende overschrijdingskans).

Steekproef	n	Alternatief	α	x_0	Beslissing	Overschr. kans
a	15	$P \neq 0{,}5$	0,05	3	.....	.....
b	15	$P \neq 0{,}5$	0,01	12	.....	.....
c	18	$P < 0{,}5$	0,05	5	.....	.....
d	20	$P > 0{,}5$	0,01	3	.....	.....
e	26	$P > 0{,}5$	0,05	19	.....	.....
f	40	$P \neq 0{,}5$	0,05	14	.....	.....
g	60	$P \neq 0{,}5$	0,01	40	.....	.....
h	80	$P > 0{,}5$	0,05	29	.....	.....
i	90	$P > 0{,}5$	0,05	54	.....	.....
j	90	$P \neq 0{,}5$	0,05	54	.....	.....
k	121	$P > 0{,}5$	0,05	70	.....	.....
l	324	$P < 0{,}5$	0,01	154	.....	.....
m	400	$P \neq 0{,}5$	0,05	220	.....	.....
n	400	$P > 0{,}5$	0,05	220	.....	.....
o	900	$P \neq 0{,}5$	0,01	486	.....	.....

6.3. Toetsing van de hypothese $P = P_0 \neq 0{,}5$

Naast de in 6.2 behandelde toets voor de hypothese $P = P_0 = 0{,}5$ is ook toetsing van de hypothese $P = P_0 \neq 0{,}5$ mogelijk. Toetsen van deze vorm noemt men wel *binomiale toetsen*, omdat de kansverdeling van de toetsingsgrootheid x een binomiale verdeling is. Deze toetsen worden dan ook op analoge wijze uitgevoerd. Hierbij zijn twee gevallen te onderscheiden, die wij achtereenvolgens behandelen.

6.3.1. EXACTE TOETSING

Als P_0 (vrij) ver van 0,5 verwijderd is en/of n betrekkelijk klein is dient exacte toetsing plaats te vinden. De verdeling van de toetsingsgrootheid x onder H_0 is dan niet symmetrisch, zodat het bepalen van tweezijdige overschrijdingskansen niet analoog aan het behandelde in 6.2 verloopt. Voorbeeld 6.9 verduidelijkt dit.

Voorbeeld

6.9. Men toetst de hypothese $P_0 = 0{,}2$ met als tegenhypothese $P_1 \neq 0{,}2$ en als drempelwaarde $\alpha = 0{,}05$. Een aselecte steekproef van 20 elementen levert $x_0 = 8$ op. Tot welke uitspraak komt men?

Er wordt tweezijdig getoetst. De verdeling van de toetsingsgrootheid x onder H_0 is te vinden in tabel 6.1 (laatste kolom). Om $P_D(x_0 = 8)$ te bepalen dient men nu alle waarden van x bijeen te zoeken, die een hoogstens even grote kans bezitten als $x_0 = 8$ en de daarbij behorende kansen te sommeren (zie 3.6.1). Dit zijn in de eerste plaats de waarden 8 t/m 20 (met als som van hun kansen 0,0322) en vervolgens de waarde 0, daar $P(0) = 0{,}0115$ kleiner is dan $P(8) = 0{,}0222$. Hieruit volgt dat $P_D(x_0 = 8) = 0{,}0322 + 0{,}0115 = 0{,}0437$. P_0 wordt dus verworpen ten gunste van $P_0 > 0{,}2$.

6.3.2. NORMALE BENADERING

Als P_0 niet te ver van 0,5 verwijderd is en/of n (vrij) groot is, kan de binomiale verdeling bevredigend worden benaderd door de bestpassende normale verdeling.

De *tweezijdige* overschrijdingskans van de waargenomen x_0 is bij benadering gelijk aan de tweezijdige kans van

$$T_0 = \frac{|x_0 - P_0 n| - 0{,}5}{\sqrt{P_0 Q_0 n}} \qquad (Q_0 = 1 - P_0). \tag{6.7}$$

Bij links éénzijdige toetsing is de gewenste kans bij benadering gelijk aan de linkse kans van

$$T_0 = \frac{x_0 - P_0 n + 0{,}5}{\sqrt{P_0 Q_0 n}}. \tag{6.8.1}$$

Bij rechts éénzijdige toetsing neemt men de rechtse kans van

$$T_0 = \frac{x_0 - P_0 n - 0{,}5}{\sqrt{P_0 Q_0 n}}. \tag{6.8.2}$$

Voorbeeld

6.10. Gegevens: H_0: $P < 0{,}6$, H_1: $P > 0{,}6$, $n = 400$, $\alpha = 0{,}01$. Men vindt $x_0 = 264$.

Er wordt rechts éénzijdig getoetst. Bij deze waarden van $P = 0{,}6$ en $n = 400$ kan de normale benadering worden toegepast. Uit (6.8.2) volgt:

$$T_0 = \frac{264 - (0{,}6)\ (400) - 0{,}5}{\sqrt{(0{,}6)\ (0{,}4)\ (400)}} = \frac{23{,}5}{9{,}8} = 2{,}4.$$

Uit tabel 6.10 blijkt, dat $P_R(2{,}4) < 0{,}01$, zodat H_0 wordt verworpen.

6.3.3. OPGAVEN

6.5. Gegevens: H_0: $P > 0{,}35$, H_1: $P < 0{,}35$, $n = 121$, $\alpha = 0{,}05$. Men vindt $x_0 = 27$. Hoe luidt de uitspraak?

6.6. Gegevens: H_0: $P < 0{,}45$, H_1: $P > 0{,}45$, $n = 100$, $\alpha = 0{,}01$. Men vindt $x_0 = 60$. Tot welke uitspraak komt men?

6.4. Sequente toetsing

6.4.1. ALGEMENE GANG VAN ZAKEN

Bij de toetsen die wij in de voorgaande paragrafen hebben behandeld, wordt een steekproef van bepaalde omvang gekozen, zonder dat daarbij acht wordt geslagen op de *tijdens* de trekking verschijnende uitkomsten. Gedurende de tweede wereldoorlog is echter door WALD en zijn medewerkers in de Statistical Research Group van de Columbia Universiteit (1945) de zg. *sequente toetsingsmethode* ontwikkeld, waarbij slechts zoveel elementen aselect worden getrokken als nodig zijn om bij tevoren gekozen waarden van α èn β tot een beslissing tussen H_0 en H_1 te komen. Dit aantal blijkt gemiddeld aanzienlijk lager te zijn dan de steekproefomvang, die men bij de gebruikelijke toetsingsmethode op grond van deze gegevens zou voorschrijven (bijv. op grond van P_0, P_1, α en β met formule (6.4)). Sequente toetsing is daardoor vooral te prefereren, indien men zo economisch mogelijk wil werken (bijv. als de elementen die men onderzoekt daarbij verloren gaan) of als men (bijv. bij een medisch onderzoek om ethische redenen) liefst zo weinig mogelijk personen in de steekproef wil opnemen.

Bij de sequente methode trekt men successievelijk aselect één element uit de populatie en verricht daaraan de waarneming. Na elke trekking komt men, op grond van de op dat moment beschikbare uitkomsten, tot één van de drie volgende beslissingen:

1. H_1 verwerpen ten gunste van H_0,
2. H_0 verwerpen ten gunste van H_1,
3. Een volgende trekking verrichten.

Teneinde te kunnen vaststellen welke beslissing men zal nemen, dient men bij elke steekproefomvang twee kansen te kennen:

$P_{0,n}$ = de kans op de verkregen uitkomst onder H_0,
$P_{1,n}$ = de kans op de verkregen uitkomst onder H_1.

Indien $P_{0,n}$ veel groter is dan $P_{1,n}$ besluit men in principe tot het verwerpen van H_1. Is $P_{0,n}$ veel kleiner dan $P_{1,n}$, dan besluit men tot het verwerpen van H_0. Verschillen $P_{0,n}$ en $P_{1,n}$ betrekkelijk weinig, dan wordt een volgende trekking verricht.

WALD (1947) heeft bewezen, dat voor sequente toetsing het volgende beslissingsschema geldt

(6.9) verwerp H_1, als $\frac{P_{1,n}}{P_{0,n}} \leqslant \frac{\beta}{1-\alpha}$,

(6.10) verwerp H_0, als $\frac{P_{1,n}}{P_{0,n}} \geqslant \frac{1-\beta}{\alpha}$,

(6.11) verricht een volgende trekking, als $\frac{\beta}{1-\alpha} < \frac{P_{1,n}}{P_{0,n}} < \frac{1-\beta}{\alpha}$.

Het steekproeftrekken wordt voortgezet, totdat beslissing 1 of beslissing 2 is genomen.

Opmerking:

De afleiding van deze ongelijkheden bevat bepaalde benaderingen. Hier wordt de som van α en β echter niet door verhoogd, terwijl de toeneming van het aantal waarnemingen tengevolge van een eventuele verkleining van deze kansen gering is.

6.4.2. RECHTS ÉÉNZIJDIGE SEQUENTE BINOMIALE TOETSING

Formules. Onderstel dat men wil beslissen tussen de hypothesen

$$H_0 : P = P_0 \text{ en } H_1 : P = P_1 (> P_0) \qquad [P = P(A)].$$

De kans op het ten onrechte verwerpen van H_0 wordt gelijk gesteld aan α en de kans op het ten onrechte verwerpen van H_1 aan β.

Er dient rechts éénzijdig te worden getoetst. Onder H_0 is de kans dat zich in een aselecte steekproef van n elementen x_n elementen met kenmerk A bevinden

$$P(\underline{x} = x_n \mid H_0) = P_{0,n} = C_{x_n}^n \cdot {P_0}^{x_n} \cdot {Q_0}^{n-x_n} \qquad (Q_0 = 1 - P_0).$$

Onder H_1 geldt

$$P(\underline{x} = x_n \mid H_1) = P_{1,n} = C_{x_n}^n \cdot {P_1}^{x_n} \cdot {Q_1}^{n-x_n} \qquad (Q_1 = 1 - P_1).$$

Hieruit volgt dat

$$(6.12) \qquad \frac{P_{1,n}}{P_{0,n}} = \frac{C_{x_n}^n \cdot {P_1}^{x_n} \cdot {Q_1}^{n-x_n}}{C_{x_n}^n \cdot {P_0}^{x_n} \cdot {Q_0}^{n-x_n}} = \left[\frac{P_1}{P_0}\right]^{x_n} \cdot \left[\frac{Q_1}{Q_0}\right]^{n-x_n}$$

De kritieke waarde, waarbij H_0 juist wordt verworpen bij een steekproefomvang n, noemen wij b_n. Uit (6.10) en (6.12) volgt

$$(6.13) \qquad \left[\frac{P_1}{P_0}\right]^{b_n} \cdot \left[\frac{Q_1}{Q_0}\right]^{n-b_n} = \frac{1-\beta}{\alpha}.$$

De kritieke waarde, waarbij H_1 juist wordt verworpen bij een steekproefomvang n, noemen wij a_n. Uit (6.9) en (6.12) volgt

$$\left[\frac{P_1}{P_0}\right]^{a_n} \cdot \left[\frac{Q_1}{Q_0}\right]^{n-a_n} = \frac{\beta}{1-\alpha}. \tag{6.14}$$

Door het nemen van logarithmen gaat (6.13) over in:

$$b_n \log \frac{P_1}{P_0} + (n - b_n) \log \frac{Q_1}{Q_0} = \log \frac{1-\beta}{\alpha},$$

$$b_n \log \frac{P_1}{P_0} + n \log \frac{Q_1}{Q_0} - b_n \log \frac{Q_1}{Q_0} = \log \frac{1-\beta}{\alpha},$$

$$b_n \left[\log \frac{P_1}{P_0} - \log \frac{Q_1}{Q_0} \right] = \log \frac{1-\beta}{\alpha} - n \log \frac{Q_1}{Q_0},$$

zodat

$$b_n = \frac{\log \frac{1-\beta}{\alpha}}{\log \frac{P_1}{P_0} - \log \frac{Q_1}{Q_0}} - n \frac{\log \frac{Q_1}{Q_0}}{\log \frac{P_1}{P_0} - \log \frac{Q_1}{Q_0}} \tag{6.15}$$

$$= \frac{\log \frac{1-\beta}{\alpha}}{\log \frac{P_1 Q_0}{P_0 Q_1}} + n \frac{\log \frac{Q_0}{Q_1}}{\log \frac{P_1 Q_0}{P_0 Q_1}}.$$

Geheel analoog volgt uit (6.14)

$$a_n = \frac{\log \frac{\beta}{1-\alpha}}{\log \frac{P_1 Q_0}{P_0 Q_1}} + n \frac{\log \frac{Q_0}{Q_1}}{\log \frac{P_1 Q_0}{P_0 Q_1}}. \tag{6.16}$$

Deze formules zien er op het eerste gezicht vervaarlijk uit, maar zij zijn eenvoudig te hanteren. Wij schrijven

$$s = \frac{\log \frac{Q_0}{Q_1}}{\log \frac{P_1 Q_0}{P_0 Q_1}} \qquad h_0 = \frac{\log \frac{\beta}{1-\alpha}}{\log \frac{P_1 Q_0}{P_0 Q_1}} \qquad h_1 = \frac{\log \frac{1-\beta}{\alpha}}{\log \frac{P_1 Q_0}{P_0 Q_1}}.$$

De vergelijkingen (6.16) en (6.15) gaan dan resp. over in

$$a_n = sn + h_0 \tag{6.17}$$

en

$$b_n = sn + h_1. \tag{6.18}$$

Uit (6.17) kan men voor elke steekproefomvang n het aantal elementen a_n met kenmerk A bepalen, waarbij H_1 kan worden verworpen met een kans op het ten onrechte verwerpen die ten hoogste gelijk aan β is. Uit (6.18) volgt het aantal b_n waarbij H_0 kan worden verworpen met een kans op een foutieve uitspraak die ten hoogste gelijk aan α is.

In tabel 6.11 zijn voor enkele regelmatig gebruikte waarden van α en β de waarden van $\log \frac{\beta}{1-\alpha}$ en $\log \frac{1-\beta}{\alpha}$ opgenomen. Verder behoeft men voor het bepalen van s, h_0 en h_1 alleen $\log Q_0/Q_1$ en $\log P_1Q_0/P_0Q_1$ te berekenen.

Tabel 6.11

α	β	$\log \frac{\beta}{1-\alpha}$	$\log \frac{1-\beta}{\alpha}$
0,01	0,01	—1,9956	1,9956
0,01	0,02	—1,6946	1,9912
0,01	0,05	—1,2967	1,9777
0,02	0,01	—1,9912	1,6946
0,02	0,02	—1,6902	1,6902
0,02	0,05	—1,2922	1,6767
0,05	0,01	—1,9777	1,2967
0,05	0,02	—1,6767	1,2922
0,05	0,05	—1,2787	1,2787

Toepassing. Wij passen de zojuist beschreven sequente toetsing toe op de gegevens van het voorbeeld in 6.1. Daar de formules op rechts éénzijdige toetsing betrekking hebben, kiezen wij

$$H_0 : P = 0{,}2 \text{ en } H_1 : P = 0{,}5.$$

Reeds eerder is gemotiveerd, dat de gevolgen van het ten onrechte verwerpen van H_1 het ernstigst worden geacht. Daarom kiezen wij $\beta = 0{,}01$ en $\alpha = 0{,}05$.

Uit tabel 6.11 volgt dat bij de gekozen waarden van α en β:

$$\log \frac{\beta}{1-\alpha} = -1{,}9777 \text{ en } \log \frac{1-\beta}{\alpha} = 1{,}2967.$$

$P_0 = 0{,}2$, $Q_0 = 0{,}8$, $P_1 = 0{,}5 = Q_1$, zodat:

$$\log Q_0/Q_1 = \log (0{,}8/0{,}5) = \log 1{,}6 = 0{,}2041$$
$$\log P_1Q_0/P_0Q_1 = \log (0{,}5 \times 0{,}8/0{,}2 \times 0{,}5) = \log 4 = 0{,}6021.$$

Zodoende is

$$s = \frac{0{,}2041}{0{,}6021} = 0{,}3390 \qquad h_0 = -\frac{1{,}9777}{0{,}6021} = -3{,}2847$$

$$h_1 = \frac{1{,}2967}{0{,}6021} = 2{,}1536$$

Vergelijking (6.17) luidt dus

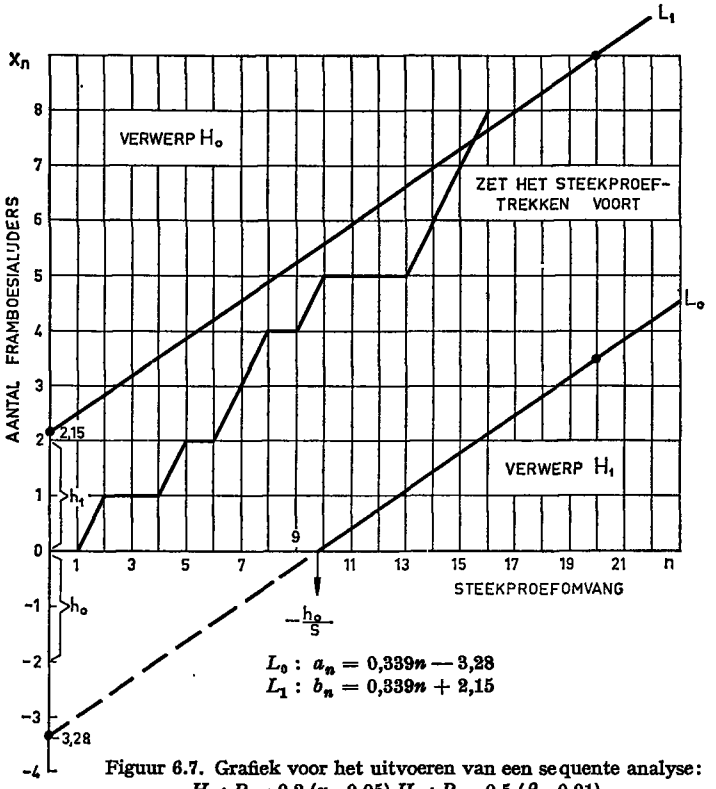
$$a_n = 0{,}339n - 3{,}28$$

en vergelijking (6.18) wordt

$$b_n = 0{,}339n + 2{,}15.$$

Het sequente beslissingsschema kan op twee manieren worden samengesteld:

a. De tabellarische methode. Voordat men tot het steekproeftrekken overgaat berekent men uit de verkregen vergelijkingen voor $n = 1, 2, 3, \ldots$ tot bijv. 30 de waarden van a_n en b_n. Daarna gaat men tot de sequente trekking over en schrijft voor $n = 1, 2, 3, \ldots$ het aantal waargenomen framboesialijders x_n in de tabel met de berekende waarden van a_n en b_n. Men verkrijgt dan een geleidelijk oplopende getallenreeks. Zodra x_n gelijk is aan b_n, wordt de hypothese $P_0 = 0{,}2$ verworpen en zodra x_n gelijk is aan a_n verwerpt men de hypothese $P_1 = 0{,}5$.



Figuur 6.7. Grafiek voor het uitvoeren van een sequente analyse: $H_0 : P = 0{,}2$ (α=0,05) $H_1 : P = 0{,}5$ (β=0,01).

b. De grafische methode. Deze is sprekender dan de tabellarische methode èn vereist minder rekenwerk. Men tekent hierbij de beide vergelijkingen in een grafiek met op de X-as de steekproefomvang n en op de Y-as het aantal framboesialijders x_n (zie figuur 6.7). Men verkrijgt dan de evenwijdige rechte lijnen L_0 (voor a_n) en L_1 (voor b_n).

De uitkomsten van de sequente trekking worden als volgt in deze grafiek getekend: men gaat, te beginnen met de eerste waarneming in de oorsprong, bij iedere getrokken persoon zonder framboesia één eenheid naar rechts en bij iedere persoon met framboesia tevens één eenheid naar omhoog, totdat men één der begrenzingslijnen, L_0 of L_1, overschrijdt. Wordt L_0 overschreden, dan besluit men tot het verwerpen van H_1; bij het overschrijden van L_1 besluit men tot het verwerpen van H_0. Onderstel bijv. dat successievelijk de volgende uitkomsten worden verkregen (+ = framboesia):

Trekking	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Uitkomst	—	+	—	—	+	—	+	+	—	+	—	—	—	+	+	+

De lijn, die met deze uitkomsten correspondeert, het *steekproefpad*, is in figuur 6.7 getekend. Bij de 16e trekking wordt L_1 gesneden. Het steekproeftrekken wordt stopgezet en de uitspraak luidt: H_0 verwerpen ten gunste van H_1.

De gemiddelde steekproefgrootte. Uit figuur 6.7 kan worden afgelezen, dat het steekproeftrekken vrij snel kan worden beëindigd, als de fractie framboesialijders in de populatie veel kleiner is dan $P_0 = 0{,}2$ of veel groter is dan $P_1 = 0{,}5$. Als alle personen in de populatie aan framboesia lijden eindigt het steekproeftrekken (met de uitspraak: H_0 verwerpen) bij $n = 4$. Deze steekproefomvang volgt ook uit de formule [1]

$$n = \frac{h_1}{1 - s} = \frac{2{,}1536}{0{,}6610} = 3{,}3 = 4 \qquad (P = 1). \tag{6.19}$$

Als er in het geheel geen personen met framboesia in de populatie aanwezig zijn, verloopt de steekproeflijn langs de X-as, zodat het steekproeftrekken dan bij $n = 10$ eindigt (met de uitspraak: H_1 verwerpen). Dit aantal kan berekend worden met

$$n = -\frac{h_0}{s} = -\frac{-3{,}2847}{0{,}3390} = 9{,}7 = 10 \qquad (P = 0). \tag{6.20}$$

Bedraagt de fractie framboesialijders P_0 (= 0,2) dan volgt n een kansverdeling, die afhangt van P_0, P_1, α en β. Bij het herhaald uitvoeren van sequente toetsing zou men dus een reeks verschillende waarden van n vinden. Bewezen kan worden, dat het gemiddelde van deze steekproefgrootten, de zg. *gemiddelde steekproefgrootte* (*GSG*), voor $P = P_0$ gelijk is aan

$$n = \frac{h_0 + \alpha(h_1 - h_0)}{P_0 - s} = \frac{-3{,}2847 + 0{,}05(2{,}1536 + 3{,}2847)}{0{,}2 - 0{,}3390} \tag{6.21}$$

$$= 21{,}7 = 22 \qquad (P = P_0).$$

[1] Bij deze en de vier volgende formules neme men steeds het kleinste gehele getal dat groter is dan de waarde, die uit de formule volgt.

Voor $P = P_1$ is de *GSG*:

$$(6.22) \quad n = \frac{h_1 + \beta(h_0 - h_1)}{P_1 - s} = \frac{2{,}1536 + 0{,}01(-3{,}2847 - 2{,}1536)}{0{,}5 - 0{,}3390} = 13 \qquad (P = P_1).$$

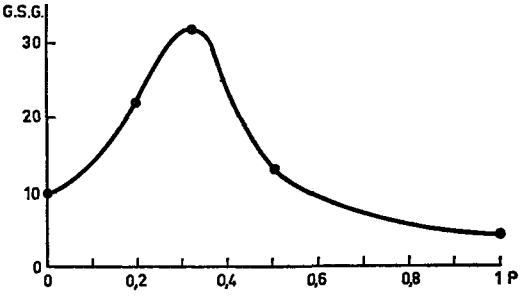
Tenslotte geldt, dat voor $P = s$ de *GSG* is:

$$(6.23) \quad n = -\frac{h_0 h_1}{s(1-s)} = -\frac{(-3{,}2847)\ (2{,}1536)}{0{,}3390\ (1 - 0{,}3390)} = 31{,}6 = 32 \qquad (P = s).$$

De berekende *GSG* zijn in tabel 6.12 samengevat. Men kan door deze 5 punten (*P*, *GSG*) een kromme schetsen, die voor de gegeven waarden $P_0(= 0{,}2)$, $P_1(= 0{,}5)$, $\alpha(= 0{,}05)$ en $\beta(= 0{,}01)$ ongeveer het verband tussen P en *GSG* aangeeft (zie figuur 6.8). Uit deze kromme blijkt, dat de *GSG* maximaal is tussen P_0 en P_1 (meestal: dicht bij $P = s$) en laag in de buurt van 0 en 1. De kromme komt hoger boven de X-as te liggen (de *GSG* neemt toe), naarmate P_0 en P_1 dichterbij elkaar komen en α en β kleiner worden. De kromme geeft dus een indruk van het aantal waarnemingen, dat gemiddeld verricht zal moeten worden om tot een beslissing te komen. Blijken deze gemiddelden groter te zijn, dan men praktisch te verwezenlijken acht, dan kan men voor de aanvang van het steekproeftrekken overwegen, of een andere keuze van P_0, P_1, α of β mogelijk is. Overigens blijkt de *GSG* steeds kleiner te zijn dan de (vaste) steekproefomvang, die men zou moeten kiezen om bij dezelfde P_0, P_1, α en β de toetsing op de gebruikelijke wijze (d.i. niet-sequent) uit te voeren.

Tabel 6.12. *GSG* voor het framboesiaprobleem

P	*GSG*
0	10
0,2	22
$s = 0{,}339$	32
0,5	13
1	4



Figuur 6.8. *GSG*-kromme.

Opmerking: Hoewel de *GSG* van de sequente binomiale toets relatief klein is, loopt men toch het risico dat het aantal waarnemingen groot wordt, voordat een beslissing kan worden genomen. Uit praktische overwegingen kan het noodzakelijk zijn een maximum n_m te stellen voor het aantal waarnemingen dat men wil verrichten. Indien men dit maximum bereikt, voordat L_0 of L_1 overschreden is, neemt men toch een beslissing, bijv. als

(6.24) $$x_{n_m} > \frac{h_0 + h_1}{2} + sn,$$ verwerpt men H_0,

en als

(6.25) $$x_{n_m} < \frac{h_0 + h_1}{2} + sn,$$ verwerpt men H_1.

Dit procédé noemt men een afgeknotte sequente toetsing (Eng. closed, truncated design). De kans op een foutieve conclusie wordt er door verhoogd. Als n_m echter niet te klein wordt gekozen (minstens op 2 à 3 maal de maximale waarde van de *GSG* wordt gesteld), is de kans op een foutieve conclusie slechts klein.

Een steekproefexperiment. Het volgende experiment, waarbij sequente toetsing wordt toegepast, leent zich goed voor klassikale uitvoering.

a. Iedere deelnemer tekent een grafiek voor het uitvoeren van een sequente analyse, gebaseerd op de volgende gegevens: $P_0 = 0{,}5$, $P_1 > 0{,}9$, $\alpha = \beta = 0{,}05$. P = de kans op kenmerk A.
(Ter controle vermelden wij, dat voor deze gegevens $s = 0{,}7325$, $h_0 = -1{,}34$, $h_1 = 1{,}34$. Merk op dat $h_0 = -h_1$ voor $\alpha = \beta$.)

b. Het steekproeftrekken wordt nagebootst door te werpen met een geldstuk (kruis = kenmerk A) of beter nog: met een dobbelsteen (1, 2 of 3 ogen = A). Iedere deelnemer werpt voor zichzelf en tekent de uitkomsten stuk voor stuk aan in de grafiek. Het aantal worpen (trekkingen) waarbij een uitspraak wordt bereikt, noteert men.

Als de gebruikte geldstukken (dobbelstenen) zuiver zijn, wordt getrokken uit een populatie met $P = 0{,}5$. Daar $\alpha = 0{,}05$ is gekozen, verwachten wij dat er op de 100 uitspraken ongeveer 5 onjuist zullen zijn (luiden: H_0 verwerpen). De verkregen steekproefgrootten kunnen voor alle uitgevoerde toetsingen in een frequentieverdeling worden samengevat. Het gemiddelde van deze verdeling kan worden vergeleken met de verwachte *GSG* volgens (6.21).

De in (*a*) gegeven waarden zijn zo gekozen, dat in de regel reeds bij een betrekkelijk klein aantal trekkingen een uitspraak wordt bereikt, zodat men eventueel ook individueel meerdere toetsingen kan verrichten, zonder dat het teveel tijd kost. Desgewenst kan men α en/of β en/of P_1 kleiner kiezen.

6.4.3. SEQUENTE BINOMIALE TOETSING MET DRIE BESLISSINGSMOGELIJKHEDEN

Deze sequente toets stemt overeen met een *tweezijdige* binomiale toets. Door DE BOER (1953) is een toets met drie beslissingsmogelijkheden ontwikkeld, waarvan het principe zeer eenvoudig is, daar zij bestaat uit twee sequente binomiale toetsen T en T' met als bepalende grootheden P_0, P_1, α en β, resp. P_0', P_1', α' en β', waarbij $P_0 < P_1 < P_0' < P_1'$. Ook deze toets kan worden afgeknot, waarbij dus tevoren een bovengrens wordt gesteld aan het aantal te verrichten waarnemingen.

Andere sequente procedures met drie beslissingsmogelijkheden zijn gegeven door ARMITAGE (1954, 1960, 1961). De eigenschappen van de afgeknotte methoden van DE BOER en ARMITAGE zijn door VAN

ELTEREN (1961) vergeleken. Daar VAN ELTEREN concludeert dat in vele gevallen de methoden van ARMITAGE te verkiezen zijn, zullen wij diens toetsen met drie beslissingsmogelijkheden aan de hand van een medisch voorbeeld in het kort bespreken[1].

Veronderstel dat men in een kliniek wil onderzoeken welke van twee behandelingsmethoden, A en B, de voorkeur verdient voor de behandeling van een bepaalde aandoening. Bij de sequente uitvoering van deze proef dienen telkens tweetallen patiënten te worden vergeleken. Zo'n tweetal kan soms worden gevonden door twee achtereenvolgens opgenomen patiënten met de beschouwde aandoening. Soms zal men tweetallen vormen die deze aandoening in ongeveer even ernstige vorm hebben, en/of van patiënten van hetzelfde geslacht of behorend tot dezelfde leeftijdsgroep. De toewijzing van de behandelingen kan voor elk tweetal aselect plaatsvinden.

Een groep patiënten die tot eenzelfde deelpopulatie behoort (met bijv. hetzelfde geslacht, dezelfde leeftijd en een ongeveer even ernstige aandoening) noemt men wel een *laag* (Eng. stratum). Men dient zo'n indeling van de patiënten in lagen (de stratificatie) niet te ver door te voeren, daar anders de mogelijkheid bestaat dat er in verschillende lagen niet-gepaarde patiënten overblijven, waardoor men waarnemingen verliest. Over de aselecte toewijzing schrijft ARMITAGE het volgende:

'The random allocation should be prepared before the start of the trial. . . . The allocation for each pair of observations can be done by tossing a coin. Alternatively, a table of random sampling numbers (aselecte getallen) can be used. If the two treatments are denoted by A and B, an *even* random digit can be taken to indicate the order AB, and *odd* digit indicating BA. Thus one random digit is required for each pair of observations.

This system of allocation has the disadvantage that, for between subject comparisons in which the treatments are not disguised in any way, the physician knows, for the second member of each pair, which treatment is to be used. He might occasionally be influenced (or be thought by a critic to have been influenced) by this knowledge in deciding whether or not a particular patient should enter the trial. To avoid this difficulty an allocation list can be prepared in which, within each stratum, each successive group of 4 (or 6 or 8) patients contains an equal number allocated to A and B. To achieve a balance after every 6, for instance, select a column of digits from a table of random numbers, and write down the first three numbers to occur out of the numbers 1—6. If these are, say, 4, 3 and 6, then treatment A will be allocated to the third, fourth and sixth patient in this group, and the others will receive B. The order for this group of 6 will thus be: $B\,B\,A\,A\,B\,A$. If the precise system used is not divulged by the person who constructs the list, the possible bias mentioned above will probably be avoided. In the sequential analysis the first A in each balanced group is paired with the first B, the second A with the second B; and so on. Thus, in the example above three pairs will be formed by the first and third subjects; by the second and fourth; and the fifth and sixth.'

[1] Hierbij komt nog dat een bijzonder duidelijk boekje van ARMITAGE (1960) verschenen is, waarin niet alleen allerlei medische aspecten bij de toepassing van sequente methoden worden besproken, maar waarin tevens een reeks sequente schema's voor praktisch gebruik geheel is uitgewerkt en van voorbeelden is voorzien.

Wanneer bij een tweetal behandelingen A het gewenste resultaat geeft en B niet (of wanneer A een beter resultaat geeft dan B), spreken wij van de uitkomst *plus*. Het omgekeerde geval levert de uitkomst *min* op. Als beide patiënten op dezelfde wijze reageren (beide gunstig of beide ongunstig) zegt dat niets over het verschil tussen de genezingskansen van de twee behandelingen, zodat elk tweetal met dit resultaat bij de analyse buiten beschouwing blijft.

Onder de hypothese dat beide behandelingen even grote kansen op een gunstig resultaat bieden is $P(+) = P = \frac{1}{2}\,[= P(-)]$. Indien A een gunstiger resultaat oplevert is $P > \frac{1}{2}$, indien B een gunstiger resultaat geeft is $P < \frac{1}{2}$. Wij veronderstellen nu dat men de hypothese H_0: $P = \frac{1}{2}$ wil toetsen met als alternatieve hypothese $H_1 : P \geqslant P_1\ (> \frac{1}{2})$ of $P \leqslant 1 - P_1$ (kiest men dus bijv. $P_1 = 0{,}75$, dan is $1 - P_1 = 0{,}25$). Verder nemen wij aan dat de kans op het ten onrechte verwerpen van H_0 ten hoogste α en het onderscheidingsvermogen van de toets ten minste $1 - \beta$ moet bedragen.

Voor $P_1 = 0{,}55(0{,}05)0{,}95$, $\alpha = 0{,}05$ of $0{,}01$ en $1 - \beta = 0{,}95$ geeft ARMITAGE (1960) de open (niet afgeknotte) sequente toetsingsschema's resp. in zijn tabellen 3.1 en 3.2.[1] Wij laten deze schema's buiten beschouwing daar de afgeknotte schema's voor de praktijk belangrijker zijn. Deze zijn voor dezelfde waarden van P_1, α en $1 - \beta$ in zijn tabellen 3.5 en 3.7, resp. 3.6 en 3.8 opgenomen. Bij de grafische uitvoering van de sequente toetsing gaat men uit van een grafiek, waarin op de X-as het aantal paren met preferentie wordt afgezet, terwijl de Y-as dient voor het aangeven van de 'score' y (aantal plussen verminderd met het aantal minnen) die bij elk opeenvolgend paar wordt bereikt (zie figuur 6.9). Uit de tabellen kan men voor de gekozen waarden van P_1 (en $1 - P_1$), α en $1 - \beta$ de coëfficiënten a en b aflezen van de vergelijkingen van de grenslijnen H en L, die luiden

$$y_H = a + bn \tag{6.26}$$

en

$$y_L = -a - bn, \tag{6.27}$$

en verder het aantal waarnemingen n_m (bij ARMITAGE met N aangeduid) waarbij de toets kan worden afgeknot.

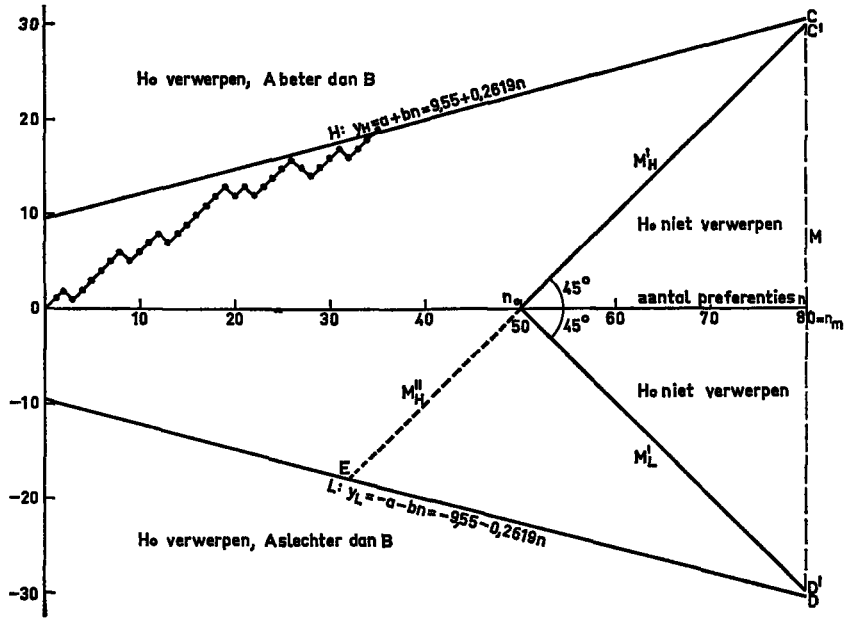
Tabel 3.5 geeft voor $\alpha = 0{,}05$ en $1 - \beta = 0{,}95$ de waarden van a en b voor $P_1 = 0{,}55(0{,}05)0{,}70$ en de waarden van n_m voor $P_1 = 0{,}55(0{,}05)0{,}95$. Tabel 3.7 geeft voor dezelfde α en $1 - \beta$ en $P_1 = 0{,}75(0{,}05)0{,}95$ exact berekende afzonderlijke *punten* voor de grenslijnen H en L.

Tabel 3.6 geeft voor $\alpha = 0{,}01$ en $1 - \beta = 0{,}95$ de waarden van a en b voor

[1] In deze en de volgende tabellen gebruikt ARMITAGE 2α, waar wij met α werken en θ_1 in plaats van P_1.

$P_1 = 0{,}55(0{,}05)0{,}75$ en de waarden van n_m voor $P_1 = 0{,}55(0{,}05)0{,}95$. Tabel 3.8 geeft voor deze α en $1 - \beta$ en $P_1 = 0{,}80(0{,}05)0{,}95$ exact berekende afzonderlijke punten voor de grenslijnen.

Voor de in de tabellen 3.7 en 3.8 opgenomen gevallen kunnen de waarden van a en b voor de *benaderende grenslijnen* desgewenst in de tabellen 3.1, resp. 3.2 onder a_1 en b worden afgelezen.



Figuur 6.9.

Kiest men bij het voorbeeld (de vergelijking van de behandelingen A en B) $P_1 = 0{,}75$, $\alpha = 0{,}01$ en $1 - \beta = 0{,}95$, dan volgen uit tabel 3.6 de waarden: $a = 9{,}55$, $b = 0{,}2619$ en $n_m = 80$. De vergelijkingen van de twee grenslijnen H en L luiden dus

$$y_H = 9{,}55 + 0{,}2619n, \qquad y_L = -9{,}55 - 0{,}2619n.$$

Uit deze vergelijkingen volgt dat voor $n = 0$ geldt dat $y_H = 9{,}55$ en $y_L = -9{,}55$, terwijl voor $n = n_m = 80$ de waarden $y_H = 9{,}55 + 0{,}2619(80) = 30{,}50$ en $y_L = -30{,}50$ verkregen worden. Met deze getallen kunnen de grenslijnen H en L worden ingetekend (zie figuur 6.9). Trekt men door het punt $n = n_m = 80$ op de X-as de verticaal M dan snijdt deze de lijn H in het punt C met $y_H = 30{,}5$ en de lijn L in het punt D met $y_L = -30{,}5$.

Men zou voor de uitvoering van de sequente analyse kunnen uitgaan van de grafiek — met de lijnen H, L en M — die nu is verkregen. Te beginnen met het eerste tweetal gaat men dan het steekproefpad in de grafiek tekenen, door voor elke *plus* (preferentie voor A)

één eenheid naar rechts (volgens de schaal op de X-as) en één eenheid naar boven (volgens de schaal op de Y-as) te gaan. Voor elke *min* (preferentie voor B) gaat men één eenheid naar rechts en één eenheid naar beneden. Zolang het steekproefpad binnen de grenslijnen H, L en M blijft wordt de proef voortgezet. Verder geldt bij overschrijding van de grenslijn:

H — H_0 wordt verworpen t.g.v. $H_1 : P_1 \geqslant 0{,}75$ (A beter dan B),
L — H_0 wordt verworpen t.g.v. $H_1 : P_1 \leqslant 0{,}25$ (A slechter dan B),
M — H_0 wordt niet verworpen.

Het schema kan echter nog worden verbeterd, door de grenslijn M te vervangen door de 'wig', gevormd door de grenslijnen M_H' en M_L'. Deze worden zo getekend dat een steekproefpad, dat M_H' (M_L') bereikt zelfs bij het doorlopen tot het maximum aantal tweetallen met uitsluitend plussen (minnen) de grenslijn $H(L)$ niet meer kan bereiken. Dit betekent dat men bij kruising van het steekproefpad met M_H' of M_L' concludeert, dat H_0 niet wordt verworpen. Daartoe dient de lijn M_H' te lopen van het punt C' op de lijn M (onder C) tot het punt n_0 op de X-as. Punt C' heeft als waarde van y het grootste gehele getal (y_0) dat kleiner is dan de y-waarde van punt C, d.i. het hoogste punt beneden de lijn H dat het steekproefpad bij $n_m = 80$ preferenties kan bereiken. Voor het voorbeeld is $y_0 = 30$ (het grootste gehele getal beneden 30,5). Verder geldt dat $n_0 = n_m - y_0 = 80 - 30 = 50$. De lijn M_L' begint bij het punt D' op de lijn M (boven D), waarvan de y-waarde – 30 is en loopt eveneens naar $n_0 = 50$. Indien men (zoals in figuur 6.9) de eenheden van de schalen op de X- en Y-as gelijk neemt maken de lijnen M_H' en M_L' hoeken van 45° met de X-as.

In figuur 6.9 is het steekproefpad uitgezet dat correspondeert met de volgende reeks preferenties:

+	+	−	+	+	+	+	+	−	+	+	+	−	+	+	+	+	+
+	−	+	−	+	+	+	+	−	−	+	+	+	−	+	+	+	

Bij de 35-ste preferentie kruist het pad de grenslijn H. Het onderzoek wordt stopgezet met de conclusie dat behandeling A beter is dan behandeling B.

Men kan ook gebruik maken van *scheve* afgeknotte schema's (skew designs). Deze berusten op het feit dat bijv. de grenslijn H niet meer kan worden bereikt, wanneer het steekproefpad het gebied tussen L en M_L' is ingegaan. Er zijn, indien het verlengde M_H'' van de lijn M_H' is gesneden, nog maar twee conclusies mogelijk, n.l. (*a*) dat B beter is dan A, of (*b*) dat de behandelingen niet significant verschillen. Nu kan het voorkomen dat men het ethisch niet verantwoord acht in dit geval de proef verder voort te zetten (als bijv. B de gebruikelijke en A een nieuwe behandeling is, zal men in ieder geval bij conclusie (*a*),

maar waarschijnlijk ook bij conclusie (*b*), behandeling *B* handhaven). Men kan dan gebruik maken van het scheve afgeknotte schema dat gevormd wordt door de grenslijnen *H*, *C'E* en *L*. Men concludeert bij overschrijding van de grenslijn:

H — *A* beter dan *B*.
L — *A* slechter dan *B*; deze conclusie is slechts mogelijk indien het steekproefpad de lijn *H* tussen de *Y*-as en punt *E* kan snijden.
C'E — *A* niet beter dan *B*.

De sequente analyse kan ook met drie of meer behandelingen (bijv. twee geneesmiddelen en een placebo) worden uitgevoerd. Voor deze methode en uitvoeriger bespreking van de voorgaande methoden, met een groot aantal praktijkvoorbeelden, verwijzen wij naar ARMITAGE (1960).

Tenslotte enkele opmerkingen betreffende de toepassing van de sequente methode op medisch gebied:

a. Door GOSLINGS (1961) is de medisch-ethische zijde van de sequente analyse besproken in een artikel, dat tevens verschillende andere aspecten van een 'controlled clinical trial' belicht. Zijn conclusie luidt dat medisch-ethisch gezien een sequente proefopzet, waarbij in zo kort mogelijke tijd een beslissing met een bekende onbetrouwbaarheid kan worden bereikt, sterk toe te juichen valt.

b. De sequente methode is goed te gebruiken wanneer men het effect van een behandeling bij acute aandoeningen wil onderzoeken, of te maken heeft met een behandeling die snel een verbetering in een chronische aandoening kan opleveren. In gevallen met een korte opnameduur en een lange observatietijd heeft de methode gewoonlijk geen zin. ARMITAGE (1960) bespreekt echter (in hoofdstuk 7) bepaalde gevallen waarin men toch sequent kan werken.

c. De sequente methode is geschikt voor proeven waarbij men de behandeling beoordeelt op één criterium (zoals bijv. het effect op de bloeddruk bij hypotensieve middelen gedurende operaties), of wanneer enkele criteria zich tot één totaaluitkomst laten combineren. Heeft men echter voor het waarderen van het effect van de behandeling een groot aantal verschillende criteria nodig, dan is deze methode (nog) niet goed bruikbaar.

6.4.4. OPGAVEN

6.7. Teken in één grafiek de lijnen L_0 en L_1 die corresponderen met onderstaande gegevens:

a. $P_0 = 0{,}5$, $P_1 = 0{,}9$, $\alpha = \beta = 0{,}05$ — *b.* Als (*a*), maar $\alpha = \beta = 0{,}01$
c. $P_0 = 0{,}5$, $P_1 = 0{,}8$, $\alpha = \beta = 0{,}05$ — *d.* Als (*c*), maar $\alpha = \beta = 0{,}01$

6.8. Bepaal voor de vier situaties in opgave 6.7 bij benadering de *GSG* krommen en vergelijk deze.

6.9. Uit DIXON en MASSEY (1957).

Men toetst de hypothese, dat de fractie defecten in een grote partij van een bepaald artikel $P_0 = 0{,}1$ bedraagt, met als alternatief: $P_1 = 0{,}3$. Men kiest $\alpha = 0{,}05$ en $\beta = 0{,}10$. Een aselecte steekproef levert — in volgorde van trekking — de volgende uitkomsten op (D = defect):

G G D G D G G G D D G G G G G D G G G
G G G D G G G G D G

a. Toets op de gebruikelijke wijze. Tot welke uitspraak komt u?
b. Pas de sequente analyse toe. Bij welk aantal waarnemingen komt u tot een uitspraak en hoe luidt deze?

6.5. Samenvatting

Dit hoofdstuk bevat een betrekkelijk groot aantal 'technische' details betreffende het uitvoeren van een toetsing. Het lijkt ons daarom nuttig de algemene gang van zaken bij het toetsen van een hypothese nog eens in enkele punten samen te vatten en daar enige opmerkingen aan te verbinden. Wij gaan uit van de situatie die in de praktijk het meest voorkomt, nl. die waarin een enkelvoudige hypothese H_0 (zoals $P = 0{,}6$) wordt getoetst met als alternatief een samengestelde hypothese H_1 (van het type $P < 0{,}6$, $P > 0{,}6$ of $P \neq 0{,}6$).

Elke toetsing berust op een aselecte steekproef van n elementen en verloopt als volgt:

1. Stipuleer de *te toetsen hypothese* H_0 en de *tegenhypothese* H_1. Ook als men geen specifieke alternatieve hypothese kan aangeven, dient men a priori te overwegen, welke mogelijke hypothesen men wil toelaten (om te kunnen uitmaken, of één- of tweezijdig wordt getoetst).
2. Kies een *toetsingsgrootheid* u, waarvan de kansverdeling onder H_0 (bij benadering) kan worden berekend.
3. Kies de *onbetrouwbaarheidsdrempel* α, waarbij wordt getoetst. Deze keuze van de maximaal toelaatbaar geachte kans op het ten onrechte verwerpen van H_0 is een zaak voor de onderzoeker en hangt af van de risico's, die een foutieve uitspraak t.a.v. H_0 met zich mee brengt.
4. Bepaal op grond van de kansverdeling van $\underline{u}$ onder H_0 een zodanige *kritieke zone* Z (één- of tweezijdig, afhankelijk van H_1) dat de kans dat $\underline{u}$ een in Z gelegen waarde aanneemt onder H_0 hoogstens gelijk is aan α.
5. *Verwerp* H_0 op grond van de getrokken steekproef, als de daarbij behorende waarde u_0 van $\underline{u}$ in Z ligt. Men zegt dan wel: de waargenomen uitkomst is significant bij een drempelwaarde α (met een onbetrouwbaarheid α). *Verwerp* H_0 *niet*, als u_0 niet in Z ligt.

Deze punten geven aanleiding tot de volgende opmerkingen:

a. Als men, bijv. op grond van de praktische mogelijkheden de steekproefomvang tevoren kiest, dient men feitelijk door middel van de werkingskromme van de toets te onderzoeken, hoe groot voor verschillende alternatieve hypothesen de kans op een fout van de tweede soort, d.i. de kans op het ten onrechte verwerpen van deze hypothesen, is. Door de keuze van α heeft men immers de kans op het ten onrechte

verwerpen van H_0 in de hand. Bij een gegeven steekproefomvang kan echter voor een H_1, die van praktische betekenis is, de kans op een fout van de tweede soort β wel zo groot worden, dat men H_1, ook als deze juist is, slechts zelden als zodanig zal erkennen. Anderzijds zou ook kunnen blijken, dat men een aanvaardbare β reeds bij veel kleinere steekproefomvang kan verkrijgen.

b. Wanneer H_1 tevoren volledig kan worden gespecificeerd, zodat ook deze hypothese enkelvoudig is, kan men in vele gevallen: 1) de noodzakelijke steekproefomvang tevoren vaststellen, of 2) gebruik maken van een sequente methode. Naar onze mening dient men, meer dan momenteel gebruikelijk is, naar deze procedure te streven.

c. Evenmin als bij een intuïtief getrokken conclusie (in de trant van: de steekproef bevat slechts 29,1% mannen, dus het is onwaarschijnlijk dat de populatie er 50% bevat) kan men via een statistische toetsing tot een zonder enige reserve zekere uitspraak komen. De exactheid en objectiviteit die door het toepassen van een toets worden verkregen bestaan daarin, dat men zelf kan bepalen met welke kans(en) op een foutieve uitspraak men genoegen wil nemen.

HOOFDSTUK 7.

SCHATTINGSTHEORIE

Bij vele praktijkproblemen beschikt men niet over (een) hypothese(n) betreffende de waarde van de parameter van de populatieverdeling, die wordt bestudeerd. Men wenst deze waarde dan zo goed mogelijk te *schatten* uit de getrokken steekproef. Het volgende voorbeeld bevat een typerend schattingsprobleem.

Voorbeeld

7.1. Een arts wordt uitgezonden naar Nieuw-Guinea, teneinde het voorkomen van framboesia in een bepaalde streek te bestuderen. Men vermoedt dat in deze streek betrekkelijk veel personen aan manifeste huidafwijkingen tengevolge van framboesia lijden, maar uit de schaarse beschikbare gegevens kan men zich geen oordeel vormen, dat voldoende betrouwbaar is om tot een verantwoorde bestrijding van het probleem te komen.

A bezoekt de streek en besluit zijn oordeel omtrent de fractie framboesialijders in de populatie te baseren op een aselecte steekproef van n personen. De fractie lijders in deze steekproef, $p_0 = x_0/n$, moet nu dienen om de populatiefractie, P, te schatten.[1]

In dit hoofdstuk behandelen wij eerst de grondslagen van de schattingstheorie. Vervolgens gebruiken wij deze voor het geven van een schatting van de parameter P van een dichotomie.

7.1. Grondslagen van de schattingstheorie

7.1.1. TERMEN EN BEGRIPPEN

Een *schatter* is een steekproeffunctie, die gebruikt wordt om de waarde van een onbekende parameter te schatten. De waarde die een schatter voor een bepaalde steekproef aanneemt, heet een *schatting* van die parameter.

Een schatter heet *zuiver*, als het gemiddelde van zijn kansverdeling gelijk is aan de parameter. Een schatter heet *bruikbaar*, indien de spreiding van zijn kansverdeling afneemt, naarmate n toeneemt; vrijwel alle gebruikelijke schatters zijn bruikbaar. Een schatter heet *doeltreffend*, als hij zuiver èn bruikbaar is en *het doeltreffendst*, als hij

[1] Het is duidelijk, dat het probleem in deze vorm beter met de realiteit zal overeenstemmen dan het overeenkomstige toetsingsprobleem in 6.1.

van alle doeltreffende schatters van een parameter de kleinste spreiding bezit.

Voorbeelden

7.2. De steekproeffunctie $\underline{p} = \underline{x}/n$ is een schatter van de parameter P van een dichotomie. De in een aselecte steekproef waargenomen fractie p_0 is een schatting van P.

De kansverdeling van $\underline{p}$ is een binomiale verdeling met gemiddelde $\mu_{\underline{p}} = P$ en spreiding $\sigma_{\underline{p}} = \sqrt{PQ/n}$ (zie 4.3.1), zodat de schatter $\underline{p}$ zuiver en bruikbaar, dus doeltreffend, is. Bij een dichotomie met $P = 0{,}5$, geldt bijv.:

voor $n = 25$ .... $\sigma_{\underline{p}} = 0{,}1$ ..... $P(0{,}30 < \underline{p} < 0{,}70) = 0{,}95$
$n = 100$ .... $\sigma_{\underline{p}} = 0{,}05$ ..... $P(0{,}40 < \underline{p} < 0{,}60) = 0{,}95$
$n = 1000$ .... $\sigma_{\underline{p}} = 0{,}0158$ .... $P(0{,}47 < \underline{p} < 0{,}53) = 0{,}95$
$n = 10000$ .... $\sigma_{\underline{p}} = 0{,}005$ ..... $P(0{,}49 < \underline{p} < 0{,}51) = 0{,}95$.

7.3. De steekproeffunctie $\underline{x}$ is een schatter van de parameter $\mu_{\underline{x}}$.

7.1.2. PUNTSCHATTINGEN

Men noemt een schatting de *aannemelijkste schatting*, indien de kans op de waargenomen uitkomst maximaal wordt bij substitutie van de schatting voor de parameter.

Voorbeeld

7.4. Onderstel dat men uit een dichotomie een steekproef van 10 elementen trekt, waarin $x_0 = 4$, zodat $p_0 = 0{,}4$. De aannemelijkste schatting van de (onbekende) parameter P is dan die waarde van P, waarvoor de kans op de uitkomst $p_0 = 0{,}4$, maximaal is. Bewezen kan worden dat $P(\underline{x} = 4 \mid n = 10)$ maximaal ($= 0{,}2508$) wordt, voor $P = p_0 = 0{,}4$; voor elke andere waarde van P blijkt deze kans kleiner te zijn. Dus p_0 is de aannemelijkste schatting van P.

Als men één waarde als schatting van de parameter wil geven (men spreekt dan van een *puntschatting*) ligt het voor de hand, daarvoor de aannemelijkste schatting te nemen. Als puntschatting van P geeft men dan p_0. Daar bij een bepaalde P de waargenomen p_0 van steekproef tot steekproef varieert en in de regel niet gelijk is aan P heeft deze wijze van schatten het nadeel, dat het merendeel van de uitspraken die men verkrijgt onjuist is. Deze onbevredigende situatie kan worden verbeterd door het geven van een intervalschatting.

7.1.3. INTERVALSCHATTINGEN

Een schatting van P die in de regel juist is, kan worden verkregen door een *schattingsinterval* te geven, waarvan de benedengrens p_1 en de bovengrens p_2 zodanig worden bepaald, dat de kans op een interval dat de (onbekende) parameter niet bevat, (hoogstens) gelijk is aan α. De kans, dat het interval de parameter wel bevat is dus (minstens) gelijk aan $1 - \alpha$. Men noemt $1 - \alpha$ de *betrouwbaarheidscoëfficiënt* en spreekt van een schattingsinterval met een betrouwbaarheid $1 - \alpha$, of van

een $100(1-\alpha)\%$ *betrouwbaarheidsinterval*. Is bijv. $\alpha = 0{,}05$, dan is de betrouwbaarheidscoëfficiënt 0,95 en verkrijgt men een 95% betrouwbaarheidsinterval. Een betrouwbaarheidsinterval is een stochastisch interval, berekend uit de waarnemingen.

Eén der meest algemene methoden om een intervalschatting met een betrouwbaarheidscoëfficiënt $1-\alpha$ te verkrijgen is te ontlenen aan de theorie van het toetsen van hypothesen. In principe kiest men dit schattingsinterval zodanig, dat het alle waarden van de parameter bevat, die bij tweezijdige toetsing met een onbetrouwbaarheid α niet kunnen worden verworpen op grond van de in de steekproef waargenomen waarde van de schatter. In 7.2.1 geven wij aan hoe men daarbij te werk gaat.

Bij het toetsen van de hypothese H_0 berust de keuze van α op de consequenties, die het ten onrechte verwerpen van H_0 met zich mee brengt. Geheel analoog is bij het schatten van een parameter de keuze van α afhankelijk van de consequenties van de fout die men maakt, als het gegeven schattingsinterval de parameter niet bevat. De keuze van de betrouwbaarheidscoëfficiënt $1-\alpha$ is dus afhankelijk van de vraag, hoe belangrijk het wordt geacht dat het interval de parameter wèl bevat. Een grotere betrouwbaarheid brengt bij een bepaalde steekproefomvang met zich mee, dat men met een groter interval (dus met een minder scherpe schatting van de parameter) genoegen moet nemen. Indien men de steekproefomvang vergroot, wordt bij een bepaalde betrouwbaarheidscoëfficiënt het schattingsinterval kleiner.

7.2. Betrouwbaarheidsgrenzen van een fractie

7.2.1. EXACTE GRENZEN

Wij behandelen eerst de procedure, volgens welke men exacte betrouwbaarheidsgrenzen van een fractie, d.i. de parameter P van een dichotomie, kan bepalen. Dit geeft ons tevens de gelegenheid, de in 7.1. ingevoerde begrippen te verduidelijken en de betekenis van een schattingsinterval te demonstreren.

Een populatie bestaat uit een groot aantal zwarte en witte knikkers. De parameter $P(\text{zwart}) = P$ is onbekend. Wij gaan nu een schema opstellen, waarmee men voor $n = 10$ voor elke waargenomen $p_0 = x_0/10$ een 95% betrouwbaarheidsinterval kan bepalen en gaan daarbij uit van de kansverdelingen van $\underline{p} = \underline{x}/10$, die voor een aantal waarden van P in tabel 7.1 zijn gegeven.

In tabel 7.1 bepalen wij voor iedere waarde van P een linker en een rechter kritieke zone, die ieder een onbetrouwbaarheid van ten hoogste $\frac{1}{2}\alpha = \frac{1}{2}(0{,}05) = 0{,}025$ bezitten (zodat zij gezamenlijk een onbetrouwbaarheid van hoogstens 0,05 hebben). Voor $P = 0{,}95$ vindt men

Tabel 7.1. Kansverdelingen van $\underline{p}$ voor $n = 10$ en verschillende waarden van P

P	p											Som der kansen
	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1	
0,95	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,010	0,075	0,315	0,599	1,000
0,90	0,000	0,000	0,000	0,000	0,000	0,002	0,011	0,057	0,194	0,387	0,349	1,000
0,85	0,000	0,000	0,000	0,000	0,001	0,008	0,040	0,130	0,276	0,347	0,197	0,999
0,80	0,000	0,000	0,000	0,001	0,006	0,026	0,088	0,201	0,302	0,268	0,107	0,999
0,75	0,000	0,000	0,000	0,003	0,016	0,058	0,146	0,250	0,282	0,188	0,056	0,999
0,70	0,000	0,000	0,001	0,009	0,037	0,103	0,200	0,267	0,233	0,121	0,028	0,999
0,65	0,000	0,001	0,004	0,021	0,069	0,154	0,238	0,252	0,176	0,072	0,013	1,000
0,60	0,000	0,002	0,011	0,042	0,111	0,201	0,251	0,215	0,121	0,040	0,006	1,000
0,55	0,000	0,004	0,023	0,075	0,160	0,234	0,238	0,166	0,076	0,021	0,003	1,000
0,50	0,001	0,010	0,044	0,117	0,205	0,246	0,205	0,117	0,044	0,010	0,001	1,000
0,45	0,003	0,021	0,076	0,166	0,238	0,234	0,160	0,075	0,023	0,004	0,000	1,000
0,40	0,006	0,040	0,121	0,215	0,251	0,201	0,111	0,042	0,011	0,002	0,000	1,000
0,35	0,013	0,072	0,176	0,252	0,238	0,154	0,069	0,021	0,004	0,001	0,000	1,000
0,30	0,028	0,121	0,233	0,267	0,200	0,103	0,037	0,009	0,001	0,000	0,000	0,999
0,25	0,056	0,188	0,282	0,250	0,146	0,058	0,016	0,003	0,000	0,000	0,000	0,999
0,20	0,107	0,268	0,302	0,201	0,088	0,026	0,006	0,001	0,000	0,000	0,000	0,999
0,15	0,197	0,347	0,276	0,130	0,040	0,008	0,001	0,000	0,000	0,000	0,000	0,999
0,10	0,349	0,387	0,194	0,057	0,011	0,002	0,000	0,000	0,000	0,000	0,000	1,000
0,05	0,599	0,315	0,075	0,010	0,001	0,000	0,000	0,000	0,000	0,000	0,000	1,000

de linker zone $P \leqslant 0{,}7$ (de kans van deze zone is 0,011; de kans van de zone $p \leqslant 0{,}8$ is 0,086 en deze zone is dus te groot). Een rechter zone kan bij $P = 0{,}95$ niet worden gekozen, daar $P(\underline{p} = 1) = 0{,}599$ ver boven 0,025 ligt. Voor $P = 0{,}9$ kan eveneens slechts een linker zone ($p \leqslant 0{,}6$) worden aangegeven, enz. De verkregen zones zijn in tabel 7.1 met verticale streepjes afgescheiden.

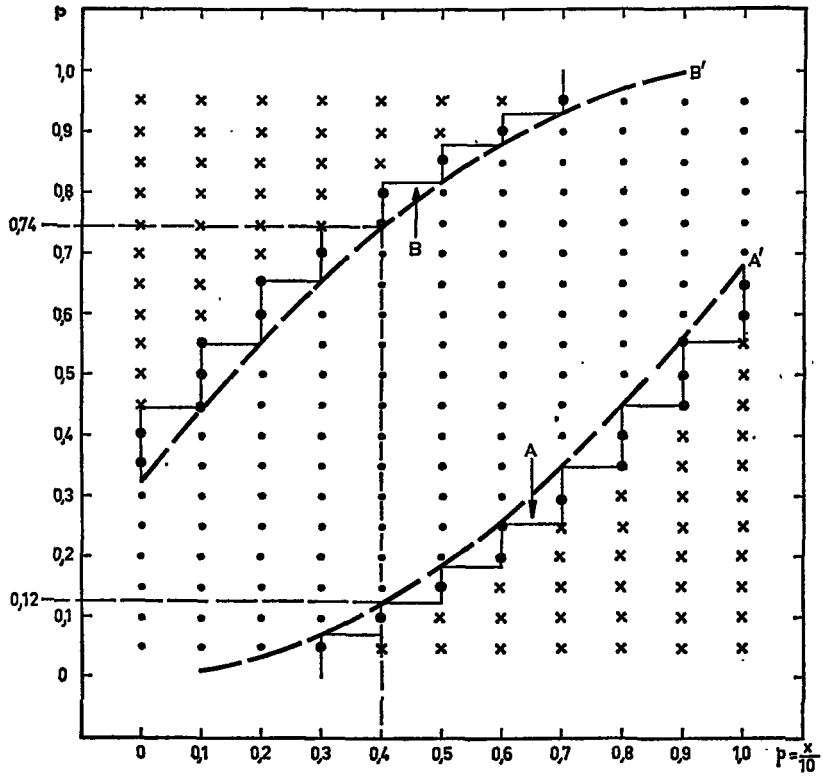
In figuur 7.1 zijn de punten (p, P) uit tabel 7.1 getekend. De waarden (= punten) in de zojuist bepaalde kritieke zones zijn met een kruisje gemerkt, de grenswaarden van deze zones (kritieke waarden) met een dikke stip. De zones zijn vervolgens met de trappenlijnen A en B afgebakend en langs deze trappenlijnen zijn de krommen A′ en B′ geschetst.[1]

Uit de verkregen grafiek kan men nu voor $n = 10$ een intervalschatting met een betrouwbaarheid van ten minste $1 - 2(0{,}025) = 0{,}95$ aflezen. Onderstel bijv. dat een aselect getrokken steekproef $p_0 = 0{,}4$ oplevert. Men richt in het punt $p = 0{,}4$ op de X-as een loodlijn op en leest op de Y-as de ordinaten af van de snijpunten van deze loodlijn met de krommen A′ en B′. Men vindt dan bij benadering

$$0{,}12 < P < 0{,}74.$$

Wij beweren dus, met een betrouwbaarheid van ten minste 0,95, dat de (onbekende) parameter P een waarde tussen 0,12 en 0,74 bezit. De

[1] Om deze trappenlijnen nauwkeurig te tekenen, dient men voor veel meer waarden van P over de kritieke waarden van de linker en rechter zones met per zone $\alpha < 0{,}025$ te beschikken. Deze zijn in tabel 7.1 niet gegeven, maar wèl in figuur 7.1 verwerkt.



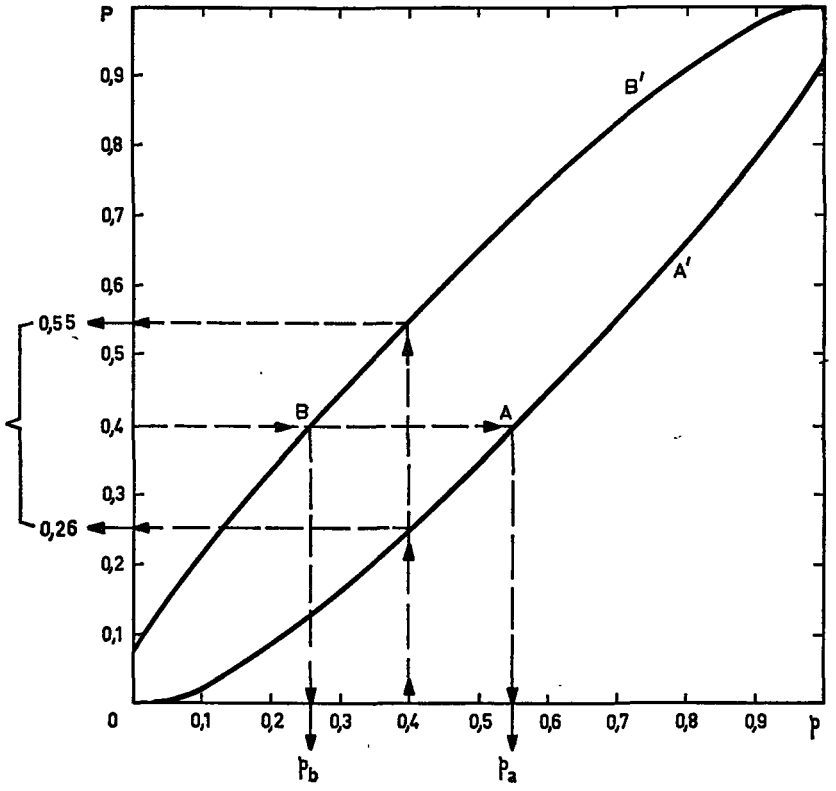
Figuur 7.1. 95% betrouwbaarheidsgrenzen van P voor $n = 10$.

betekenis van deze uitspraak is de volgende: bij aselecte trekking van een steekproef van 10 exemplaren is de kans, dat het op basis van p_0 uit figuur 7.1 bepaalde betrouwbaarheidsinterval P bevat, tenminste gelijk aan 0,95. D.w.z. hoewel de juistheid van een bepaalde uitspraak (één schattingsinterval) onzeker is, zal de relatieve frequentie van het aantal juiste uitspraken bij een groot aantal steekproeven (ongeveer) 0,95 zijn. Men mag echter *niet* beweren: de kans dat P in het verkregen interval ligt is 0,95. P is immers geen stochastische grootheid, maar een constante.

Op welke grond berust deze uitspraak nu? Wij demonstreren dit met figuur 7.2, waarin de grenslijnen A' en B' zijn getekend voor $n = 50$, eveneens met een betrouwbaarheid van 0,95. Evenals in figuur 7.1 zijn de punten, waardoor deze grenslijnen lopen, kritieke waarden, die zo zijn gekozen dat voor elke waarde van P de kans, dat de schatter $\underline{p}$ een waarde binnen deze grenzen aanneemt, ten minste 0,95 is.

Onderstel nu, dat wij te doen hebben met een populatie waarin $P = 0{,}4$. Uit de figuur is dan onmiddellijk af te lezen, dat de kans, dat

$\hat{p}$ een waarde tussen p_B en p_A inneemt, ten minste 0,95 is. Maar tevens blijkt, dat juist voor deze waarden tussen p_B en p_A (zoals bijv. $p_0 = 0{,}4$) een betrouwbaarheidsinterval verkregen wordt, dat P bevat. Voor $\hat{p} \leqslant p_B$ en voor $\hat{p} \geqslant p_A$ verkrijgt men een interval, waar P niet in ligt. Deze redenering geldt voor iedere waarde van P, zodat de grenslijnen altijd een 95% betrouwbaarheidsinterval opleveren.



Figuur 7.2. 95% betrouwbaarheidsgrenzen van P voor $n = 50$.

Het gebied tussen de grenslijnen noemt men wel een *betrouwbaarheidsgordel*. De gordel in figuur 7.2 kan worden gebruikt om bij $n = 50$ een tweezijdig begrensd interval af te lezen met een betrouwbaarheid van ten minste 0,95, of voor het aflezen van een bovengrens, resp. benedengrens met een betrouwbaarheid van ten minste 0,975.

7.2.2. GRAFIEKEN MET BETROUWBAARHEIDSGRENZEN VAN P

Als n niet te klein is en P niet te dicht bij 0 of 1 ligt kan men betrouwbaarheidsgrenzen van P aflezen uit een grafiek, waarin bij een

bepaalde betrouwbaarheidscoëfficiënt voor verschillende waarden van n de betrouwbaarheidsgordels zijn getekend. Voor waarden van n, die niet in de grafiek zijn opgenomen dient men op het oog te interpoleren tussen de krommen met lagere en hogere n. Dergelijke grafieken zijn in verschillende leerboeken en tabellenverzamelingen opgenomen. In PEARSON & HARTLEY (1954) vindt men zeer duidelijke met een betrouwbaarheid van 0,95, resp. 0,99. EISENHART c.s. (1947) geven grafieken met een betrouwbaarheid van 0,99, 0,95, 0,90 en 0,80. De STATISTISCHE TABELLEN EN NOMOGRAMMEN (1951) bevatten grenzen met een gemiddelde betrouwbaarheid van 0,90 en 0,98.[1]

In figuur 7.3 is een grafiek met 95% betrouwbaarheidsgordels opgenomen. Om ruimte te sparen bevat deze figuur alleen de linkerhelft van de betrouwbaarheidsgordels. Wij demonstreren het gebruik ervan aan de hand van enkele voorbeelden.

Voorbeelden

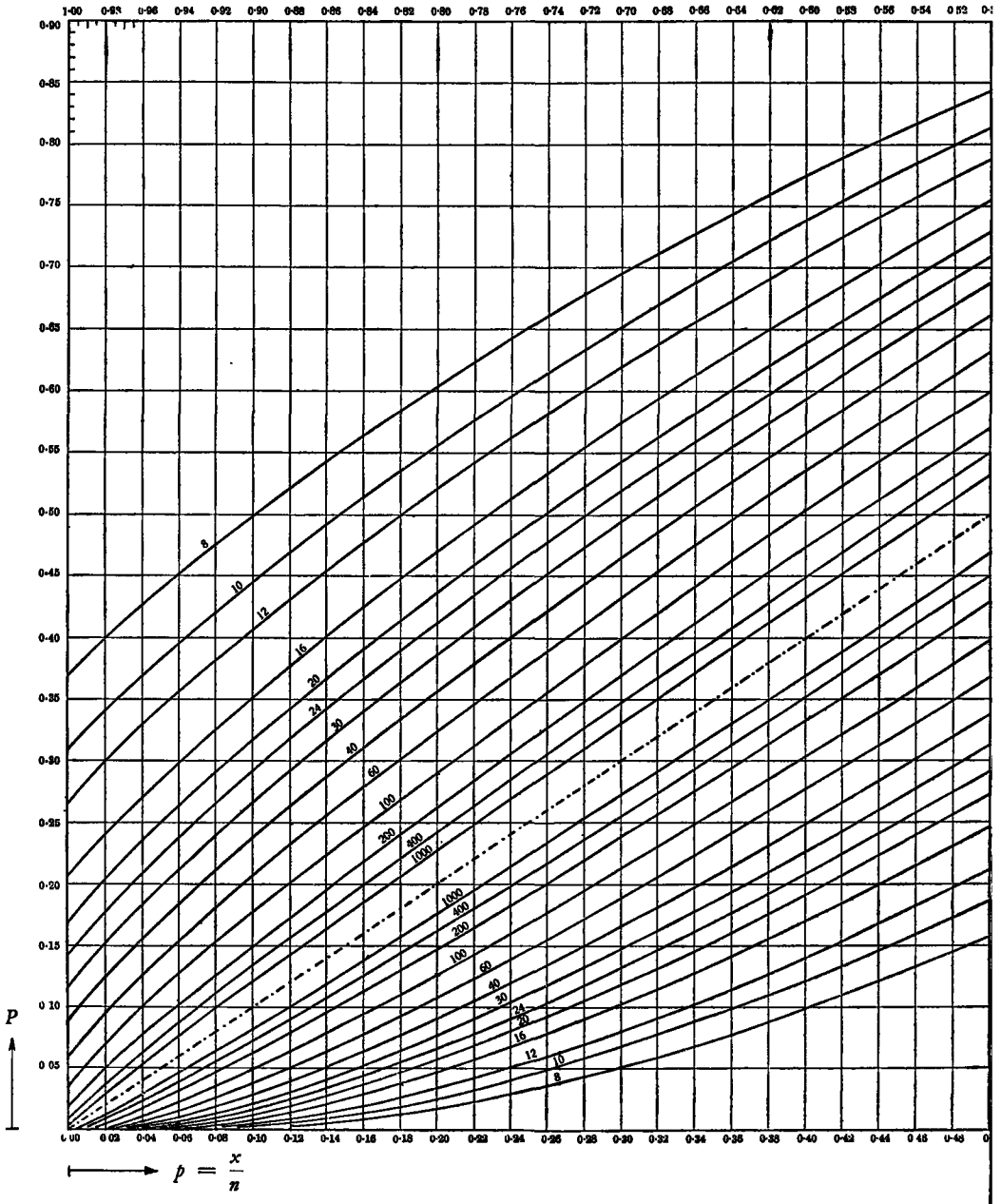
7.5. $p_0 < 0{,}5$: Een aselect getrokken steekproef van 100 elementen levert $p_0 = 0{,}28$ op. Het 95% betrouwbaarheidsinterval van P wordt als volgt bepaald. Men richt in het punt $p = 0{,}28$ op de X-as een loodlijn op en bepaalt van de snijpunten van deze loodlijn met de grenslijnen voor $n = 100$ de ordinaten op de *linker* verticale schaal: 0,19 en 0,38. Het 95% betrouwbaarheidsinterval is dan bij benadering $0{,}19 < P < 0{,}38$.

7.6. $p_0 > 0{,}5$: Onderstel dat $n = 50$ en $p_0 = 0{,}7$. Men richt in het punt $p = 1 - 0{,}7 = 0{,}3$ op de X-as een loodlijn op, bepaalt de snijpunten met de grenslijnen voor $n = 50$ (deze zijn niet getekend, maar moeten tussen die voor $n = 40$ en $n = 60$ worden geïnterpoleerd, iets naar $n = 60$ toe). De ordinaten van deze snijpunten leest men nu af op de *rechter* verticale schaal: (ongeveer) 0,82 en 0,56. Het 95% betrouwbaarheidsinterval is dus bij benadering $0{,}56 < P < 0{,}82$.

7.2.3. TABELLEN MET BETROUWBAARHEIDSGRENZEN VAN P

Als n klein is en P dicht bij 0 of 1 ligt, zodat x_0 klein is, kan men de betrouwbaarheidsgrenzen beter in een tabel opzoeken. Er zijn verschillende tabellen, waarin voor een reeks waarden van x en n (soms voor x en $n - x$) 95% en 99% betrouwbaarheidsgrenzen van P worden gegeven. Het uitvoerigst zijn de tabellen van MAINLAND (1952, 1956) en GEIGY (1960); laatstgenoemde tabellenverzameling geeft 95% en 99% betrouwbaarheidsgrenzen van $100P$ (het *percentage* elementen met kenmerk A in de populatie) voor $n = 2(1)100$, 105, 110(10)200, 250, 300(100)1000. Wij hebben hier drie tabellen met betrouwbaarheidsgrenzen aan ontleend:

[1] Deze zijn gebaseerd op tabellen van HAMAKER en geven betrouwbaarheidsintervallen, die zo zijn samengesteld dat de fractie juiste uitspraken gemiddeld gelijk is aan de betrouwbaarheid. Zij zijn daardoor iets scherper dan de intervallen, die op de gebruikelijke (door ons beschreven) wijze zijn bepaald en waarbij de fractie juiste uitspraken ten minste gelijk is aan de onbetrouwbaarheid. Een voordeel van de HAMAKER-grafieken is, dat zij ook voor kleine fracties kunnen worden gebruikt.



Figuur 7.3. 95% betrouwbaarheidsgrenzen van P voor verschillende waarden van n [uit PEARSON & HARTLEY (1954)].

1. Tabel R – 1 geeft de 95% en 99% betrouwbaarheidsgrenzen van het percentage elementen met kenmerk A in de populatie op basis van x elementen met dit kenmerk in een aselecte steekproef van n elementen, voor $n = 5, 6, 7, 8, 9, 10, 12, 14, 16, 20, 25, 30, 40, 50$ en 75. Bij even waarden van n vindt men deze grenzen voor waarden van x tot en met $\frac{1}{2}n$, bij oneven waarden van n voor waarden van x tot en met $\frac{1}{2}(n + 1)$. Voor grotere waarden van x kan men de gewenste betrouwbaarheidsgrenzen verkrijgen door de bij $n - x$ behorende grenzen af te lezen en deze van 100 af te trekken. Tabel R – 1 bevat slechts een beperkt aantal waarden van n en is dan ook bedoeld om de lezer een indruk te geven van de uitspraken omtrent P die bij steekproeven van betrekkelijk kleine omvang mogelijk zijn.

Voorbeelden

7.7. Voor $n = 10$ en $x_0 = 4$ ($p_0 = 0{,}4$) leest men in de 95% kolom van tabel R – 1 de getallen 12,2 en 73,8 af. Het 95% betrouwbaarheidsinterval van P is dus bij benadering $0{,}12 < P < 0{,}74$.

Voor $n = 10$ en $x_0 = 6$ ($= 10 - 4$) zijn de procentuele grenswaarden $100 - 12{,}2 = 87{,}8$ en $100 - 73{,}8 = 26{,}2$. In dit geval is het 95% betrouwbaarheidsinterval van P bij benadering $0{,}26 < P < 0{,}88$.

7.8. Men wil het 99% betrouwbaarheidsinterval van P bepalen voor $n = 75$ en $x_0 = 60$ ($p_0 = 0{,}8$). Tabel R – 1 levert voor $n = 75$ en $x_0 = 75 - 60 = 15$ in de 99% kolom de grenswaarden 9,7 en 34,3. Daar $100 - 9{,}7 = 90{,}3$ en $100 - 34{,}3 = 65{,}7$ is het 99% betrouwbaarheidsinterval van P bij benadering $0{,}66 < P < 0{,}90$.

2. Tabel R – 2 geeft 95% en 99% betrouwbaarheidsgrenzen van het populatiepercentage ($100P$) voor een reeks waarden van x tussen 0 en 50 en steekproeven van de omvang $n = 100, 150, 200, 300, 500$ en 1000.

3. Tabel R – 3 geeft 95% en 99% betrouwbaarheidsgrenzen van het populatiepercentage voor een waargenomen *percentage* $a = 0(1)50$ en steekproeven van de omvang $n = 100, 200, 500$ en 1000. De grenzen voor $a = 51(1)100$ kunnen worden verkregen door de grenzen van $n - a$ af te lezen en deze van 100 af te trekken.

De tabellen R – 2 en R – 3 zijn speciaal vervaardigd ten behoeve van schattingen bij de differentiatie van cellen in een uitstrijkpreparaat. Dit onderwerp wordt in hoofdstuk 22 uitvoeriger behandeld.

Voorbeelden

7.9. Een aselecte steekproef van 200 kinderen uit een populatie van grote omvang bevat 15 kinderen met folliculosis. Tabel R – 2 levert de volgende schattingsintervallen voor de fractie folliculosis in de populatie:

95% interval: $0{,}043 < P < 0{,}121$,
99% interval: $0{,}034 < P < 0{,}136$.

7.10. Onder 500 erythrocyten worden 4 reticulocyten gevonden. Het 95% betrouwbaarheidsinterval van het percentage reticulocyten in het bloed kan volgens tabel R – 2 op 0,22 — 2,0 pct worden gesteld.

7.11. Een aselecte steekproef van 1000 personen uit een populatie van zeer grote omvang bevat 480 vrouwen. Het 95% betrouwbaarheidsinterval van de fractie vrouwen in de populatie is volgens tabel R – 3: $0{,}449 < P < 0{,}511$ (vergelijk deze uitkomst met de aflezing uit figuur 7.3).

7.2.4. BEREKENING VAN BETROUWBAARHEIDSGRENZEN VAN P

Indien P niet te dicht bij 0 of 1 ligt en n niet te klein is, kan een *benedengrens* p_1 met een onbetrouwbaarheid $\frac{1}{2}\alpha$ bij benadering worden berekend met de formule

$$p_1 = \frac{x_0 + \frac{T_{\frac{1}{2}\alpha}^2}{2} - \sqrt{T_{\frac{1}{2}\alpha}^2\left[\frac{x_0(n-x_0)}{n} + \frac{T_{\frac{1}{2}\alpha}^2}{4}\right]}}{n + T_{\frac{1}{2}\alpha}^2}. \tag{7.1}$$

Een *bovengrens* p_2 met dezelfde onbetrouwbaarheid volgt uit

$$p_2 = \frac{x_0 + \frac{T^2_{1-\frac{1}{2}\alpha}}{2} + \sqrt{T^2_{1-\frac{1}{2}\alpha}\left[\frac{x_0(n-x_0)}{n} + \frac{T^2_{1-\frac{1}{2}\alpha}}{4}\right]}}{n + T^2_{1-\frac{1}{2}\alpha}}. \tag{7.2}$$

De te gebruiken waarden van $T_{\frac{1}{2}\alpha}^2$ volgen uit tabel B. Gemakshalve zijn zij voor de meest voorkomende waarden van α in tabel 7.2 vermeld.

Tabel 7.2. Waarden van $T_{\frac{1}{2}\alpha}^2 = T^2_{1-\frac{1}{2}\alpha}$

$\frac{1}{2}\alpha$	$T_{\frac{1}{2}\alpha}$	$T_{\frac{1}{2}\alpha}^2 = T^2_{1-\frac{1}{2}\alpha}$
0,05	—1,65	2,71
0,025	—1,96	3,84
0,01	—2,33	5,43
0,005	—2,58	6,63

Een tweezijdig begrensd betrouwbaarheidsinterval met onbetrouwbaarheid α (betrouwbaarheid $1 - \alpha$) is te verkrijgen, door (7.1) en (7.2) tegelijk te gebruiken.

Voor grote n en grote x_0 kunnen de termen $T_{\frac{1}{2}\alpha}^2/2$ en $T_{\frac{1}{2}\alpha}^2/4$ in de teller en $T_{\frac{1}{2}\alpha}^2$ in de noemer van (7.1), resp. de hiermede overeenkomende termen met $T^2_{1-\frac{1}{2}\alpha}$ in (7.2), zonder bezwaar worden verwaarloosd, zodat de formules veel eenvoudiger worden en overgaan in

$$p_1 = \frac{x_0 + T_{\frac{1}{2}\alpha}\sqrt{\frac{x_0(n-x_0)}{n}}}{n} = p_0 + T_{\frac{1}{2}\alpha}\sqrt{\frac{p_0 q_0}{n}}, \tag{7.3}$$

$$p_2 = \frac{x_0 + T_{1-\frac{1}{2}\alpha}\sqrt{\frac{x_0(n-x_0)}{n}}}{n} = p_0 + T_{1-\frac{1}{2}\alpha}\sqrt{\frac{p_0 q_0}{n}}. \tag{7.4}$$

Voorbeeld

7.12. Een aselecte steekproef van 225 elementen uit een dichotomie levert $x_0 = 135$ ($p_0 = 0{,}6$) op. Gevraagd wordt het 95% betrouwbaarheidsinterval van P.

Nu is: $1 - \alpha = 0{,}95$, dus $\alpha = 0{,}05$ en $\frac{1}{2}\alpha = 0{,}025$, zodat $T_{\frac{1}{2}\alpha}^2 = T^2_{1-\frac{1}{2}\alpha} = 3{,}84$. Substitutie in (7.1), resp. (7.2), levert op:

$$p_1 = \frac{135 + \frac{3{,}84}{2} - \sqrt{3{,}84\left[\frac{(135)\,(90)}{225} + \frac{3{,}84}{4}\right]}}{225 + 3{,}84} = \frac{136{,}92 - 14{,}52}{228{,}84} = 0{,}535,$$

$$p_2 = \frac{136{,}92 + 14{,}52}{228{,}84} = 0{,}662.$$

Het gevraagde 95% betrouwbaarheidsinterval is $0{,}535 < P < 0{,}662$. Bij toepassing van (7.3) en (7.4) vindt men

$$p_1 = 0{,}6 - 1{,}96\sqrt{\frac{0{,}6 \times 0{,}4}{225}} = 0{,}536 \text{ en}$$

$$p_2 = 0{,}6 + 1{,}96\sqrt{\frac{0{,}6 \times 0{,}4}{225}} = 0{,}664.$$

Het verschil met de nauwkeuriger berekende uitkomsten is dus gering.

7.2.5. BEREKENING VAN DE STEEKPROEFOMVANG, NODIG VOOR HET VERKRIJGEN VAN EEN BETROUWBAARHEIDSINTERVAL VAN GEGEVEN BREEDTE

Onderstel dat men een steekproef wil trekken uit een dichotomie en vermoedt dat P in de buurt van p' ligt. Hoe groot moet deze steekproef worden gekozen, indien men een betrouwbaarheidsinterval met betrouwbaarheid $1 - \alpha$ wil verkrijgen waarvan de breedte ongeveer gelijk is aan i?

Uit (7.3) en (7.4) volgt dat bij benadering

$$p_1 = p' + T_{\frac{1}{2}\alpha}\sqrt{\frac{p'q'}{n}},\quad p_2 = p' + T_{1-\frac{1}{2}\alpha}\sqrt{\frac{p'q'}{n}} \quad (q' = 1 - p'),$$

zodat

$$i = p_2 - p_1 = T_{1-\frac{1}{2}\alpha}\sqrt{\frac{p'q'}{n}} - T_{\frac{1}{2}\alpha}\sqrt{\frac{p'q'}{n}} = -2T_{\frac{1}{2}\alpha}\sqrt{\frac{p'q'}{n}}.$$

Hieruit volgt

$$n = \frac{4T_{\frac{1}{2}\alpha}^2 p'q'}{i^2}. \tag{7.5}$$

Voorbeeld

7.13. Onderstel dat men vermoedt dat P in de buurt van 0,7 ligt en dat men een 99% betrouwbaarheidsinterval wenst te verkrijgen, dat ongeveer gelijk is aan 0,05. Dan is: $1 - \alpha = 0{,}99$, $\alpha = 0{,}01$, $\frac{1}{2}\alpha = 0{,}005$, zodat $T_{\frac{1}{2}\alpha^2} = 6{,}63$ (tabel 7.2). Verder is $p' = 0{,}7$, $q' = 0{,}3$ en $i = 0{,}05$, zodat uit (7.5) volgt

$$n = \frac{4(6{,}63)\ (0{,}7)\ (0{,}3)}{(0{,}05)^2} = 2228.$$

Deze steekproefomvang is zo groot, dat men besluit met een 95% betrouwbaarheidsinterval genoegen te nemen. Als de overige gegevens ongewijzigd blijven, vindt men

$$n = \frac{4(3{,}84)\ (0{,}7)\ (0{,}3)}{(0{,}05)^2} = 1291.$$

Men trekt een steekproef van deze omvang en deze levert $p_0 = 0{,}65$ op, zodat het 95% betrouwbaarheidsinterval, berekend met (7.3) en (7.4), bij benadering $0{,}65 \pm 0{,}0260$ is. De breedte van dit interval is dus, ondanks het verschil tussen p_0 en p', vrijwel gelijk aan de gestipuleerde breedte, 0,05.

7.2.6. OPGAVEN

7.1. *a.* Hoe kan men een kleiner betrouwbaarheidsinterval met een grotere betrouwbaarheidscoëfficiënt verkrijgen?

b. Wat gebeurt er met de betrouwbaarheidscoëfficiënt, indien men bij een bepaalde steekproefomvang het betrouwbaarheidsinterval kleiner maakt?

7.2. Betreft voorbeeld 7.1. A trekt aselect een steekproef van 250 personen. Hierin bevinden zich 105 framboesialijders. Geef het 95% betrouwbaarheidsinterval van P, de fractie framboesialijders in de populatie:

a. Uit figuur 7.3 *b.* Met de formules (7.1) en (7.2)

c. Met de formules (7.3) en (7.4).

7.3. Betreft voorbeeld 7.1. Onderstel dat A een steekproef van zodanige omvang wil trekken dat het 95% betrouwbaarheidsinterval, berekend op basis van deze steekproef, niet groter is dan 0,10. Hoe groot dient hij deze steekproef (ongeveer) te kiezen als hij vermoedt dat P in de buurt ligt van:

a. 0,2 *b.* 0,4 *c.* 0,5 *d.* 0,55 *e.* 0,6.

7.4. Uit een populatie van patiënten met een bepaalde ziekte trekt men aselect een steekproef van 50 personen. Hieronder bevinden zich 10 mannen. Bereken de bovengrens van de fractie mannen in de populatie met $\alpha = 0{,}05$:

a. Met formule (7.2) *b.* Met formule (7.4).

7.5. Bepaal uit onderstaande steekproefuitkomsten bij benadering de 95% en 99% betrouwbaarheidsintervallen van P:

Steekproef	a	b	c	d	e	f	g	h
n	400	900	1600	2500	225	625	900	3600
p_0	0,4	0,6	0,8	0,2	0,5	0,5	0,9	0,9

7.6. Men wil een steekproef trekken uit een dichotomie en vermoedt dat P ongeveer gelijk is aan 0,8. Hoe groot moet deze steekproef zijn, opdat een 95% betrouwbaarheidsinterval verkregen wordt dat niet groter is dan:

a. 0,10 *b.* 0,05 *c.* 0,02 *d.* 0,01.

7.7. Bepaal het 95% en 99% betrouwbaarheidsinterval van P, indien men $p_0 = 0{,}2$ heeft waargenomen in een steekproef van de omvang:
a. 20 *b.* 40 *c.* 50 *d.* 75 *e.* 100 *f.* 200.

7.8. Bij de differentiatie van 100 witte bloedcellen werden 11 basophiele granulocyten aangetroffen. Tussen welke grenzen ligt, behoudens een onbetrouwbaarheid van 5%, het percentage basophielen in het bloed?

7.3. Schatten en toetsen

Uit formule (7.5) blijkt, dat bij een schattingsprobleem de steekproefomvang afhankelijk is van drie factoren:

a. De breedte van het gewenste schattingsinterval.

b. De betrouwbaarheidscoëfficiënt.

c. De waarde van P. Bij vele praktijkproblemen is hier de grootteorde wel ongeveer van bekend, bijv. uit de literatuur, uit eigen voorgaande onderzoeken, uit een vooronderzoek, e.d.

Bij het bepalen van een betrouwbaarheidsinterval komen in principe dezelfde kwesties aan de orde als bij het toetsen van een hypothese. Men kan ook twee soorten fouten onderscheiden:

1. Het interval bevat de waarde van de parameter niet. Deze fout is analoog aan een fout van de eerste soort bij het toetsen van hypothesen. De kans op het vermijden van deze fout heeft men in de hand door de keuze van de betrouwbaarheidscoëfficiënt.

2. Het interval bevat te veel waarden die onjuist zijn, d.w.z. het is te breed, waardoor de schatting onscherp wordt. Deze fout is analoog aan de fout van de tweede soort bij het toetsen van hypothesen. WALLIS & ROBERTS (1956) karakteriseren deze fout zeer duidelijk als volgt: 'It is of little help to know that a needle you are seeking is in a certain haystack, however precisely you may be told the location of the haystack. To know which cubic inch the needle is in, however, may be all that is needed for the practical purpose of finding it'.

Ook bij het schatten zal men een compromis moeten zoeken en wel tussen: *a.* Een klein interval, dat veel informatie geeft, maar onbetrouwbaar is ($1 - \alpha$ laag zodat kans dat het interval de parameter bevat klein is), en *b.* Een groot interval, dat weinig pertinente informatie geeft, maar zeer betrouwbaar is ($1 - \alpha$ hoog zodat de kans dat het interval de parameter bevat groot is). Evenals bij het toetsen kan men de beste keuze maken door a priori te overwegen hoe nauwkeurig men wenst te schatten (keuze van de breedte van het interval) en hoe betrouwbaar men wil schatten (keuze van de betrouwbaarheidscoëfficiënt). Men kan dan — eventueel voor een aantal waarden van de parameter, die men mogelijk acht — met (7.5) ongeveer de noodzakelijke steekproefomvang bepalen. Deze handelwijze voorkomt dat men

een te grote steekproef trekt (met als uitkomst een schatting die veel nauwkeuriger is dan nodig is), of dat men een te kleine steekproef trekt (met als uitkomst een schatting die zo onnauwkeurig is dat men er weinig of niets mee kan uitvoeren).

Een betrouwbaarheidsinterval verstrekt meer informatie dan een toetsing. In plaats van een keuze tussen twee tevoren gestelde hypothesen leidt het tot een splitsing van alle mogelijke hypothesen in twee groepen: een groep van hypothesen die (bij een bepaalde betrouwbaarheid) met de steekproefuitkomst te verenigen zijn en een groep van hypothesen waarvoor dit niet het geval is. De eerstgenoemde vallen binnen de betrouwbaarheidsgrenzen. Dit verschil blijkt duidelijk uit het volgende voorbeeld.

Voorbeeld

7.14. Men toetst de hypothese $P = 0{,}5$ met als alternatief $P < 0{,}5$ en kiest $\alpha = 0{,}025$. Een aselecte steekproef van 40 elementen levert $x_0 = 16$ op ($p_0 = 0{,}4$). Uit tabel I blijkt, dat men H_0 bij de gekozen drempelwaarde niet verwerpt, daar de kritieke waarde $x_L = 13$ is.

Uit tabel R – 1 kan bij benadering het 95% betrouwbaarheidsinterval $0{,}25 < P < 0{,}57$ worden afgelezen. Hieruit blijkt eveneens dat H_0 niet kan worden verworpen. Verder blijkt dat zelfs de hypothesen $P = 0{,}56$ of $P = 0{,}26$ niet kunnen worden verworpen. D.w.z. de steekproef geeft, door zijn te kleine omvang, te weinig informatie. Dit onbevredigende feit komt uit de intervalschatting duidelijker naar voren dan uit de toetsing, die slechts de uitspraak 'H_0 niet verwerpen' oplevert. NATRELLA (1960) merkt dan ook in een beschouwing betreffende de relatie tussen schatten en toetsen op:

'The confidence interval not only gives a Yes or No answer, but also, by its width, gives an indication of whether the answer should be whispered or shouted.'

C

METHODEN DER MATHEMATISCHE STATISTIEK

1. VERDELINGSVRIJE METHODEN

'Can we in any of the problems of medical statistics reach satisfactory results by means of relatively simple numerical methods only? Can we satisfactorily test hypotheses and draw deductions from data that have been analysed by means of such simple methods? Personally I believe the answer to these questions is an unqualified yes, that many of the figures included in medical papers can by relatively simple statistical methods be made to yield information of value, and that even the simplest statistical analysis carried out logically and carefully is an aid to clear thinking with regard to the meaning and limitations of the original records.'

A. B. HILL

Tabel 8.1. Systematisch overzicht van de verdelingsvrije toetsen in de hoofdstukken 9 en 10

Proefopzet		Aard van de kenmerken	
		A. Kwalitatief	B. Kwantitatief
		Categorisch systeem met r (2) categorieën *)	Waarden van een grootheid (Rangnummers)
Eén aselecte steekproef met één waarneming per element		9.1. Inleiding 9.2. Toetsing van de hypoth. $P_1=P_2=P_3$ 9.3. χ^2-toets betreffende een populatie met r categorieën	10.1. χ^2-toets voor aanpassing 10.2. Toets van KOLMOGOROV-SMIRNOV
Eén steekproef met herhaalde waarnemingen per element (verwante waarnemingsreeksen)	2	9.4. Tekentoets	10.3. Rang-tekentoets (symmetrietoets van WILCOXON)
	k	9.5. Toets van COCHRAN	10.4. Toets van FRIEDMAN (methode der m rangschikkingen)
Aselecte steekproeven of groepen met één waarneming per element (aselecte waarnemingsreeksen)	2	9.6. Toets van FISHER (2×2 tabel) 9.7. χ^2-toets voor de $2 \times r$ tabel	10.5. Toets van WILCOXON (voor twee steekproeven)
	k	9.8. χ^2-toets voor de $k \times 2$ tabel 9.9. χ^2-toets voor de $k \times r$ tabel	10.6. Toets van KRUSKAL en WALLIS
Eén aselecte steekproef met waarnemingsparen		9.10. χ^2-toetsen voor onafhankelijkheid	10.7. Rangcorrelatietoets van SPEARMAN 10.8. Rangcorrelatietoets van KENDALL

*) Zie voor toetsen betreffende een *geordend* categorisch systeem hoofdstuk 21.

HOOFDSTUK 8

OVERZICHT VAN DE VERDELINGSVRIJE METHODEN

8.1. Mathematisch-statistische methoden

De mathematisch-statistische methoden kunnen worden toegepast wanneer men op basis van aselecte steekproeven uitspraken wil doen omtrent de populaties, waaruit deze steekproeven afkomstig zijn. Zij kunnen worden verdeeld in verdelingsvrije (of parametervrije) en klassieke (of parametrische) methoden.

1. De *verdelingsvrije methoden* danken hun naam aan de eigenschap dat bij hun afleiding geen veronderstellingen behoeven te worden gemaakt over de vorm van de populatieverdelingen (zij zijn dus vrij van veronderstellingen t.a.v. de vorm van deze verdelingen). Het zijn primair methoden voor het toetsen van hypothesen, die wij zullen onderscheiden in:

a. Kwalitatieve verdelingsvrije methoden. Deze kunnen worden gebruikt voor populatieverdelingen van categorische systemen met 2 of meer (r) kenmerken.

b. Kwantitatieve verdelingsvrije methoden. Deze zijn in principe bestemd voor populatieverdelingen van continu verdeelde grootheden. Zij kunnen echter bij benadering worden gebruikt voor discreet verdeelde populaties. Een deel van deze methoden (de z.g. rangorde of ranginvariante methoden) kunnen ook worden toegepast indien de waarnemingen uit rangnummers bestaan.

De verdelingsvrije methoden die wij in de hoofdstukken 9 t/m 12 behandelen worden in de volgende paragraaf ingeleid.

2. De *klassieke methoden* zijn afgeleid onder de veronderstelling dat de populatieverdelingen een *normaal verdeelde* grootheid betreffen. Zij zijn, zowel voor continu als voor discreet verdeelde grootheden met een pseudo-normale verdeling, bij benadering geldig. Bovendien zijn zij voor grotere steekproeven bruikbaar wanneer de populatieverdelingen ééntoppig en niet te sterk asymmetrisch zijn. De klassieke methoden dienen voor het toetsen van hypothesen, resp. het geven van schattingen, betreffende gemiddelden en standaardafwijkingen (of varianties) van populatieverdelingen. Een reeks klassieke methoden wordt in de hoofdstukken 14 t/m 20 behandeld en in hoofdstuk 13 ingeleid.

De kwalitatieve verdelingsvrije methoden kunnen zowel op geordende als op niet-geordende categorische systemen (zie 1.1.2) worden toegepast. Voor *geordende categorische systemen* kan men echter ook — en in vele gevallen zelfs beter — gebruik maken van een klassieke of een kwantitatieve verdelingsvrije methode. Wij bespreken deze mogelijkheden in hoofdstuk 21, dus pas nadat de klassieke methoden zijn behandeld.

Opmerking. Men kan een reeks waargenomen waarden van een grootheid altijd in (kwalitatieve) categorieën omzetten, bijv. door deze waarden te splitsen in de categorieën 'boven de mediaan' en 'ten hoogste gelijk aan de mediaan'. Daarna kan men dan een kwalitatieve verdelingsvrije methode toepassen. Wij laten deze procedure (die tot de z.g. *mediaantoetsen* leidt) verder buiten beschouwing. Zij leidt weliswaar tot eenvoudiger berekeningen, maar tevens tot een ons inziens onverantwoord verlies aan informatie.

Wanneer de waarnemingen in één of meer steekproeven een continu of discreet verdeelde grootheid betreffen, kan zich de vraag voordoen of men een klassieke of een (verdelingsvrije) rangorde methode zal gebruiken. In hoofdstuk 23 zullen wij deze twee soorten methoden met elkaar vergelijken en nagaan, hoe men in de praktijk tot een verantwoorde keuze kan komen.

8.2. Verdelingsvrije methoden

8.2.1. DE VERDELINGSVRIJE METHODEN IN DE HOOFDSTUKKEN 9 EN 10

In tabel 8.1 is een systematisch overzicht gegeven van de verdelingsvrije methoden die in de hoofdstukken 9 en 10 worden behandeld. Deze methoden zijn daartoe volgens twee criteria ingedeeld:

1. *Naar de aard van de waarnemingen* (verticale ingang). De kwalitatieve methoden zijn in hoofdstuk 9, de kwantitatieve in hoofdstuk 10 ondergebracht.

2. *Naar de proefopzet* (horizontale ingang). Uitgaande van het aantal steekproeven en het aantal waarnemingen per element zijn enkele eenvoudige proefopzetten onderscheiden, die in de praktijk herhaaldelijk voorkomen. Deze proefopzetten zijn in tabel 8.1 aan geduid met een korte omschrijving. Wij zullen deze toelichten en van enkele voorbeelden voorzien:

a. Eén aselecte steekproef met één waarneming per element. Men trekt uit de te onderzoeken populatie aselect een steekproef van n elementen en bepaalt van elk element welk kenmerk van een categorisch systeem, resp. welke waarde van een grootheid het bezit.

Voorbeelden

8.1. Men wil onderzoeken of een dobbelsteen zuiver is. Men verricht daartoe 360 worpen en noteert de uitkomsten (deze kunnen worden beschouwd als een aselect met teruglegging getrokken steekproef van 360 elementen). Zie ook voorbeeld 9.1.

8.2. Uit een bepaalde populatie trekt men aselect een steekproef van 200 kinderen. Van elk kind meet men de lichaamslengte (in cm). Men wil op basis van de steekproefuitkomsten de hypothese toetsen dat de lichaamslengte in de populatie normaal verdeeld is.

b. Eén aselecte steekproef met herhaalde waarnemingen per element. Men trekt uit de te onderzoeken populatie aselect een steekproef van n elementen en verricht aan elk element op systematische wijze k waarnemingen betreffende hetzelfde categorische systeem, of dezelfde grootheid. Wij zullen deze proefopzet kortweg met de term *k verwante waarnemingsreeksen* aanduiden [verricht men twee waarnemingen per element ($k = 2$) dan spreekt men van *gepaarde* of *paarsgewijze* waarnemingen (of waarnemingsreeksen)]. In tabelvorm zien deze k verwante waarnemingsreeksen er als volgt uit:

Elementen in de steekproef	Herhaalde waarnemingen					
	1	2	.....	j	.....	k
1	x_{11}	x_{12}	.....	x_{1j}	.....	x_{1k}
2	x_{21}	x_{22}	.....	x_{2j}	.....	x_{2k}
.	.	.		.		.
.	.	.		.		.
.	.	.		.		.
.	.	.		.		.
i	x_{i1}	x_{i2}	.....	x_{ij}	.....	x_{ik}
.	.	.		.		.
.	.	.		.		.
.	.	.		.		.
.	.	.		.		.
n	x_{n1}	x_{n2}	.....	x_{nj}	.....	x_{nk}

De k op één rij staande waarnemingen betreffen steeds hetzelfde element.

Voorbeelden

8.3. Men wil het effect van een geneesmiddel bestuderen. Bij een aselecte steekproef van patiënten uit de te onderzoeken populatie verricht men daartoe bij elke patiënt één waarneming vóór en één waarneming op een zeker tijdstip na de toediening van het middel. Men meet bijv. de systolische bloeddruk van elke patiënt vóór en na de toediening van een middel, waarvan men verwacht dat het bloeddrukverlagend werkt (wil men tevens weten, hoelang het middel effect heeft, dan dient men de waarnemingen op meerdere tijdstippen na toediening te verrichten).

8.4. Men wil de constantheid van een meettechniek onderzoeken. Men verricht dan aan n objecten, die geen verandering ondergaan, metingen op k verschillende tijdstippen.

8.5. Men onderzoekt de gelijkwaardigheid van k onderzoekers bij gebruik van dezelfde meetmethode. Men laat dan elke onderzoeker aan elk van n objecten, onafhankelijk van de andere onderzoekers, één waarneming verrichten.

8.6. Men vergelijkt twee behandelingen met de methode der gepaarde controles. Tegenover ieder individu met behandeling A wordt een individu met behandeling B geplaatst, waarbij ernaar wordt gestreefd dat de twee individuen in een paar in zoveel mogelijk opzichten overeenstemmen (men kiest daartoe bijv. tweeling-dieren van hetzelfde geslacht). Per paar wordt door loting uitgemaakt welk individu behandeling A krijgt.

Opmerking. Een modificatie van proefopzet (*b*) is de opzet waarbij per element elke waarneming uit *twee* of *meer* onderling onafhankelijke metingen (bij 8.4 of 8.5 bijv. duplobepalingen) bestaat. Wanneer men een verdelingsvrije methode wil gebruiken kan men op proefopzet (*b*) terugkomen door de *gemiddelden* van deze metingen te nemen.

c. 1. *Twee of meer aselecte steekproeven met één waarneming per element.* Men trekt uit elke te onderzoeken populatie aselect een steekproef (geeft men deze populaties met de nummers 1, 2, . . ., j, . . ., k aan, dan kan de omvang van deze steekproeven resp. met n_1, n_2, . . ., n_j, . . ., n_k worden aangeduid; voor het totale aantal elementen Σn_j gebruiken wij het symbool n). Aan elk element verricht men één waarneming betreffende een categorisch systeem of een grootheid.

c. 2. *Twee of meer aselect samengestelde groepen met één waarneming per element.* Men splitst een steekproef van n elementen uit de te onderzoeken populatie aselect in k groepen, resp. van de omvang n_1, n_2, . . ., n_k. Deze proefopzet doet zich voor bij experimenten; de groepen ontvangen dan een verschillende behandeling.

Wij geven deze proefopzetten kortweg met de term k *aselecte steekproeven* (resp. *groepen*) aan. In tabelvorm zien zij er als volgt uit:

Steekproeven (groepen)					
1	2	.....	j	.....	k
x_{11}	x_{21}	.....	x_{j1}	.....	x_{k1}
x_{12}	x_{22}	.....	x_{j2}	.....	x_{k2}
.	.		.		.
.	.		.		.
.	.		.		.
.	.		.		.
x_{1n_1}	x_{2n_2}	.....	x_{jn_j}	.....	x_{kn_k}

De verticale lijnen in de tabel geven nadrukkelijk aan, dat men hier met verschillende steekproeven (groepen) te doen heeft.

Bij *c.* 1 en *c.* 2 kunnen dezelfde statistische methoden worden toegepast.

Voorbeelden

8.7. Men wil onderzoeken of de gemiddelde hemoglobinegehalten van kinderen behorend tot vier verschillende populaties gelijk zijn. Uit elke populatie trekt men aselect een steekproef van 100 kinderen en van elk kind wordt het hemoglobinegehalte volgens dezelfde methode bepaald.

8.8. Voor muizen van een bepaalde stam wil men onderzoeken of de gemiddelde narcoseduur bij inspuiting van eenzelfde dosis van twee slaapmiddelen, A en B, hetzelfde is. Men splitst daartoe een groep van 40 muizen van deze stam aselect in twee groepen van 20 muizen en geeft de ene groep middel A en de andere middel B.

d. Eén aselecte steekproef met één waarnemingspaar per element. Men trekt uit de te onderzoeken populatie aselect een steekproef van n elementen. Per element verricht men twee waarnemingen, bijv. één waarneming betreffende de grootheid x en één waarneming betreffende de grootheid y, zodat men per element het *waarnemingspaar* (x, y) verkrijgt.

Voorbeeld

8.9. Voor een populatie van patiënten met een bepaalde aandoening wil men onderzoeken of er een samenhang bestaat tussen de leeftijd en de systolische bloeddruk. Men trekt uit de populatie aselect een steekproef van 50 patiënten, waarvan men per patiënt de leeftijd (x) en de bloeddruk (y) vaststelt.

Vanzelfsprekend zijn vele andere — minder eenvoudige — proefopzetten mogelijk. Enkele hiervan komen in 8.2.3 ter sprake, terwijl er ook een aantal bij de klassieke methoden wordt behandeld.

Het is gewenst dat men de voorgaande proefopzetten zo goed leert kennen, dat men voor gegeven waarnemingsmateriaal direct kan vaststellen of één daarvan is gebruikt. Men kan dan immers, rekening houdend met de aard van de waarnemingen, in tabel 8.1 nagaan welke verdelingsvrije methode kan worden gebruikt. Men kan met deze proefopzetten vertrouwd raken door, voordat men tot lezing van de hoofdstukken 9 en 10 overgaat, de daarin opgenomen voorbeelden en de opgaven te bestuderen.

8.2.2. DE VERDELINGSVRIJE METHODEN IN HOOFDSTUK 11

Hoofdstuk 11 bevat een aantal verdelingsvrije toetsingsmethoden, die kunnen worden toegepast wanneer men de *fluctuatie* in een reeks *opeenvolgende waarnemingen* wil bestuderen. Men kan hiermee de hypothese toetsen dat de waarnemingen in aselecte volgorde staan, met als alternatieve hypothese dat dit niet het geval is, resp. dat een *verloop* (bijv. in de vorm van een geleidelijke stijging) optreedt. Verdere bijzonderheden geeft de inleiding tot deze toetsen in 11.1.

8.2.3. DE VERDELINGSVRIJE TOETSEN IN HOOFDSTUK 12

Bij de proefopzet met k *verwante waarnemingsreeksen* (proefopzet b in 8.2.1) kan het voorkomen dat de waarnemingsreeksen van 1 t/m k in een bepaalde volgorde staan, bijv. wanneer zij betrekking hebben op k opeenvolgende tijdstippen van waarneming. In 12.2 bespreken

wij een kwantitatieve verdelingsvrije toets waarmee men kan onderzoeken of de reeksen in de volgorde 1, 2, . . ., k gemiddeld een stijgend of dalend verloop vertonen. In 12.3 komt de toepassing van deze toets op waarnemingen betreffende een dichotomie aan de orde.

Bij de proefopzet met k *aselecte steekproeven* of *groepen* (proefopzet c in 8.2.1) kan het voorkomen dat de steekproeven of groepen volgens een bepaald criterium in de volgorde 1, 2, . . ., k staan. Betreffen de steekproeven patiënten, dan kan dit criterium bijv. zijn: hun leeftijd, het klinisch stadium waarin zij verkeren, of de tijd die verlopen is tussen het begin van de ziekte en de aanvang van een behandeling. Bij k groepen proefdieren kan de rangschikking van de groepen bijv. berusten op de verschillende hoeveelheden, die van een stof zijn ingespoten. In 12.4 behandelen wij een kwantitatieve verdelingsvrije toets, die kan worden gebruikt om te onderzoeken of er gemiddeld in de volgorde 1, 2, . . ., k een stijgend of dalend verloop aanwezig is. In 12.5 bespreken wij de toepassing van deze toets op waarnemingen die een dichotomie betreffen.

De toetsen in 12.2 en 12.4 kunnen ook worden gebruikt voor waarnemingen die een geordend categorisch systeem van kenmerken betreffen. Deze toepassing wordt in hoofdstuk 21, tezamen met een aantal andere toetsen betreffende zo'n systeem, behandeld.

8.2.4. ANDERE VERDELINGSVRIJE METHODEN

In de hoofdstukken 9 t/m 12 zijn niet alle verdelingsvrije methoden behandeld. In de eerste plaats is gestreefd naar het geven van methoden die voor een bepaald geval een zo groot mogelijk onderscheidingsvermogen bezitten, daarom zijn bijv. de mediaantoetsen (zie de opmerking in 8.1) en de hoektoets voor onafhankelijkheid (zie MATHEMATISCH CENTRUM (1952)) weggelaten. Verder zijn methoden die voornamelijk van theoretisch belang zijn, of methoden die in de praktijk weinig worden gebruikt, eveneens weggelaten. Het is ons uit de reacties op de eerste druk gebleken, dat het gewenst is enkele methoden toe te voegen. Deze zijn om technische redenen in hoofdstuk 26 gegeven.

8.2.5. DE ONTWIKKELING VAN VERDELINGSVRIJE METHODEN

De (kwalitatieve) χ^2-toetsen zijn door K. PEARSON reeds in 1900 aangegeven. Enkele kwantitatieve verdelingsvrije methoden zijn eveneens reeds van oudere datum; zo dateert bijv. de rangcorrelatiemethode van SPEARMAN uit 1904 en wordt de tekentoets in 1925 door FISHER in de eerste druk van zijn 'Statistical Methods' reeds terloops vermeld. De belangrijke ontwikkelingen op dit gebied zijn echter pas

enkele jaren voor de tweede wereldoorlog begonnen, o.m. met het werk van FRIEDMAN (1937), PITMAN (1937), KENDALL (1938) en MOOD (1940). Daarna is het steeds meer in de belangstelling gekomen en nog altijd worden nieuwe methoden ontwikkeld en reeds bestaande gemodificeerd of uitgebreid.

SIEGEL en TUKEY (1960) geven bijv. een verdelingsvrije methode voor het vergelijken van de spreidingen van twee populatieverdelingen op basis van aselecte steekproeven. Een bezwaar van deze methode is echter, dat zij strikt genomen geldt voor verdelingen met gelijke gemiddelden. Het onderscheidingsvermogen van de toets (voor verschillen in spreiding) neemt nl. af, naarmate het verschil tussen de populatiegemiddelden groter wordt.

GLASSER (1962) beschrijft een verdelingsvrije toets die kan worden toegepast, als men beschikt over een aselecte steekproef van n paren *verwisselbare* waarnemingen (zoals bijv. de lengten van paren identieke tweelingen) en wil onderzoeken, of deze (stochastisch) onafhankelijk zijn.

In Nederland heeft vooral de Statistische Afdeling van het Mathematisch Centrum te Amsterdam zich na 1945 intensief met de verdelingsvrije methoden bezig gehouden. Niet alleen heeft men er bestaande methoden van een praktische handleiding en van tabellen met exacte kansen en/of kritieke waarden voorzien, maar tevens heeft men er verschillende nieuwe methoden ontwikkeld. Deze activiteiten zijn in vele rapporten en artikelen, alsmede in een aantal proefschriften tot uitdrukking gebracht.

HOOFDSTUK 9.

KWALITATIEVE VERDELINGSVRIJE TOETSEN

9.1. Populatieverdelingen met r categorieën

Beschouw een populatie, waarvan elk der elementen één van de kenmerken van een categorisch systeem met r categorieën draagt. Stel dat er N_1 elementen zijn met kenmerk A, N_2 met kenmerk B, ..., N_r met kenmerk R en dat $N_1 + N_2 + \ldots + N_r = N$.

9.1.1. STEEKPROEVEN ZONDER TERUGLEGGING

Wanneer men aselect zonder teruglegging een steekproef van n elementen trekt uit de bovengenoemde populatie, is de kans op een steekproef bestaande uit x_1 elementen met kenmerk A, x_2 elementen met kenmerk $B, \ldots, x_r$ elementen met kenmerk R

$$(9.1) \qquad P(f_1, f_2, \ldots, f_r) = \frac{C_{f_1}^{N_1} \times C_{f_2}^{N_2} \times \ldots \times C_{f_r}^{N_r}}{C_n^N}$$

$$(f_r = n - f_1 - f_2 - \ldots - f_{r-1}).$$

Deze kansverdeling noemen wij de *algemene hypergeometrische verdeling.*

De verdeling volgens formule (4.1) is hier een speciaal geval van met $r = 2$.

Voorbeeld

9.1. Men trekt 13 kaarten uit een goed geschud spel van 52 kaarten. De kans dat zich hierin 10 schoppen, 2 harten, 1 ruiten en 0 klaveren bevinden is volgens (9.1)

$$C_{10}^{13} \times C_2^{13} \times C_1^{13} \times C_0^{13} / C_{13}^{52}.$$

9.1.2. STEEKPROEVEN MET TERUGLEGGING

Stelt men voor de bovengenoemde populatie: $P_1 = N_1/N$, $P_2 = N_2/N, \ldots, P_r = N_r/N$ dan is bij trekking met teruglegging van n elementen de kans op een steekproef met x_1 elementen A, x_2 elementen $B, \ldots, x_r$ elementen R

$$(9.2) \qquad P(f_1, f_2, \ldots, f_r) = \frac{n!}{f_1! f_2! \ldots f_r!} P_1^{f_1} \times P_2^{f_2} \times \ldots \times P_r^{f_r}.$$

Deze kansverdeling heet de *multinomiale verdeling*. De binomiale verdeling is hiervan een speciaal geval met $k = 2$. Voor elke f_i geldt dat

$$\mu_{f_i} = nP_i \tag{9.3}$$

en

$$\sigma^2_{f_i} = nP_iQ_i \qquad (Q_i = 1 - P_i). \tag{9.4}$$

Voor $P_1 = P_2 = \ldots = P_r = P$ gaat (9.2) over in

$$P(f_1, f_2, \ldots, f_r) = \frac{n!}{f_1!f_2! \ldots f_r!} P^n. \tag{9.5}$$

Indien de steekproefomvang aanzienlijk kleiner is dan de populatieomvang ($n \ll N$) kan men de kansverdeling volgens (9.1) benaderen door de kansverdeling volgens (9.2).

Voorbeeld

9.2. Wij gaan uit van een populatie van grote omvang van de volgende samenstelling:

Kenmerk	A	B	C
Kans	1/3	1/3	1/3

Men trekt aselect zonder teruglegging een steekproef van $n = 6$ elementen. Welke kansverdeling volgen de verschillende steekproefuitkomsten?

Daar de steekproefomvang klein is ten opzichte van de populatieomvang kan de gevraagde kansverdeling volgens (9.5) worden benaderd door de multinomiale verdeling

$$P(f_1, f_2, f_3) = \frac{6!}{f_1!f_2!f_3!} \left(\frac{1}{3}\right)^6.$$

Zo vindt men bijv. dat de kans op 5 elementen met kenmerk A, 1 met kenmerk B en 0 met kenmerk C ($f_1 = 5$, $f_2 = 1$, $f_3 = 0$) gelijk is aan

$$P(5, 1, 0) = \frac{6!}{5!1!0!} \left(\frac{1}{3}\right)^6 = \frac{6}{729}.$$

In tabel 9.1A zijn alle verschillende uitkomsten bij trekking van 3 elementen zonder teruglegging met de bijbehorende kansen opgenomen.

Tabel 9.1*A*. De multinomiale verdeling

$$P(f_1, f_2, f_3) = \frac{6!}{f_1!\, f_2!\, f_3!}\left(\frac{1}{3}\right)^6$$

A f_1	B f_2	C f_3	$729P(f_1, f_2, f_3)$
6	0	0	1
0	6	0	1
0	0	6	1
5	1	0	6
5	0	1	6
0	5	1	6
0	1	5	6
1	5	0	6
1	0	5	6
4	2	0	15
4	0	2	15
0	4	2	15
0	2	4	15
2	4	0	15
2	0	4	15
4	1	1	30
1	4	1	30
1	1	4	30
3	3	0	20
3	0	3	20
0	3	3	20
3	2	1	60
3	1	2	60
2	3	1	60
2	1	3	60
1	3	2	60
1	2	3	60
2	2	2	90
Totaal			729

9.2. Toetsing van de hypothese $P_1 = P_2 = P_3$

9.2.1. EXACTE TOETSING

Men wil voor een populatie van grote omvang, waarvan de elementen het kenmerk A, B of C dragen, de hypothese H_0: $P_1 = P_2 = P_3 = 1/3$ toetsen met $\alpha = 0{,}10$. Men trekt aselect een steekproef van 6 elementen en deze bevat $f_1 = 5$ elementen met kenmerk A, $f_2 = 0$ elementen met kenmerk B en $f_3 = 1$ element met kenmerk C.

Intuïtief zal men geneigd zijn om tot het verwerpen van H_0 over te gaan, wanneer de in de steekproef waargenomen fracties $p_1 = f_1/n$, $p_2 = f_2/n$ en $p_3 = f_3/n$ 'duidelijk' van 1/3 afwijken, d.i. als de *waargenomen frequenties* f_1, f_2 en f_3 'duidelijk' verschillen van de (onder H_0) *verwachte frequenties* $e_1 = e_2 = e_3 = \frac{1}{3}n = 2$ (e van 'expected').

Voor $n = 6$ zijn de uitkomsten die het sterkst van deze verwachtingen afwijken:

f_1	f_2	f_3
6	0	0
0	6	0
0	0	6

Deze uitkomsten zijn van hetzelfde *type*, dat wij met (6, 0, 0) zullen aanduiden. Volgens tabel 9.1 is onder H_0 de kans op dit type gelijk aan $3 \times \frac{1}{729} = \frac{3}{729}$.

Het uitkomsttype dat qua mate van discrepantie t.o.v. de verwachte frequenties op het type (6, 0, 0) volgt is (5, 1, 0). Uit tabel 9.1 blijkt dat onder H_0 de kans op dit type gelijk is aan $6 \times \frac{6}{729} = \frac{36}{729}$. De kansen van de overige uitkomsttypen volgen eveneens uit tabel 9.1 en men verkrijgt nu de kansverdeling in tabel 9.2, waarin de typen (ongeveer) in volgorde van toenemende discrepantie t.o.v. het type (2, 2, 2) zijn geplaatst.[1]

Tabel 9.2. Kansverdeling van het uitkomsttype onder $H_0 : P_1 = P_2 = P_3 = 1/3$ voor $n = 6$

Type	Kans	Cumulatieve kans
2, 2, 2	90/729 = 0,1235	0,1235
3, 2, 1	360/729 = 0,4938	0,6173
3, 3, 0	60/729 = 0,0823	0,6996
4, 1, 1	90/729 = 0,1235	0,8231
4, 2, 0	90/729 = 0,1235	0,9466
5, 1, 0	36/729 = 0,0494	0,9960
6, 0, 0	3/729 = 0,0041	1,0001

Nu de kansverdeling van het uitkomsttype onder H_0 bekend is, kan men volgens de in 6.1.4 gegeven richtlijnen tot het opstellen van een beslissingsschema overgaan. De kritieke zone (waarin wij tot het verwerpen van H_0 zullen besluiten) dient te bestaan uit uitkomsttypen, die slecht met H_0 verenigbaar zijn. Het eerste type dat voor plaatsing in deze zone in aanmerking komt is vanzelfsprekend (6, 0, 0) met een kans van 0,0041. Voegt men ook type (5, 1, 0) met een kans van 0,0494 toe dan ontstaat een zone met een kans van 0,0041 + +0,0494 = 0,0535. Het type (4, 2, 0) kan niet meer bij de zone worden gevoegd, daar deze, met een kans van 0,0535 + 0,1235 = 0,1770, de

[1] Ongeveer, want deze volgorde is niet steeds evident [vergelijk bijv. de typen (3, 3, 0) en (4, 1, 1)]. Dit levert echter bij het volgende betoog geen bezwaar op.

gekozen waarde van α (= 0,10) ver zou overschrijden. De kritieke zone bestaat dus uit de typen (6, 0, 0) en (5, 1, 0) en daar de steekproefuitkomst $f_1 = 5$, $f_2 = 0$ en $f_3 = 1$ tot het laatstgenoemde type behoort wordt H_0 verworpen.

De toets die in het voorgaande is ontwikkeld kan in principe voor elke hypothese omtrent een populatie met k categorieën worden gebruikt. Het is dan echter noodzakelijk dat een objectief criterium voor de discrepantie tussen de waargenomen en de verwachte frequenties wordt gebruikt. Dit criterium wordt in 9.2.2 geïntroduceerd. Verder is reeds nu in te zien dat voor populaties met een groter

Tabel 9.1*B*. De multinomiale verdelingen:

I. $P(f_1, f_2, f_3) = \frac{6!}{f_1!\,f_2!\,f_3!}(0{,}8)^{f_1}(0{,}1)^{f_2}(0{,}1)^{f_3}$

II. $P(f_1, f_2, f_3) = \frac{6!}{f_1!\,f_2!\,f_3!}(0{,}1)^{f_1}(0{,}1)^{f_2}(0{,}8)^{f_3}$

A	B	C	I	II
f_1	f_2	f_3	$10^6P(f_1, f_2, f_3)$	$10^6P(f_1, f_2, f_3)$
6	0	0	262144	1
0	6	0	1	1
0	0	6	1	262144
5	1	0	196608	6
5	0	1	196608	48
0	5	1	6	48
0	1	5	6	196608
1	5	0	48	6
1	0	5	48	196608
4	2	0	61440	15
4	0	2	61440	960
0	4	2	15	960
0	2	4	15	61440
2	4	0	960	15
2	0	4	960	61440
4	1	1	122880	240
1	4	1	240	240
1	1	4	240	122880
3	3	0	10240	20
3	0	3	10240	10240
0	3	3	20	10240
3	2	1	30720	480
3	1	2	30720	3840
2	3	1	3840	480
2	1	3	3840	30720
1	1	2	480	3840
1	2	3	480	30720
2	2	2	5760	5760
Totaal			1000000	1000000

aantal categorieën en/of grotere steekproeven het uitrekenen van de exacte kansverdeling praktisch onmogelijk wordt. In 9.2.3 behandelen wij een verdeling die dan als benadering kan worden gebruikt. De uitvoering van de toetsing komt in 9.3 aan de orde.

In 6.2 hebben wij uiteengezet dat men feitelijk steeds het onderscheidingsvermogen van een toets t.o.v. de verschillende mogelijke alternatieve hypothesen dient te bestuderen. Bij de aldaar beschreven toetsing van de hypothese $P = P_0$ voor een populatie met twee categorieën kan men dit doen door de werkingskromme van de toets op te stellen, d.w.z. door voor een reeks alternatieve waarden van P na te gaan hoe groot de kans op een fout van de tweede soort β (resp. het onderscheidingsvermogen $1 - \beta$) bij een bepaalde steekproefomvang en bij een bepaalde kritieke zone (behorend bij de gekozen waarde van α) is. Bij de toetsing van een hypothese betreffende een populatie met $r(> 2)$ categorieën, zoals in het voorgaande is beschreven, ligt de zaak minder eenvoudig. Wij zullen dit aan de hand van een voorbeeld laten zien.

Wij gaan nogmaals uit van de toetsing van de hypothese $H_0 : P_1 = P_2 = P_3 = 1/3$, maar veronderstellen dat de alternatieve hypothese luidt $H_1 : P_1 = 0{,}8,\ P_2 = 0{,}1,\ P_3 = 0{,}1$.

Voor een aselect zonder teruglegging getrokken steekproef van 6 elementen is de kansverdeling per uitkomst onder H_0 in tabel 9.1*A* gegeven. De kansverdeling onder H_1 staat in tabel 9.1*B* onder I.

Volgens de in 6.1.4 gegeven procedure stellen wij nu een kritieke zone op bestaande uit uitkomsten met een kleine kans onder H_0 en een grote kans onder H_1 en wel zodanig, dat de som van de kansen van deze uitkomsten onder H_0 (in de verdeling in tabel 9.1*A*) ten hoogste gelijk is aan $\alpha = 0{,}10$, terwijl de som van de kansen van deze uitkomsten onder H_1 (in de verdeling in tabel 9.1*B* - I) zo groot mogelijk wordt. De lezer kan nu zelf vaststellen dat deze zone als volgt is opgebouwd:

Tabel 9.1*C*

Uitkomsten in Z			Kans onder	Kans onder
A	B	C	H_0	H_1
6	0	0	1/729	0,262144
5	1	0	6/729	0,196608
5	0	1	6/729	0,196608
4	2	0	15/729	0,061440
4	0	2	15/729	0,061440
4	1	1	30/729	0,122880
Totaal			73/729 α_Z	0,901120 $1 - \beta_Z$

Met deze kritieke zone bereikt men, met een kans op het ten onrechte verwerpen van H_0 van praktisch 0,10 ($\alpha_Z = 73/729$), een onderscheidingsvermogen t.o.v. H_1 van iets meer dan 0,90. Uit tabel 9.1*B* blijkt direct dat de kritieke zone die wij eerder samenstelden en die bestaat uit de *uitkomsttypen* (6, 0, 0) en (5, 1, 0) t.o.v. H_1 slechts een onderscheidingsvermogen van ongeveer 0,655 bezit:
[0,262144 + 2(0,000001) + 2(0,196608) + 2(0,000006) + 2(0,000048) = 0,655470].

Het is duidelijk, dat de speciaal voor deze ene alternatieve hypothese gekozen kritieke zone in tabel 9.1*C* voor vele andere alternatieve hypothesen een (zeer) klein onderscheidingsvermogen zal bezitten. In tabel 9.1*B* is bijv. onder II de kansverdeling van de uitkomsten opgenomen, die geldt voor het alternatief $H_1 : P_1 = 0{,}1$, $P_2 = 0{,}1$, $P_3 = 0{,}8$. Voor dit alternatief heeft de zone in tabel 9.1*C*, zoals men zelf kan nagaan, slechts een onderscheidingsvermogen van 0,001270. De zone bestaande uit de uitkomsttypen bezit echter t.o.v. dit alternatief *eveneens* een onderscheidingsvermogen van ongeveer 0,655. Hetzelfde geldt voor de alternatieve hypothese $H_1 : P_1 = 0{,}1$, $P_2 = 0{,}8$, $P_3 = 0{,}1$.

In de praktijk is bij het toetsen van een hypothese omtrent een populatie met r categorieën zelden of nooit een specifieke alternatieve hypothese aanwezig. Zelfs de vorm van zo'n alternatieve hypothese (in de zin van $P_1 = P_2 < P_3$, $P_1 > P_2 > P_3$, e.d.) kan slechts bij uitzondering worden aangegeven. Men bereikt dan bij een gekozen waarde van α en een bepaalde steekproefomvang een 'redelijk' onderscheidingsvermogen voor allerlei alternatieve hypothesen door de kritieke zone van de toets op het uitkomsttype (of beter gezegd, zoals wij in 9.2.2 zullen zien, op een daaruit volgende toetsingsgrootheid G) te baseren. Dit onderscheidingsvermogen zal toenemen, naarmate de steekproefomvang groter wordt.

9.2.2. G ALS MAAT VOOR DE DISCREPANTIE TUSSEN WAARGENOMEN EN VERWACHTE FREQUENTIES

Evenals in 9.2.1 gebruiken wij in het vervolg de symbolen:

f_i = de waargenomen frequentie in categorie i. Als er r categorieën zijn, is dus $\Sigma f_i = n$.

$e_i = nP_i$ = de verwachte frequentie in categorie i onder H_0. Dus ook $\Sigma e_i = n$.

Daar $\Sigma f_i = \Sigma e_i$, is $\Sigma(f_i - e_i) = 0$ en zodoende onbruikbaar als maat voor de discrepantie tussen het waargenomen en het verwachte uitkomsttype. Dit bezwaar kan worden opgeheven door de verschillen $f_i - e_i$ te kwadrateren. Elke waarde $(f_i - e_i)^2$ dient echter te worden beschouwd in betrekking tot e_i, daar een verschil van bijv. 10 bij

$e_i = 10$ veel belangrijker is dan een zelfde verschil bij $e_i = 100$. Deze redenering brengt ons tot de grootheid

$$G = \frac{(f_1 - e_1)^2}{e_1} + \frac{(f_2 - e_2)^2}{e_2} + \ldots + \frac{(f_r - e_r)^2}{e_r},$$

dus

$$G = \sum_{i=1}^{r} \frac{(f_i - e_i)^2}{e_i} \tag{9.6}$$

als maat voor de discrepantie tussen de in de steekproef waargenomen en de onder H_0 verwachte frequenties. G is een discrete grootheid, die geen negatieve waarden aanneemt. Als alle verschillen $f_i - e_i$ nul zijn (d.i. als voor elke categorie de waargenomen en de verwachte frequentie geheel overeenstemmen) is $G = 0$. Naarmate de overeenstemming slechter wordt zijn de absolute waarden van de verschillen $f_i - e_i$ groter en neemt G een grotere waarde aan.

Voor het uitkomsttype (4, 1, 1) (zie 9.2.1) verloopt de berekening van de bijbehorende waarde van G als volgt:

Categorie	f_i	e_i	$f_i - e_i$	$(f_i - e_i)^2$	$\frac{(f_i - e_i)^2}{e_i}$
A	4	2	2	4	2,0
B	1	2	—1	1	0,5
C	1	2	—1	1	0,5
			0		3,0 = G

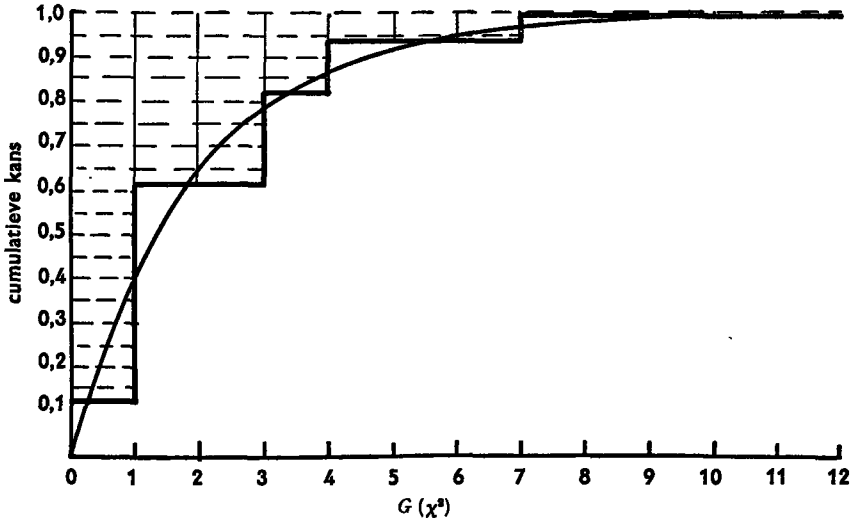
Van elk uitkomsttype kan op deze wijze de corresponderende waarde van G worden berekend. Men vindt de waarden in tabel 9.3; de kans van elke waarde kan aan tabel 9.2 worden ontleend. Daar de uitkomsttypen (4, 1, 1) en (3, 3, 0) dezelfde waarde van G opleveren zijn hun kansen in tabel 9.3 gecombineerd. Men ziet, dat $G = 0$ voor het type (2, 2, 2) en dat G de grootste waarde aanneemt voor het type (6, 0, 0) dat hier het meest van afwijkt.

Tabel 9.3. Exacte kansverdeling van G voor het probleem in 9.2.1

$n = 6$, $P_1 = P_2 = P_3 = \frac{1}{3}$

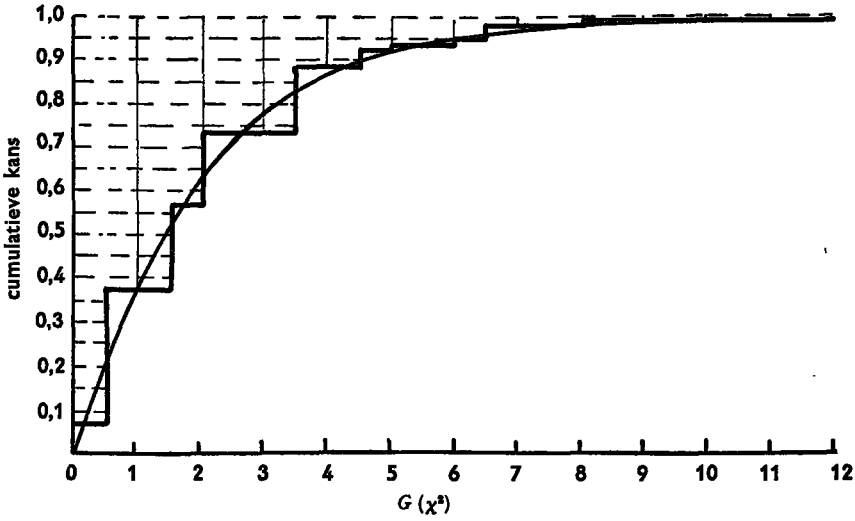
Type	G	P	Cumulatieve kans
2,2,2	0	0,1235	0,1235
3,2,1	1	0,4938	0,6173
3,3,0	3	0,2058	0,8231
4,1,1	3		
4,2,0	4	0,1235	0,9466
5,1,0	7	0,0494	0,9960
6,0,0	12	0,0041	1,0001

Het blijkt nu, dat de discrete kansverdeling van $\underline{G}$ (zoals deze in tabel 9.3 voor het toetsingsprobleem in 9.2.1 is opgenomen) benaderd kan worden door de best-passende χ^2-verdeling (d.i. een theoretische continue kansverdeling, die wij in 9.2.3 bespreken). De benadering is reeds bevredigend, als de *verwachte frequenties* ten minste gelijk zijn aan 5 en zij is zeer nauwkeurig voor grote steekproeven. Hoewel de exacte verdeling van $\underline{G}$ in tabel 9.3 slechts betrekking heeft op een steekproef van 6 elementen met verwachte frequenties, die aanzienlijk lager zijn dan 5, vertoont de best-passende χ^2-verdeling zelfs in dit geval reeds een vrij redelijke overeenstemming met de verdeling van $\underline{G}$ (zie figuur 9.1). Figuur 9.2 laat zien, dat de benadering beter wordt bij toenemende n. De (exact berekende) somfrequentieverdeling van $\underline{G}$ voor $n = 12$ is in deze figuur vergeleken met de best-passende cumulatieve χ^2-verdeling. Men dient hierbij te bedenken, dat de nauwkeurigheid van de aanpassing vooral van belang is voor betrekkelijk grote waarden van $\underline{G}$, dus bij cumulatieve kansen tussen 0,90 en 1.



Figuur 9.1. Cumulatieve verdeling van $\underline{G}$ en de best-passende cumulatieve χ^2-verdeling bij toetsing van de hypothese $P_1 = P_2 = P_3 = \frac{1}{3}$, $n = 6$.

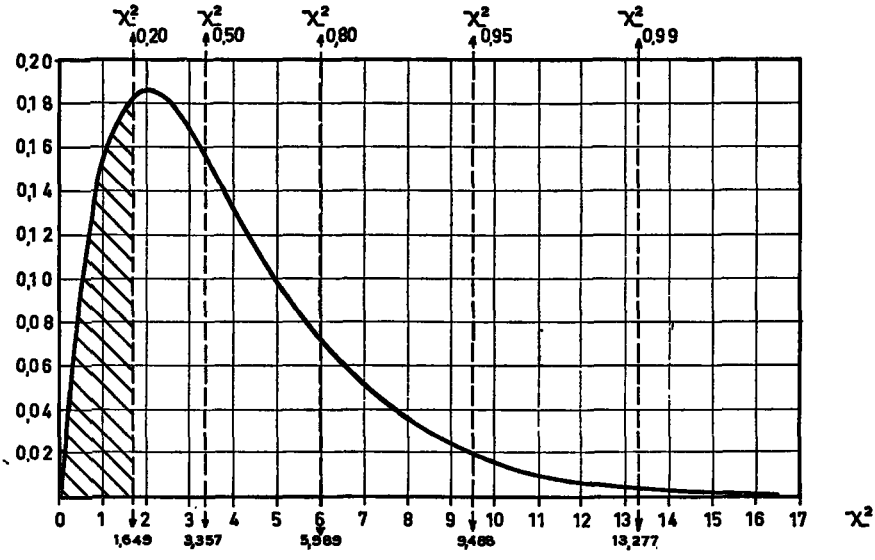
Bij het toetsen van hypothesen betreffende populaties met r categorieën kan men dus, indien het aantal waarnemingen niet te klein is, gebruik maken van de toetsingsgrootheid G, waarvan de exacte kansverdeling benaderd wordt met de best-passende χ^2-verdeling. Dit maakt het mogelijk een benaderende overschrijdingskans van G_0 (de uit een steekproef berekende waarde van $\underline{G}$) direct af te lezen uit tabellen, waarin fractielen van χ^2-verdelingen zijn opgenomen.



Figuur 9.2. Als figuur 9.1, maar voor $n = 12$.

9.2.3. χ^2-VERDELINGEN

Het gebruik van tabel D. Een χ^2-verdeling is een kansverdeling, die volledig gespecificeerd wordt door één parameter, het *aantal vrijheidsgraden* ν (de Griekse letter nu; men gebruikt ook wel d van het Engelse 'degrees of freedom'). In tabel D zijn verschillende fractielen

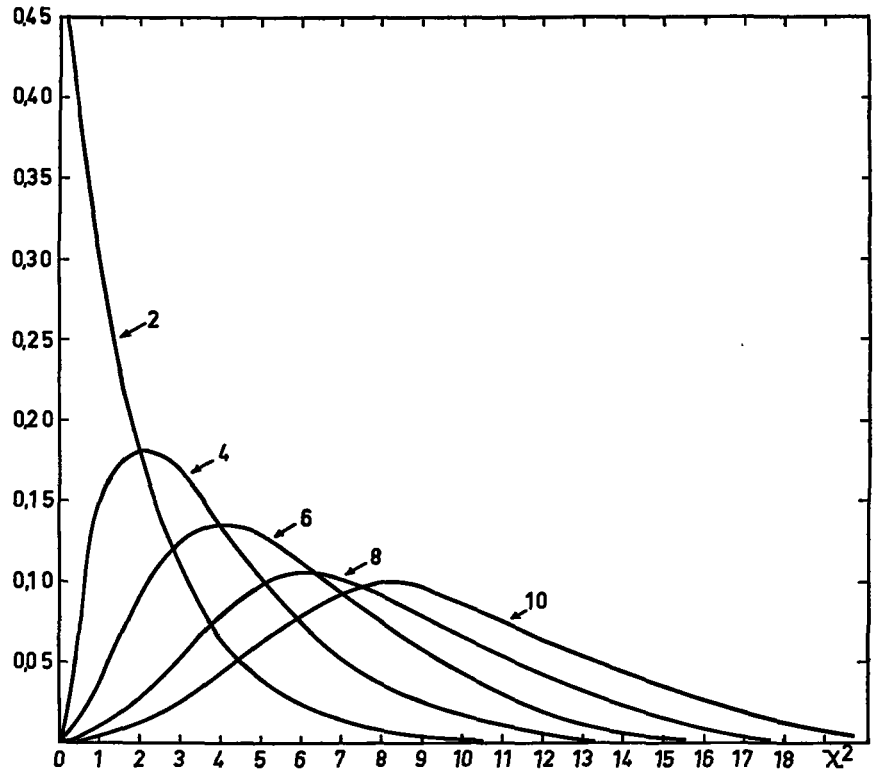


Figuur 9.3. χ^2-verdeling met 4 vrijheidsgraden.

van de χ^2-verdelingen met $\nu = 1(1)30$ gegeven. Elke regel in deze tabel bevat fractielen van één χ^2-verdeling. Beschouw bijv. de χ^2-verdeling met $\nu = 4$, die in figuur 9.3 is getekend. In tabel D leest men op de regel $\nu = 4$ en in de kolom, waarboven $\chi^2_{0,20}$ staat, de waarde 1,649 af, dus kortweg: $\chi^2_{0,20}(\nu = 4) = 1{,}649$. Dat wil zeggen, dat voor *deze* χ^2-verdeling geldt: het fractiel 0,20 (20e percentiel, 2e deciel) is 1,649, d.w.z. 20% van het oppervlak onder de kromme in figuur 9.3 ligt links van de ordinaat in het punt 1,649 op de X-as. Men kan ook schrijven:
$P(\underline{\chi}^2 < 1{,}649) = P_L(1{,}649) = 0{,}20$ en $P(\underline{\chi}^2 > 1{,}649) = P_R(1{,}649) = 0{,}80$.

Is van deze χ^2-verdeling een waarde gegeven, dan kan men in het algemeen met tabel D de linkse of rechtse kans van deze waarde niet precies bepalen. Van $\chi^2(\nu = 4) = 11$ kan men bijv. slechts zeggen. dat de linkse kans groter is dan 0,95 en kleiner dan 0,99 (zie figuur 9.3).

In figuur 9.4 zijn de χ^2-verdelingen met 2, 4, 6, 8 en 10 vrijheidsgraden getekend. Deze figuur toont aan, dat de vorm van een χ^2-verdeling afhangt van de parameter ν. De verdelingen met een klein



Figuur 9.4. χ^2-verdelingen met 2, 4, 6, 8 en 10 vrijheidsgraden.

aantal vrijheidsgraden zijn (zeer) scheef naar rechts, maar die met een groter aantal vrijheidsgraden neigen reeds tot symmetrie. Tenslotte vermelden wij, dat het gemiddelde van een χ^2-verdeling gelijk is aan ν, dat haar modus $\nu - 2$ is (behoudens voor $\nu \leqslant 2$) en haar variantie 2ν.

Normale benaderingen. FISHER heeft aangetoond dat men voor $\nu > 30$ de rechtse kans van een waarde van χ^2 bij benadering kan bepalen door.

$$T = \sqrt{2\chi^2} - \sqrt{2\nu - 1} \tag{9.7}$$

te berekenen en de rechtse kans van T in tabel A op te zoeken. Neemt men bijv. $\chi^2 = 43{,}773$ en $\nu = 30$ dan vindt men

$$T = \sqrt{2(43{,}773)} - \sqrt{2(30) - 1} = 1{,}675, \quad P_R(1{,}675) = 0{,}047.$$

Uit tabel D blijkt dat de exacte rechtse kans gelijk is aan 0,05.

WILSON en HILFERTY (1932) hebben een normale benadering gegeven, die ook voor kleine waarden van ν bevredigend is. Hierbij neemt men

$$T = \frac{\sqrt[3]{\frac{\chi^2}{\nu}} - \left(1 - \frac{2}{9\nu}\right)}{\sqrt{\frac{2}{9\nu}}} = 3\left[\sqrt[3]{\frac{\chi^2}{\nu}} + \frac{2}{9\nu} - 1\right]\sqrt{\frac{\nu}{2}}. \tag{9.8}$$

Voor $\chi^2 = 4{,}168$ ($\nu = 9$) vindt men met (9.8)

$$T = 3[\sqrt[3]{4{,}168/9} + 2/9^2 - 1]\sqrt{9/2} = 3(0{,}7737 + 0{,}0247 - 1)\sqrt{4{,}5} = -1{,}28.$$

Uit tabel A blijkt dat $P_L(-1{,}28) = 0{,}1003$. Tabel D geeft als exacte rechtse kans 0,10.

Vrijheidsgraden. Onderstel dat men zonder enige beperking 5 getallen opschrijft. Elk der getallen kan volkomen vrij worden gekozen. Men drukt dit uit door te zeggen: er zijn 5 vrijheidsgraden.

Onderstel vervolgens, dat gevraagd wordt 5 getallen op te schrijven met de restrictie, dat hun som een bepaalde waarde, zeg 50, dient te hebben. Er kunnen slechts 4 getallen vrij worden gekozen, want het vijfde is reeds bepaald door de gestelde voorwaarde: $x_1 + x_2 + x_3 + x_4 + x_5 = 50$. Het aantal vrij te kiezen getallen is $5 - 1 = 4$ dus zijn er 4 vrijheidsgraden.

Als gevraagd wordt 5 getallen op te schrijven, zodanig dat hun som gelijk is aan 50 èn de som van de eerste twee getallen gelijk is aan 20, dan zijn er twee onafhankelijke restricties: $x_1 + x_2 + x_3 + x_4 + x_5 = 50$ en $x_1 + x_2 = 20$. Zodra x_1 gekozen is, ligt x_2 vast: $x_2 = 20 - x_1$. En

$x_3 + x_4 + x_5 = 50 - 20 = 30$, zodat hiervan slechts twee getallen vrij kunnen worden gekozen. Het aantal vrijheidsgraden is hier dus $1 + 2 = 3$, d.i. het totale aantal 'keuzen' verminderd met het aantal onafhankelijke restricties: $5 - 2 = 3$.

Voor elk statistisch probleem, waarbij vrijheidsgraden betrokken zijn, kan men het aantal daarvan op overeenkomstige wijze vastellen, d.w.z. door het totale aantal 'keuzen' met het aantal daaraan verbonden onafhankelijke restricties te verminderen. Beschouw bijv. het probleem in voorbeeld 9.2. De populatie bestaat uit elementen, die het kenmerk A, B of C dragen. De elementen in de steekproef behoren dus tot één der categorieën A, B of C. De hiermede corresponderende steekproeffrequenties f_1, f_2 en f_3 zijn onderhevig aan de restrictie $f_1 + f_2 + f_3 = 6$, d.w.z. als de frequenties van twee categorieën bekend zijn, ligt de frequentie van de overblijvende categorie vast. Er zijn bij dit probleem dus $3 - 1 = 2$ vrijheidsgraden. Als men een steekproef van 500 elementen trekt, is $f_1 + f_2 + f_3 = 500$ en het aantal vrijheidsgraden blijft $3 - 1 = 2$. Voor *elk* probleem, waarbij een hypothese wordt getoetst betreffende de samenstelling van een populatie, die 3 categorieën telt en waarbij de enige restrictie is gelegen in de steekproefomvang, zijn er dus 2 vrijheidsgraden. Bevat de populatie in zo'n geval r categorieën, dan zijn er $r - 1$ vrijheidsgraden. In de volgende paragrafen komen echter ook problemen met r categorieën en m onafhankelijke restricties ter sprake: het aantal vrijheidsgraden is dan $r - m$.

De best-passende χ^2-verdeling. In 9.2.2 hebben wij reeds gezegd, dat de verdeling van de toetsingsgrootheid $\underline{G}$ bij het toetsen van de hypothese $P_1 = P_2 = P_3 = \frac{1}{3}$ onder bepaalde voorwaarden kan worden benaderd door de best-passende χ^2-verdeling.

Bij toetsingsproblemen van dit type (met één restrictie: de steekproefomvang) is dit de *χ^2-verdeling met $r - 1$ vrijheidsgraden.* De keuze van deze best-passende verdeling berust dus uitsluitend op r en niet op n. Bij problemen, waarbij r hetzelfde is zal de χ^2-verdeling echter een betere benadering geven naarmate de steekproefomvang n groter is.

De uitvoering van de toetsing met de χ^2-benadering bespreken wij nu in 9.3.

9.2.4. OPGAVEN

9.1. Verifieer de volgende kansen met tabel D:

a. $P(\underline{\chi}^2 < 2{,}706 \mid \nu = 1) = 0{,}90$
b. $P(\underline{\chi}^2 > 3{,}841 \mid \nu = 1) = 0{,}05$
c. $P(\underline{\chi}^2 > 5{,}991 \mid \nu = 2) = 0{,}05$
d. $P(\underline{\chi}^2 < 11{,}070 \mid \nu = 5) = 0{,}95$
e. $P(\underline{\chi}^2 > 30{,}578 \mid \nu = 15) = 0{,}01$
f. $P(0{,}103 < \underline{\chi}^2 < 5{,}991 \mid \nu = 2) = 0{,}90$
g. $P(\underline{\chi}^2 > 0{,}711 \mid \nu = 4) = 0{,}95$
h. $P(\underline{\chi}^2 > 9{,}342 \mid \nu = 10) = 0{,}50$
i. $0{,}02 > P(\underline{\chi}^2 > 16{,}3 \mid \nu = 6) > 0{,}01$
j. $0{,}025 < P(\underline{\chi}^2 > 20 \mid \nu = 11) < 0{,}05$.

9.2. Bepaal met tabel D de grenzen, waartussen de volgende kansen liggen:

a. $P(\underline{\chi}^2 > 7{,}1 \mid \nu = 1)$
b. $P(\underline{\chi}^2 > 7{,}1 \mid \nu = 2)$
c. $P(\underline{\chi}^2 > 7{,}1 \mid \nu = 4)$
d. $P(\underline{\chi}^2 > 7{,}1 \mid \nu = 8)$
e. $P(\underline{\chi}^2 > 7{,}1 \mid \nu = 16)$
f. $P(\underline{\chi}^2 < 5{,}5 \mid \nu = 4)$
g. $P(\underline{\chi}^2 < 22 \mid \nu = 10)$
h. $P(\underline{\chi}^2 > 42{,}5 \mid \nu = 25)$
i. $P(\underline{\chi}^2 < 8{,}75 \mid \nu = 20)$
j. $P(\underline{\chi}^2 > 22{,}5 \mid \nu = 5)$.

9.3. Bepaal bij benadering de volgende kansen:

a. $P(\underline{\chi}^2 > 57{,}51 \mid \nu = 35)$ *b.* $P(\underline{\chi}^2 > 45 \mid \nu = 41)$.

9.3. De χ^2-toets betreffende een populatie met r categorieën

9.3.1. UITVOERING VAN DE TOETSING

De χ^2-benadering betreft toetsing van hypothesen van de vorm H_0: $P_1 = P_2 = \ldots = P_r$ of $P_1 : P_2 : \ldots : P_r = a : b : \ldots : m$. De toetsingsgrootheid is

$$G = \sum_{i=1}^{r} \frac{(f_i - e_i)^2}{e_i} \qquad (e_i = P_i n). \tag{9.9.1}$$

Deze grootheid volgt onder H_0 bij benadering een χ^2-verdeling met $r - 1$ vrijheidsgraden, indien n niet te klein is, zodat de verwachte frequenties e_i minstens gelijk zijn aan 5.

De toetsing geschiedt rechts éénzijdig, daar een slechte overeenstemming tussen de waargenomen en verwachte frequenties tot hoge waarden van $\underline{G}$ leidt. Bij een onbetrouwbaarheidsdrempel α zal men tot het verwerpen van H_0 besluiten voor

$$G_0 \geqslant \chi^2_{1-\alpha} \ (\nu = r - 1).$$

G_0 is de uit de steekproef berekende waarde van $\underline{G}$. De kritieke waarde $\chi^2_{1-\alpha}$ kan voor verschillende waarden van α en $\nu = 1(1)30$ uit tabel D worden afgelezen. Bijv.:

α	0,05	0,025	0,02	0,01	0,005	0,001
$\chi^2_{1-\alpha}$	$\chi^2_{0,95}$	$\chi^2_{0,975}$	$\chi^2_{0,98}$	$\chi^2_{0,99}$	$\chi^2_{0,995}$	$\chi^2_{0,999}$
$\nu = 2$	5,991	7,378	7,824	9,210	10,597	13,815
$\nu = 3$	7,815	9,348	9,837	11,345	12,838	16,268
$\nu = 5$	11,070	12,832	13,388	15,086	16,750	20,517

Toetst men bijv. bij een drempelwaarde 0,01 en telt de populatieverdeling 6 categorieën, dan wordt H_0 verworpen indien $G_0 \geqslant 15{,}086$ $[= \chi^2_{0,99} \ (\nu = 5)]$ is.

Voorbeelden

9.3. Men wil onderzoeken of een dobbelsteen zuiver is en werpt daartoe 360 keer met deze steen. De volgende uitkomsten worden waargenomen:

Aantal ogen $= x_i$	1	2	3	4	5	6	Totaal
f_i	67	52	39	72	55	75	360

De toetsing verloopt als volgt:

a. De te toetsen hypothese luidt $H_0 : P(1) = P(2) = \ldots = P(6) = {}^1/_6$.

b. Men kiest $\alpha = 0{,}05$, d.w.z. men is bereid H_0 te verwerpen, indien de overschrijdingskans van de waargenomen waarde van de toetsingsgrootheid $< 0{,}05$ is.

c. Onder H_0 is de verwachte frequentie per categorie (aantal ogen): $e_i = 360 \times {}^1/_6 = 60$. De toetsingsgrootheid is

$$G = \sum_{i=1}^{6} \frac{(f_i - e_i)^2}{e_i} = \frac{\Sigma(f_i - e_i)^2}{e_i}.$$

De grootheid G volgt bij benadering een χ^2-verdeling met $6 - 1 = 5$ vrijheidsgraden.

d. Er wordt rechts éénzijdig getoetst. Uit tabel D blijkt, dat H_0 bij de gekozen drempelwaarde 0,05 wordt verworpen, indien

$$G_0 > 11{,}07 \text{ [daar } \chi^2_{0{,}95}(\nu = 5) = 11{,}070\text{]}.$$

e. Bereken

$$G_0 = \frac{(67 - 60)^2 + (52 - 60)^2 + (39 - 60)^2 + \ldots + (75 - 60)^2}{60} = 15{,}8.$$

f. Verwerp H_0, daar G_0 groter is dan de kritieke waarde, 11,07.
(Uit tabel D blijkt, dat de overschrijdingskans van G_0 kleiner is dan 0,01.)

Opmerking: De uitkomsten betreffen een opzettelijk vervalste dobbelsteen. In het middelste oog van het vlakje met 3 ogen werd een gaatje geboord, waarin een stukje lood werd aangebracht.

9.4. Bij kruising van twee soorten erwten (met gele ronde, resp. met groene gerimpelde zaden) vond MENDEL de in tabel 9.4 onder f_i vermelde frequenties voor de verschillende zaadsoorten. Volgens het genetisch model zouden deze frequenties de verhouding 9 : 3 : 3 : 1 moeten hebben. Men wil de overeenstemming onderzoeken tussen dit model en de waarnemingen.

Tabel 9.4. Frequenties van de zaadsoorten bij kruising van twee soorten erwten

Categorie	Waargenomen f_i	Verwacht e_i	$f_i - e_i$	$(f_i - e_i)^2/e_i$
A. Rond en geel	315	312,75	2,25	0,016
B. Gerimpeld en geel	101	104,25	—3,25	0,101
C. Rond en groen	108	104,25	3,75	0,135
D. Gerimpeld en groen	32	34,75	—2,75	0,218
Totaal	556	556,00	0,00	$G_0 = 0{,}470$

De toetsing verloopt als volgt:

a. H_0: $P(A) = {}^9/_{16}$, $P(B) = {}^3/_{16}$, $P(C) = {}^3/_{16}$, $P(D) = {}^1/_{16}$.

b. Kies $\alpha = 0{,}05$.

c. Onder H_0 volgt de grootheid G [volgens (9.3)] bij benadering een χ^2-verdeling met $\nu = 3$.

d. De toetsing is rechts éénzijdig. Uit tabel D blijkt, dat H_0 bij de drempelwaarde 0,05 wordt verworpen voor $G_0 > 7{,}815$.

e. Bereken G_0. Deze berekening is in tabel 9.4 uitgevoerd en levert $G_0 = 0{,}470$ op.

f. H_0 wordt niet verworpen. De rechtse kans van G_0 ligt tussen 0,90 en 0,95, zodat de overeenstemming tussen de waarnemingen en het genetisch model bijzonder fraai is.

Als de waarnemingsuitkomsten in fracties (of percentages) zijn gegeven kan men G_0 als volgt berekenen:

$$G_0 = n \sum_{i=1}^{r} \frac{(p_i - P_i)^2}{P_i}, \tag{9.9.2}$$

waarbij p_i = de waargenomen fractie in categorie i, P_i = de kans van categorie i onder H_0. Uit deze formule blijkt, dat de waarde van G bij gelijkblijvende verhouding van de waargenomen frequenties recht evenredig is aan n. Verder toont zij, dat men de waarde van G niet kan berekenen, als uitsluitend fracties of percentages bekend zijn en n onbekend is. Voert men in zo'n geval de berekening bijv. met de gegeven percentages uit, dan handelt men alsof $n = 100$. Is de steekproefomvang in werkelijkheid groter, dan is de berekende G_0 een onderschatting van de juiste waarde; bij een kleinere steekproef verkrijgt men een geflatteerde uitkomst.

9.3.2. POPULATIEVERDELINGEN MET 2 CATEGORIEËN

Als de populatieverdeling dichotoom is en de categorieën A en B bevat, kan men de hypothesen omtrent deze verdeling, van de vorm H_0: $P(A) = P = P_0$ toetsen volgens de in 6.2 en 6.3 behandelde methoden. In plaats van de aldaar gegeven normale benadering kan men ook de in deze paragraaf besproken χ^2-benadering toepassen. De best-passende χ^2-verdeling bezit dan $2 - 1 = 1$ vrijheidsgraad.

Onderstel dat men de hypothese $P = P_0$ toetst en dat een aselecte steekproef van de omvang n bestaat uit x_0 elementen A en $n - x_0$ elementen B. Dit zijn dus de waargenomen frequenties; onder H_0 zijn de verwachte frequenties P_0n elementen A en $Q_0n = n - P_0n$ elementen B. Substitutie in (9.9.1) levert op

$$\begin{aligned} G_0 &= \frac{(x_0 - P_0n)^2}{P_0n} + \frac{\{(n - x_0) - (n - P_0n)\}^2}{Q_0n} \\ &= \frac{(x_0 - P_0n)^2 Q_0}{P_0Q_0n} + \frac{(P_0n - x_0)^2 P_0}{P_0Q_0n} \\ &= \frac{(x_0 - P_0n)^2 (P_0 + Q_0)}{P_0Q_0n} = \frac{(x_0 - P_0n)^2}{P_0Q_0n}. \end{aligned}$$

In dit geval, met $\nu = 1$, kan de χ^2-benadering worden verbeterd door het aanbrengen van een continuïteitscorrectie, die bestaat uit een vermindering van de absolute waarde van $x_0 - P_0n$ met $\frac{1}{2}$. Men verkrijgt dan:

$$G_0 = \frac{(|\, x_0 - P_0n \,| - \frac{1}{2})^2}{P_0Q_0n}. \tag{9.10}$$

Vergelijkt men deze formule met die voor de normale benadering in 6.3, formule (6.7), dan blijkt te gelden: $G_0 = T_0^2$. Voor een *dichotome* populatie is dus de χ^2-benadering met $\nu = 1$ identiek aan de normale benadering bij *tweezijdige* toetsing.

Wil men via de χ^2-benadering éénzijdig toetsen bij een drempelwaarde α dan geldt bij:
H_0: $P < \frac{1}{2}$: Indien $x_0 > P_0n$ wordt H_0 verworpen voor $G_0 > \chi^2_{1-2\alpha}$ $(\nu = 1)$
H_0: $P > \frac{1}{2}$: Indien $x_0 < P_0n$ wordt H_0 verworpen voor $G_0 > \chi^2_{1-2\alpha}$ $(\nu = 1)$
Bij éénzijdige toetsing met $\alpha = 0{,}05$ verwerpt men dus H_0, indien $G_0 > \chi^2_{0,90}$ $(\nu = 1)$, d.i. voor $G_0 > 2{,}706$.

9.3.3. TOEPASSING OP KWANTITATIEVE WAARNEMINGEN

Men kan de χ^2-toets ook toepassen, als men de hypothese wil toetsen dat een steekproef afkomstig is uit een populatie van een bepaald type (bijv. normaal, binomiaal, POISSON). Wij bespreken deze toepassing in 10.1.

9.3.4. COMBINATIE VAN CATEGORIEËN BIJ χ^2-TOETSEN

Om te lage verwachte frequenties te voorkomen zal men soms tot het combineren van categorieën moeten overgaan. Daar hierbij verlies aan informatie optreedt, is het aanbevelenswaardig niet meer te combineren dan strikt noodzakelijk is. In principe mag deze combinatie van categorieën niet plaatsvinden op grond van de waargenomen uitkomsten.

Voorbeeld

9.5. Een proefpersoon krijgt de opdracht willekeurig 500 getallen van drie cijfers op te schrijven. Hierbij verkrijgt men voor het laatste cijfer van deze getallen de frequentieverdeling in tabel 9.5 (f). Men kan nu de hypothese toetsen dat ieder cijfer dezelfde kans heeft om genoteerd te worden, met als tegenhypothese dat bij de proefpersoon een preferentie voor sommige cijfers bestaat. Als men deze proef voor de eerste keer verricht zal men echter a priori niet weten welke cijfers dit zijn.

Tabel 9.5. Verdeling van de laatste cijfers van 500 willekeurig genoteerde getallen van drie cijfers

Laatste cijfer	0	1	2	3	4	5	6	7	8	9	Totaal
f	51	59	44	41	46	43	46	48	60	62	500
e	50	50	50	50	50	50	50	50	50	50	500
$f-e$	1	9	−6	−9	−4	−7	−4	−2	10	12	0
$(f-e)^2$	1	81	36	81	16	49	16	4	100	144	528

$$G_0(9) = \frac{528}{50} = 10{,}56 \quad (P \approx 0{,}3)$$

Het achter G_0 tussen haakjes geplaatste getal geeft het aantal vrijheidsgraden van de best-passende χ^2-verdeling aan.

Laatste cijfer	f	e	$(\lvert f-e\rvert -0{,}5)^2$
0, 1, 8, 9	232	200	992,25
2, 3, 4, 5, 6, 7	268	300	992,25

$$G_0(1) = \frac{992{,}25}{200} + \frac{992{,}25}{300} = 8{,}27, \qquad P < 0{,}01$$

In tabel 9.5 is eerst G_0 onder de te toetsen hypothese berekend. De uitkomst, $G_0(9) = 10{,}56$ geeft geen reden tot het verwerpen van H_0. Bestudeert men echter de afwijkingen $f_i - e_i$, dan is daarin een zeker

systeem te onderkennen: de cijfers 2 t/m 7 vertonen te lage waargenomen frequenties, de overige cijfers te hoge. Vormt men op grond hiervan de categorieën 0, 1, 8, 9 en 2 t/m 7 en toetst men opnieuw, dan is de uitkomst $G_0(1) = 8{,}27$, met een overschrijdingskans tussen 0,005 en 0,001, zodat men tot verwerpen van H_0 zou besluiten. Deze handelwijze is echter niet geoorloofd, omdat men de tegenhypothese (preferentie voor 0, 1, 8 en 9) heeft verkregen door selectie uit waarnemingsuitkomsten, die bij toetsing met de juiste tegenhypothese (preferentie voor ?) niet significant zijn gebleken. Wil men verantwoord te werk gaan, dan dient men een nieuwe proef uit te voeren. Voor de daarbij verkregen uitkomsten kan men de hypothese 'geen preferentie' toetsen met als alternatief 'preferentie voor 0, 1, 8 en 9'.

9.3.5. OPGAVEN

9.4. Deze opgave is bestemd voor klassikale uitvoering. Ieder trekt aselect 5 (10) steekproeven van 25 elementen uit de nevenstaande populatie. De trekking kan worden uitgevoerd met een tabel met aselecte getallen, waarbij men neemt:

Categorie	A	B	C
P	0,4	0,4	0,2

$A = 0, 1, 2, 3$, $B = 4, 5, 6, 7$ en $C = 8$ en 9.

Bereken voor elke steekproef G_0 en stel van alle gevonden waarden een frequentieverdeling op. Vergelijk $G_{0,90}$, $G_{0,95}$ en $G_{0,99}$ van deze verdeling met $\chi^2_{0,90}$, $\chi^2_{0,95}$ en $\chi^2_{0,99}$ van de best-passende χ^2-verdeling ($\nu = 2$).

9.5. Toets de hypothese, dat onderstaande steekproef afkomstig is uit een populatie met $P(A) = P(B) = P(C) = P(D) = \frac{1}{4}$. Kies $\alpha = 0{,}05$.

Categorie	A	B	C	D	Totaal
f	65	40	42	53	200

9.6. Bij 300 worpen met een dobbelsteen verkrijgt men de volgende uitkomsten:

x	1	2	3	4	5	6
f	45	40	38	53	60	64

a. Welke hypothese toetst u? Tot welke conclusie komt u?
Op grond van deze uitkomsten besluit men nogmaals 300 worpen te verrichten. De uitkomsten zijn:

x	1	2	3	4	5	6
f	41	36	44	63	56	60

b. Welke hypothese toetst u nu? Geef uw conclusie en commentaar.

9.7. Een kruisingsproef levert een nakomelingschap met planten van de vier typen $i = 1, 2, 3$ en 4 in de aantallen $n_1 = 90$, $n_2 = 36$, $n_3 = 26$, $n_4 = 4$. Toets (met $\alpha = 0{,}05$) de hypothese dat deze aantallen een verwezenlijking zijn van een kansverdeling waarin de kansen zich verhouden als 9 : 3 : 3 : 1.

9.8. Uit: Certificate Examination, Royal Statistical Society, 1948. Nevenstaande tabel vermeldt het aantal dagen (f), waarop de regenval in een bepaalde plaats in een periode van een jaar, R inches overschreed. Toets de hypothese, dat de regenval de 'wet': $\log_{10} f = 2{,}47 - 1{,}98R$ volgt.

R	0,00	0,04	0,10	0,20	0,50	1,00
f	296	246	187	119	30	3

9.4. De tekentoets

9.4.1. TOEPASSING

Deze toets ontleent zijn naam aan het feit, dat hij kan worden toegepast als men aan de verschillen tussen *gepaarde* waarnemingen een teken kan toekennen, d.i. kan aangeven welke van de twee waarnemingen 'groter' is. De tekentoets dient voor het toetsen van de nulhypothese dat de verschillen afkomstig zijn uit een populatie, waarvan de mediaan gelijk is aan nul. De praktische hypothese die op deze wijze getoetst wordt is dat de verschillen tussen de gepaarde waarnemingen toevallig zijn.

Voorbeeld

9.6. Naar VAN ELTEREN en RÜMKE (1956). In een cosmetisch preparaat kan als grondstof de stof A, dan wel B, worden gebruikt. Bij applicatie van het preparaat bestaat het gevaar, dat de gebruikte stof per ongeluk in het oog komt. Welke stof zou in dat geval het minst prikkelend werken?

Om dit te onderzoeken zijn beide stoffen verdund tot de hoogste concentratie, die praktisch in het oog zou kunnen komen. Vervolgens is bij 17 konijnen in het ene oog een druppel van de oplossing van stof A en in het andere een druppel van die van stof B gebracht. De volgende dag is nagegaan, welke van beide ogen het sterkst gereageerd had. Hoewel bij beide ogen een geringe reactie optrad, bleek het toch mogelijk steeds een uitspraak te doen. De proef leverde op, dat bij 16 konijnen de irritatie bij A het sterkst was (+), terwijl bij 1 dier stof B de sterkste irriatie opleverde (—).

9.4.2. DE TOETSINGSGROOTHEID

Als toetsingsgrootheid wordt gekozen het aantal plussen, x. Voor het voorbeeld is dus $x_0 = 16$.

9.4.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

Onder de nulhypothese is voor elk der n onafhankelijke waarnemingsparen de kans op een verschil dat groter is dan nul gelijk aan de kans op een verschil dat kleiner is dan nul, of:

$$H_0 : P(+) = P(-) = \tfrac{1}{2}.$$

Onder H_0 volgt zodoende de grootheid $\underset{\sim}{x}$, het aantal plussen bij n gepaarde waarnemingen, de binomiale verdeling

$$P(\underset{\sim}{x} = x \mid H_0) = C_x^n \cdot 2^{-n}.$$

Luidt de alternatieve hypothese, dat de mediaan van de verschillen ongelijk is aan nul, dan is $P(+) \neq \frac{1}{2}$ en wordt tweezijdig getoetst. Voor de alternatieve hypothese, dat de mediaan der verschillen kleiner is dan nul, is $P(+) < \frac{1}{2}$ en dan vindt links éénzijdige toetsing plaats. Voor de alternatieve hypothese, dat de mediaan der verschillen groter is dan nul, is $P(+) > \frac{1}{2}$, zodat rechts éénzijdig wordt getoetst.

9.4.4. UITVOERING VAN DE TOETSING

Hiervoor kunnen wij verwijzen naar 6.2, waarin de toets voor de hypothese $H_0 : P = \frac{1}{2}$ uitvoerig is behandeld. Wij volstaan hier met te memoreren dat men:

a. De overschrijdingskans van x_0 exact kan bepalen (bijv. uit tabellen met binomiale kansen).

b. Voor $n \leqslant 100$ gebruik kan maken van tabel I, waarin voor tweezijdige toetsing en de drempelwaarden 0,01, 0,02, 0,05 en 0,10 de linker kritieke waarden x_L van de grootheid $\underset{\sim}{x}$ zijn opgenomen ($x_R = n - x_L$). Bij halvering van deze drempelwaarden kan de tabel ook voor éénzijdige toetsing worden gebruikt.

c. Voor $n > 30$ de normale benadering via (6.5) of (6.6) kan toepassen. Omdat tabel I ter beschikking staat zal men hiervan gewoonlijk slechts gebruik maken, wanneer $n > 100$ is.

Bij voorbeeld 9.6 is $x_0 = 16$. Er wordt tweezijdig getoetst en wij onderstellen dat $\alpha = 0{,}05$ wordt gekozen. In tabel I vinden wij bij $n = 17$ en $\alpha = 0{,}05$ dat $x_L = 4$ zodat $x_R = 17 - 4 = 13$. Daar x_0 deze kritieke waarde overschrijdt wordt H_0 verworpen ten gunste van het alternatief dat bij stof A een sterkere prikkeling optreedt dan bij B.

Opmerking: De exacte tweezijdige overschrijdingskans van x_0 kan in dit geval nog snel worden berekend, daar:

$$P_D(\underset{\sim}{x} = 16) = 2\,|P(\underset{\sim}{x} = 17) + P(\underset{\sim}{x} = 16)| = 2\,[C_{17}^{17} \cdot 2^{-17} + C_{16}^{17} \cdot 2^{-17}]$$
$$= 2(1 + 17)2^{-17} = 9 \times 2^{-15} = 0{,}00027.$$

Door MACKINNON (1959) is voor de tekentoets een compacte tabel met kritieke waarden voor $n < 1000$ opgesteld. Deze tabel is — voorzien van voorbeelden — ook gepubliceerd in Statistica Neerlandica (1960), blz. 161—164.

Onbesliste gevallen. Deze ontstaan als het verschil tussen de gepaarde waarnemingen nul is of een zo kleine omvang heeft, dat er ten gevolge van de beperkte nauwkeurigheid van de waarnemingsmethode geen uitspraak 'plus' of 'min' kan plaatsvinden. In het algemeen geldt dat

men deze gevallen het beste buiten beschouwing kan laten: de toetsing wordt dan uitgevoerd met het aantal paren, waarbij wel een uitspraak mogelijk is. Deze methode is beter (d.w.z. leidt tot een toets met een groter onderscheidingsvermogen) dan die, waarbij men de helft van de onbesliste gevallen positief en de helft negatief telt.

Er kunnen zich echter situaties voordoen waarbij het aantal onbesliste gevallen relatief groot en van bijzonder karakter is, zoals bijv. het aantal personen in de categorie 'geen mening' bij een opinie-onderzoek. Het kan dan noodzakelijk zijn wèl rekening met deze gevallen te houden. Zie hiervoor HEMELRIJK & VAN ELTEREN [(1953) 7.4 t/m 7.6)].

9.4.5. TOEPASSING OP KWANTITATIEVE WAARNEMINGEN

Wanneer de (relatieve) grootte van de verschillen bekend is, leidt het toepassen van de tekentoets tot een verlies aan informatie. Men kan dan beter de — eveneens verdelingsvrije — rangtekentoets gebruiken.

9.4.6. OPGAVEN

9.9. Uit HEMELRIJK & VAN ELTEREN (1953).
Een dierenhandelaar krijgt een offerte van een nieuw soort kattenbrood (B), dat iets duurder is dan het kattenbrood, dat hij tevoren steeds verkocht heeft (soort A). Om te beslissen of hij soort B zal kopen neemt hij een proef met 30 katten, die hij tegelijkertijd een bakje A en een bakje B voorzet. Hij wil soort B alleen kopen als blijkt, dat B beter is dan A. Bij de proef bleken 10 katten A te prefereren boven B, terwijl 20 katten meer B aten. Wat zal volgens de tekentoets de beslissing van de handelaar zijn, indien hij een kans 0,05 wil riskeren om de offerte te accepteren als A niet slechter is dan B?

9.10. Uit N.T.v.G. (1957), 1912. 80 patiënten, die een oordeel uitspraken over het resultaat van een behandeling, gaven de volgende waarderingen: wel geholpen — 48, niet geholpen — 21, zieker gemaakt — 2, geen oordeel — 9. Wat kunt u hieruit concluderen?

9.11. Bepaal het aantal plussen, dat bij tweezijdige toepassing van de tekentoets nodig is, om juist tot het verwerpen van H_0 over te gaan (met $\alpha = 0{,}05$) voor de volgende waarden van n:
a. $n = 150$ *b.* $n = 250$ *c.* $n = 500$.

9.5. De toets van Cochran

9.5.1. TOEPASSING

Deze toets kan worden toegepast als men beschikt over k verwante waarnemingsreeksen (zie 8.2.1), waarbij de waarnemingen een dichotomie (met de kenmerken A en B) betreffen. De nulhypothese luidt dat per element de kans op de uitkomst A constant is. De toets is gevoelig voor een overwegend overeenstemmend verloop van de kansen op kenmerk A van de n elementen [1].

[1] Wil men onderzoeken of deze kansen een overwegend stijgend of dalend verloop vertonen, dán kan men gebruik maken van de toets in par. 12.3.

Tabel 9.6. Koude rillingen als bloedtransfusie-reactie bij 4 verschillende manieren van bloedvoorbereiding: 0 = geen rillingen, 1 = rillingen

Patiënt nr.	Bereidingswijze				R_i	R_i^2
	A	*B*	*C*	*D*		
1	1	1	0	0	2	4
2	1	0	0	0	1	1
3	0	1	1	0	2	4
4	1	1	0	0	2	4
5	1	1	1	0	3	9
6	1	1	0	1	3	9
7	1	0	1	0	2	4
8	1	1	1	1	4	16
9	1	1	0	1	3	9
10	0	0	0	0	0	0
11	1	0	1	0	2	4
12	1	0	0	0	1	1
13	0	0	0	0	0	0
14	1	1	1	0	3	9
15	1	0	0	0	1	1
K_j	12	8	6	3	$29 = \Sigma R_i = \Sigma K_j = S$	$75 = \Sigma R^2$
K_j^2	144	64	36	9	$253 = \Sigma K^2$	$k = 4$

$$Q_0 = \frac{3[4 \times 253 - (29)^2]}{4 \times 29 - 75} = \frac{513}{41} = 12{,}51$$

Voorbeeld

9.7. De (fictieve) gegevens in tabel 9.6 betreffen 15 patiënten, waarop vier bloedtransfusies werden toegepast met bloed dat op vier verschillende manieren, *A*, *B*, *C* en *D*, was voorbereid. Van elke patiënt werd genoteerd of bij de transfusie koude rillingen optraden. Men wil de hypothese toetsen, dat het optreden van deze bloedtransfusie-reactie onafhankelijk is van de wijze van voorbereiding van het toegediende bloed.

9.5.2. DE TOETSINGSGROOTHEID EN HAAR VERDELING ONDER H_0

Alle waarnemingen in de ene categorie worden met 1, die in de andere met 0 aangeduid. COCHRAN kiest als toetsingsgrootheid[1]:

$$(9.11) \qquad Q = \frac{(k-1)\,\Sigma(K_j - \overline{K})^2}{\Sigma R_i - \Sigma R_i^2/k} = \frac{(k-1)\,\{k \,.\, \Sigma K^2 - (S)^2\}}{k \cdot S - \Sigma R^2}.$$

De betekenis van de symbolen in deze formule blijkt uit tabel 9.6. Onder deze tabel is Q_0 voor het voorbeeld berekend.

De verdeling van Q onder H_0 kan worden benaderd door de χ^2-verdeling met $k - 1$ vrijheidsgraden, mits n niet te klein is.

[1] De keuze van deze toetsingsgrootheid wordt later besproken (zie 10.4.5).

9.5.3. UITVOERING VAN DE TOETSING

Er dient rechts éénzijdig te worden getoetst, omdat $Q = 0$ voor $K_1 = K_2 = \ldots = K_k$ [zie (9.11)] en hogere waarden aanneemt, naarmate de waarden van K verder uiteenlopen. H_0 wordt zodoende bij een drempelwaarde α verworpen, indien $Q_0 \geqslant \chi^2_{1-\alpha}$ ($\nu = k - 1$). Laatstgenoemde kritieke waarde kan uit tabel D worden afgelezen.

Kiest men bij de toetsing van de uitkomsten in tabel 9.6 een onbetrouwbaarheid 0,05, dan is de kritieke waarde $\chi^2_{0,95}(\nu = 3) = 7{,}815$. De waargenomen $Q_0 = 12{,}51$ is aanzienlijk groter (heeft een overschrijdingskans die kleiner is dan 0,01), zodat H_0 wordt verworpen.

Onbesliste gevallen. Deze gevallen, waarbij in elke rij uitsluitend nullen of enen voorkomen, verschaffen t.a.v. het verschil tussen de behandelingen geen aanwijzingen. Men kan ze dan ook direct weglaten, zonder dat daardoor Q_0 verandert. In tabel 9.6 komen 2 patiënten met 0000 en 1 patiënt met 1111 voor. Bereken Q_0 voor de overblijvende 12 patiënten.

Voor $k = 2$ beschikt men over gepaarde waarnemingen. Men kan aan de uitkomst 01 een plus en aan de uitkomst 10 een min toekennen (00 en 11 worden weggelaten) en daarna de tekentoets toepassen. Maakt men in dit geval gebruik van formule (9.11), dan heeft de verkregen Q_0 dezelfde waarde als de T_0^2, die men bij toepassing van de normale benadering van de tekentoets (zonder continuïteitscorrectie) verkrijgt.

9.6. De toets van Fisher (methode der 2 × 2 tabel)

9.6.1. TOEPASSING

Deze toets kan worden gebruikt als men de juistheid wil onderzoeken van de nulhypothese dat twee aselecte steekproeven uit dezelfde dichotomie stammen.

Voorbeelden

9.8. Een groep van twaalf muizen wordt aselect in twee even grote steekproeven, 1 en 2, verdeeld. De dieren in steekproef 1 worden intraveneus ingespoten met een bepaalde dosis van stof X en de dieren in steekproef 2 krijgen een zelfde dosis van stof Y. Na verloop van zekere tijd wordt vastgesteld, hoeveel dieren in elke steekproef nog in leven zijn. De uitkomsten zijn in onderstaande 2 × 2 tabel (vierveldentabel) vastgelegd:

		Steekproeven 1 Stof X	Steekproeven 2 Stof Y	
Uitkomst	Gestorven	1	6	7
	Levend	5	0	5
		6	6	12

De te toetsen hypothese luidt dat de twee 'behandelingen' dezelfde sterftekans bezitten.

9.9. Bij een onderzoek heeft men uit de kinderen van 7, 8 en 9 jaar van een gemeente aselect een steekproef van 100 jongens en van 100 meisjes getrokken. In deze steekproeven worden de volgende struma-frequenties geconstateerd:

	Steekproeven Jongens	Meisjes	
Struma	31	49	80
Geen struma	69	51	120
	100	100	200

De te toetsen hypothese is, dat struma bij jongens en meisjes even frequent voorkomt.

9.6.2. DE TOETSINGSGROOTHEID

De vierveldentabel wordt in het algemeen van de volgende symbolen voorzien:

		Steekproeven 1	2	
Kenmerk	A	a	c	$r = a + c$
	B	b	d	$s = b + d$
		n_1	n_2	$n = a + b + c + d$

Als toetsingsgrootheid wordt gekozen a, het aantal elementen met kenmerk A in steekproef 1.

9.6.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

Noem de kansen op de uitkomst A in de populaties waaruit de steekproeven 1 en 2 afkomstig zijn resp. P_1 en P_2. Bij de gegeven *randfrequenties* n_1, n_2, r en s volgt onder de hypothese $H_0 : P_1 = P_2$ de grootheid $\underline{a}$ de hypergeometrische verdeling

$$(9.12) \qquad P(\underline{a} = a) = \frac{C_a^r \cdot C_b^s}{C_{n_1}^n} = \frac{\dfrac{r! \cdot s!}{a!c!\ b!d!}}{\dfrac{n!}{n_1!n_2!}} = \frac{r!s!n_1!n_2!}{n!} \cdot \frac{1}{a!b!c!d!}$$

met gemiddelde

$$(9.13) \qquad \mu_a = \frac{n_1 r}{n}$$

en variantie

$$(9.14) \qquad \sigma_a^2 = \frac{n_1 n_2 r s}{n^2(n-1)}.$$

Men kan de toetsing van H_0 uitvoeren met één van de drie volgende alternatieve hypothesen:

1. $P_1 < P_2$. Links éénzijdige toetsing, H_0 wordt verworpen indien $P_L(a_0) \leqslant \alpha$ is.
2. $P_1 \neq P_2$. Tweezijdige toetsing. H_0 wordt verworpen indien $P_D(a_0) \leqslant \alpha$ is.
3. $P_1 > P_2$. Rechts éénzijdige toetsing. H_0 wordt verworpen indien $P_R(a_0) \leqslant \alpha$ is.

9.6.4. EXACTE TOETSING

Met (9.12) kan de kansverdeling van $\underline{a}$ onder H_0 worden berekend. In tabel 9.7 is deze berekening voor voorbeeld 9.8 uitgevoerd.

Tabel 9.7. Verdeling van de toetsingsgrootheid a onder H_0 voor voorbeeld 9.8

2×2 tabel	a	Berekening	$P(\underline{a} = a)$
$\frac{1 \mid 6}{5 \mid 0}$	1	$\frac{6!6!7!5!}{12!} \cdot \frac{1}{1!5!6!0!} = \frac{1}{132}$	$= 0{,}0076$
$\frac{2 \mid 5}{4 \mid 1}$	2	$\frac{6!6!7!5!}{12!} \cdot \frac{1}{2!4!5!1!} = \frac{15}{132}$	$= 0{,}1136$
$\frac{3 \mid 4}{3 \mid 2}$	3	$\frac{6!6!7!5!}{12!} \cdot \frac{1}{3!3!4!2!} = \frac{50}{132}$	$= 0{,}3788$
$\frac{4 \mid 3}{2 \mid 3}$	4	$\frac{6!6!7!5!}{12!} \cdot \frac{1}{4!2!3!3!} = \frac{50}{132}$	$= 0{,}3788$
$\frac{5 \mid 2}{1 \mid 4}$	5	$\frac{6!6!7!5!}{12!} \cdot \frac{1}{5!1!2!4!} = \frac{15}{132}$	$= 0{,}1136$
$\frac{6 \mid 1}{0 \mid 5}$	6	$\frac{6!6!7!5!}{12!} \cdot \frac{1}{6!0!1!5!} = \frac{1}{132}$	$= 0{,}0076$
			1,0000

Uit de tabel blijkt dat de tweezijdige overschrijdingskans van de gevonden uitkomst, $a_0 = 1$, gelijk is aan $P(\underline{a} = 1) + P(\underline{a} = 6) = 2(0{,}0076) = 0{,}0152$. H_0 kan bij tweezijdige toetsing bij een 5% drempel worden verworpen: de conclusie is dat de sterftekans bij stof Y hoger is dan die bij stof X.

Wij geven nog een tweede voorbeeld van exacte toetsing, maar nu met steekproeven van ongelijke omvang.

Voorbeeld

9.10. Onderstel dat twee aselecte steekproeven de nevenstaande aantallen elementen met een zekere afwijking A vertonen. De hypothese die men wenst te toetsen luidt weer, dat deze steekproeven uit dezelfde dichotomie stammen; men kiest $\alpha = 0{,}05$.

	Steekproeven 1	2	
A	1	8	9
Niet-A	7	4	11
	8	12	20

Onder H_0 volgt de grootheid $\underline{a}$ (het aantal elementen met kenmerk A in steekproef 1) de volgende verdeling:

$$P(\underline{a} = a \mid H_0) = \frac{9!11!8!12!}{20!} \cdot \frac{1}{a!b!c!d!} \quad (a + c = 9).$$

De mogelijke uitkomsten zijn:

0 | 9	1 | 8	2 | 7	3 | 6	4 | 5	5 | 4	6 | 3	7 | 2	8 | 1
8 | 3	7 | 4	6 | 5	5 | 6	4 | 7	3 | 8	2 | 9	1 | 10	0 | 11

Met (9.12) kan de verdeling van $\underline{a}$ onder H_0 worden berekend. Deze verdeling:

a	0	1	2	3	4	5	6	7	8	Totaal
$P(a)$	0,001	0,023	0,132	0,309	0,331	0,165	0,037	0,003	0,000	1,001

is niet symmetrisch. Ook in dit geval wordt tweezijdig getoetst. Wij zoeken dus alle waarden van $\underline{a}$ bijeen, die een hoogstens even grote kans bezitten als de waargenomen $a_0 = 1$; dit zijn de waarden 0, 1, 7 en 8. De tweezijdige overschrijdingskans van a_0 is gelijk aan de som van de bij deze waarden behorende kansen (zie 3.6):

$$P_D(1) = 0{,}001 + 0{,}023 + 0{,}003 + 0{,}000 = 0{,}027.$$

Deze kans is kleiner dan de gekozen onbetrouwbaarheid 0,05, zodat H_0 wordt verworpen. De conclusie luidt dat de populatie, waaruit steekproef 2 afkomstig is, een hogere fractie A's bevat dan de populatie van steekproef 1.

Het berekenen van exacte kansen is bij een wat grotere steekproefomvang dan in onze voorbeelden een tijdrovende bezigheid. Men kan echter gebruik maken van tabellen:

a. FINNEY (1948) heeft voor $n \leqslant 30$ tabellen samengesteld met exacte éénzijdige kansen, die ook in PEARSON en HARTLEY (1954) en in enigszins gewijzigde vorm in SIEGEL (1956) zijn opgenomen. LATSCHA (1950) heeft deze tabellen tot $n \leqslant 40$ uitgebreid.

b. MAINLAND c.s. (1956) geven tabellen waarin men voor $n \leqslant 40$ en $\alpha = 0{,}05$ kritieke waarden bij tweezijdige toetsing kan aflezen. Deze tabellen staan ook in RÜMKE en VAN EEDEN (1961).

c. De tabellenverzameling van GEIGY (1960) geeft voor $n \leqslant 60$ kritieke waarden bij tweezijdige toetsing met $\alpha = 0{,}20$, 0,10, 0,05, 0,02, 0,01 en 0,002, die dus ook voor éénzijdige toetsing met $\alpha = 0{,}10$, 0,05, 0,01 en 0,001 kunnen worden gebruikt.

In 9.6.5 en 9.6.6 behandelen wij een normale benadering, die men reeds bij $n \geqslant 30$ kan toepassen, mits de steekproeven niet te sterk in omvang verschillen en indien de fractie r/n niet te dicht bij 0 of 1 ligt. In 9.6.7 geven wij aan hoe men voor niet te kleine steekproeven een benaderend betrouwbaarheidsinterval voor het verschil tussen de parameters P_1 en P_2 kan berekenen. Tenslotte volgt in 9.6.8 een toets voor de hypothese $P_1 = P_2$ die men kan gebruiken voor kleine kansen en grote steekproeven van (ongeveer) dezelfde omvang.

9.6.5. DE NORMALE BENADERING

Verdeling onder H_0. De grootheid $\underline{a}$ volgt onder H_0 een hypergeometrische verdeling, die voor niet te kleine n en een fractie r/n die niet te dicht bij 0 of 1 ligt, door een normale verdeling kan worden benaderd. Uit (9.13) en (9.14) volgt dat dan onder H_0 de grootheid

$$\underline{T} = \frac{\underline{a} - \mu_a}{\sigma_a} \tag{9.15.1}$$

bij benadering standaardnormaal verdeeld is. Men kan deze formule herleiden tot

$$\underline{T} = \frac{(ad - bc)\sqrt{n-1}}{\sqrt{n_1 n_2 r s}}. \tag{9.15.2}$$

Uitvoering van de toetsing. Men kan de normale benadering verbeteren, door het aanbrengen van een continuïteitscorrectie, die in (9.15.1) gelijk is aan 0,5 en in (9.15.2) gelijk is aan $n/2$. Deze correctie blijkt gewoonlijk iets te sterk te zijn, zodat men een te kleine waarde van $\underline{T}$ (en dus een te grote overschrijdingskans, vergeleken met de exact berekende) verkrijgt. Men kan dit effect compenseren door de waarde van $\underline{T}$ met $\sqrt{n/(n-1)}$ te vermenigvuldigen.

Men berekent dus

$$T_0 = \frac{(|\, ad - bc \,| - n/2)\sqrt{n}}{\sqrt{n_1 n_2 r s}} \tag{9.16.1}$$

en $P_D(T_0)$ bij tweezijdige toetsing,

$$T_0 = \frac{(ad - bc + n/2)\sqrt{n}}{\sqrt{n_1 n_2 r s}} \tag{9.16.2}$$

en $P_L(T_0)$ bij links éénzijdige toetsing,

of

$$T_0 = \frac{(ad - bc - n/2)\sqrt{n}}{\sqrt{n_1 n_2 r s}} \tag{9.16.3}$$

en $P_R(T_0)$ bij rechts éénzijdige toetsing,

en verwerpt H_0 indien de gevonden overschrijdingskans $\leqslant \alpha$ is.

Opmerking. In het algemeen kan men stellen, dat de normale benadering kan worden toegepast indien de verwachte frequenties onder H_0 (berekend volgens de in tabel 9.8 aangegeven methode) ten minste gelijk zijn aan 5.

Bij voorbeeld 9.9 wordt tweezijdig getoetst. Met *(9.16.1)* vindt men

$$T_0 = \frac{(|\, 31 \times 51 - 49 \times 69 \,| - 200/2)\sqrt{200}}{\sqrt{100 \times 100 \times 80 \times 120}} = 2{,}454 = 2{,}45,$$

met een tweezijdige overschrijdingskans die volgens tabel A gelijk is aan 0,0142. Bij toetsing met $\alpha = 0{,}05$ wordt dus de nulhypothese verworpen en men concludeert dat in de gemeente waar het onderzoek is verricht struma bij meisjes meer voorkomt dan bij jongens.

Men kan ook als volgt redeneren. Onder H_0 is $P_1 = P_2 = P$. Veronderstel dat P niet te dicht bij 0 of 1 ligt en dat n_1 en n_2 niet te klein zijn. Onder H_0 volgt dan de grootheid p_1 een bij benadering normale verdeling met gemiddelde P en variantie PQ/n_1, terwijl de grootheid p_2 een bij benadering normale verdeling volgt met gemiddelde P en variantie PQ/n_2 (zie 4.3.1 en 5.4.2). Hieruit volgt dat onder H_0 de grootheid $p_1 - p_2$ eveneens bij benadering normaal verdeeld is, met als gemiddelde $P - P = 0$ en als variantie $PQ/n_1 + PQ/n_2 = PQ(n_1 + n_2)/n_1n_2 = PQn/n_1n_2$, zodat de grootheid

$$\underline{T} = \frac{\underline{p}_1 - \underline{p}_2}{\sqrt{PQ}} \sqrt{\frac{n_1n_2}{n}}$$

bij benadering standaardnormaal verdeeld is. Schat men nu P uit de waarnemingen door het gewogen gemiddelde p_0 te berekenen van de waargenomen waarden p_1 en p_2, dus

$$p_0 = \frac{n_1p_1 + n_2p_2}{n_1 + n_2},$$

dan kan men als toetsingsgrootheid nemen

$$\underline{T} = \frac{\underline{p}_1 - \underline{p}_2}{\sqrt{p_0q_0}} \sqrt{\frac{n_1n_2}{n}} \quad (q_0 = 1 - p_0). \tag{9.17}$$

Wanneer men in (9.17) een continuïteitscorrectie ter grootte van $2/n$ aanbrengt, verkrijgt men dezelfde uitkomst als met (9.16). Bij voorbeeld 9.9 vindt men

$$T_0 = \frac{|\,0{,}31 - 0{,}49\,| - 2/200}{\sqrt{0{,}40 \times 0{,}60}} \sqrt{\frac{100 \times 100}{200}} = 2{,}454.$$

9.6.6. DE χ^2-BENADERING

Verdeling onder H_0. Voor elk 'veld' van de vierveldentabel kan men, uitgaande van de randfrequenties, *verwachte frequenties* onder H_0 berekenen. Noemt men de met de waargenomen frequenties a, b, c en d corresponderende verwachte frequenties resp. α, β, γ en δ, dan neemt men

$$\alpha : \beta = \gamma : \delta = r : s,$$

zodat bijv. (zie tabel 9.8A) $\alpha : n_1 = r : n$, $\alpha = n_1r/n$. De bij elk veld behorende verwachte frequentie is dus gelijk aan het product van de bij dat veld behorende randfrequenties, gedeeld door n. In tabel 9.8B zijn de verwachte frequenties bij voorbeeld 9.9 berekend.

Tabel 9.8. Waargenomen en verwachte frequenties bij de 2 × 2 tabel

A. Gebruikte symbolen

Waargenomen frequenties

Kenmerk	Steekproef 1	Steekproef 2	
A	a	c	r
B	b	d	s
	n_1	n_2	n

Verwachte frequenties onder H_0

	Steekproef 1	Steekproef 2	
A	$\alpha = n_1 r/n$	$\gamma = n_2 r/n$	r
B	$\beta = n_1 s/n$	$\delta = n_2 s/n$	s
	n_1	n_2	n

B. Bij voorbeeld 9.9

	J.	M.	
Struma	31	49	80
Geen struma	69	51	120
	100	100	200

	Jongens	Meisjes	
Struma	$100 \times 80/200 = 40$	$80 - 40 = 40$	80
Geen struma	$100 - 40 = 60$	$100 - 40 = 60$	120
	100	100	200

Per veld beschikt men nu over een waargenomen en een verwachte frequentie, zodat men formule (9.6) kan toepassen. Men vindt dan

$$G_0 = \frac{(a_0 - \alpha)^2}{\alpha} + \frac{(b_0 - \beta)^2}{\beta} + \frac{(c_0 - \gamma)^2}{\gamma} + \frac{(d_0 - \delta)^2}{\delta}.$$

Deze formule kan worden vereenvoudigd daar, zoals eenvoudig is in te zien, steeds geldt dat $(a - \alpha)^2 = (b - \beta)^2 = (c - \gamma)^2 = (d - \delta)^2$. Men kan dus schrijven

$$G_0 = (a_0 - \alpha)^2 \left[\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta}\right].$$

Voor niet te kleine waarden van n en verwachte frequenties die ten minste gelijk zijn aan 5 (d.i. als r/n niet te dicht bij 0 of 1 ligt) volgt onder H_0 de grootheid G bij benadering een χ^2-verdeling met 1 vrijheidsgraad (bij gegeven randfrequenties kan slechts van één van de vier velden de frequentie vrij worden ingevuld). De benadering kan worden verbeterd door een continuïteitscorrectie van 0,5 aan te brengen en te nemen

$$(9.18) \qquad G_0 = (|\, a_0 - \alpha \,| - \tfrac{1}{2})^2 \left(\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta}\right).$$

G_0 berekend met (9.18) blijkt gelijk te zijn aan T_0^2 berekend met (9.16). Men kan dus de toets ook uitvoeren door te nemen

$$(9.19) \qquad T_0 = (a_0 - \alpha \pm \tfrac{1}{2}) \sqrt{\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta}},$$

waarbij de continuïteitscorrectie van $-\frac{1}{2}$ of $+\frac{1}{2}$ op de bij formule (9.16) aangegeven manier wordt aangebracht. Deze formule vereist

weinig rekenwerk, daar men slechts α behoeft te berekenen; β, γ en δ volgen dan door aftrekking van de randfrequenties. De reciproken van de verwachtingen kan men uit een tabel aflezen [zeer geschikt zijn hiervoor BARLOW's TABLES (1954)]. Bij voorbeeld 9.9 vindt men met (9.19)

$$T_0 = (|\,31 - 40\,| - \tfrac{1}{2})\sqrt{\frac{1}{40} + \frac{1}{60} + \frac{1}{40} + \frac{1}{60}} = 2{,}454.$$

Voorbeelden

9.11. Voor de twee steekproeven links in tabel 9.9 wil men de hypothese $P_1 = P_2$ toetsen met als alternatieve hypothese $P_1 > P_2$, zodat rechts éénzijdig wordt getoetst. Men kiest $\alpha = 0{,}05$. Rechts in de tabel zijn de verwachte frequenties berekend (gewoonlijk kan men hierbij volstaan met één decimaal). Daarna is T_0 met (9.16) en met (9.19) berekend.

Tabel 9.9. Gebruik van de benaderingen bij rechts éénzijdige toetsing

Waargenomen frequenties

	Steekproeven 1	2	
A	$a = 15$	$c = 14$	29
B	$b = 5$	$d = 16$	21
	20	30	50

Verwachte frequenties

	Steekproeven 1	2	
A	$\alpha = 11{,}6$	$\gamma = 17{,}4$	29
B	$\beta = 8{,}4$	$\delta = 12{,}6$	21
	20	30	50

(9.16) $$T_0 = \frac{(15 \times 16 - 5 \times 14 - 50/2)\sqrt{50}}{\sqrt{20 \times 30 \times 29 \times 21}} = 1{,}696$$

(9.19) $$T_0 = (15 - 11{,}6 - 0{,}5)\sqrt{\frac{1}{11{,}6} + \frac{1}{8{,}4} + \frac{1}{17{,}4} + \frac{1}{12{,}6}} = 1{,}696$$

Uit tabel A blijkt dat $P_R(1{,}70) = 0{,}045$, zodat de nulhypothese wordt verworpen ten gunste van het alternatief, $P_1 > P_2$.

9.12. Als de twee steekproeven dezelfde omvang hebben, is de verdeling van de grootheid a symmetrisch. De normale benadering is dan meestal bij kleine waarden van n bevredigend, mits r/n niet te ver van 0,5 ligt. Berekent men bijv. voor de nevenstaande 2×2 tabel de exacte linkse kans van $a_0 = 2$ dan vindt men met (9.12)

	Steekproeven 1	2	
+	2	7	9
−	8	3	11
	10	10	20

$$\frac{10!10!9!11!}{20!}\left[\frac{1}{0!10!9!11!} + \frac{1}{1!9!8!2!} + \frac{1}{2!8!7!3!}\right] = 0{,}0349.$$

Met (9.16) vindt men

$$T_0 = \frac{(2 \times 3 - 8 \times 7 + 20/2)\sqrt{20}}{\sqrt{10 \times 10 \times 9 \times 11}} = -1{,}80,$$

met een linkse kans van ongeveer 0,0359.

Opmerking. Bij gebruik van formule (9.15) met continuïteitscorrectie vindt men $T_0 = -1{,}75$ met een (te grote) linkse kans van ongeveer 0,04. Past men deze formule toe zonder continuïteitscorrectie dan vindt men $T_0 = -2{,}19$ met een (veel te kleine) linkse kans van 0,014.

9.6.7. BETROUWBAARHEIDSGRENZEN VAN $P_1 - P_2$

Veronderstel dat steekproef 1 van de omvang n_1 afkomstig is uit een populatie waarin de kans op A gelijk is aan P_1 en dat steekproef 2 van de omvang n_2 stamt uit een populatie waarin deze kans P_2 bedraagt. De fractie p_1 volgt dan een bij benadering normale verdeling met gemiddelde P_1 en variantie P_1Q_1/n_1. De fractie p_2 volgt een bij benadering normale verdeling met gemiddelde P_2 en variantie P_2Q_2/n_2. Wanneer n_1 en n_2 niet te klein zijn en P_1 en P_2 niet te dicht bij 0 of 1 liggen kan men uit de twee waargenomen fracties p_1 en p_2 een *benaderend* $100(1-\alpha)\%$ betrouwbaarheidsinterval van $P_1 - P_2$ verkrijgen door te nemen

$$(9.20)\quad (p_1 - p_2) + T_{\frac{1}{2}\alpha}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}} < P_1 - P_2 < (p_1 - p_2) + T_{1-\frac{1}{2}\alpha}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}.$$

Met deze formule vindt men bij voorbeeld 9.9 het volgende benaderende 95% betrouwbaarheidsinterval ($\alpha = 0{,}05$, $T_{0{,}025} = -T_{0{,}975} = -1{,}96$, zie tabel B)

$$(0{,}49 - 0{,}31) \pm 1{,}96\sqrt{\frac{0{,}31 \times 0{,}69}{100} + \frac{0{,}49 \times 0{,}51}{100}},$$

$$0{,}05 < P_2 - P_1 < 0{,}31.$$

9.6.8. TOETSING VOOR KLEINE KANSEN BIJ EVEN GROTE STEEKPROEVEN

Door KLERK-GROBBEN & PRINS (1954) is een zeer eenvoudige toets voor de hypothese $H_0 : P_1 = P_2 = P$ beschreven, die kan worden uitgevoerd als men te doen heeft met kleine kansen en even grote steekproeven.

Onderstel bijv., dat men de hypothese wil toetsen dat een vrij zeldzame afwijking in twee verschillende populaties evenveel voorkomt. Men trekt uit elke populatie aselect een steekproef van 1000 personen en deze steekproeven leveren op:

	1	2	
Afwijking	$a = 6$	$c = 20$	$26 = a + c = m$
Geen afwijking	994	980	
	1000	1000	$2000 = 2n$

Onder H_0 volgt de grootheid $\underline{a}$ bij benadering de binomiale verdeling

$$P(\underline{a} = a) = C_a^m \cdot 2^{-m},$$

mits de omvang der steekproeven n veel groter dan m is. De grootte van n heeft overigens op de toets geen invloed. Deze toetsing kan een-

voudig met tabel I worden uitgevoerd. Uit deze tabel blijkt dat voor $m = 26$ de tweezijdige overschrijdingskans van $a_0 = 6$ hoogstens gelijk is aan 0,02. H_0 kan dus met een onbetrouwbaarheid $\leqslant 0{,}02$ worden verworpen.

Hoewel de toets strikt genomen slechts geldig is, indien beide steekproeven dezelfde omvang n bezitten, geeft hij een goede benadering voor het geval dat $n_1 \neq n_2$ als n_1 en n_2 groot zijn en slechts weinig verschillen (bijv. $n_1 = 1000$ en $n_2 = 985$).

Voor verdere bijzonderheden verwijzen wij naar bovengenoemd artikel, waarin ook het onderscheidingsvermogen van de toets en de keuze van n wordt besproken.

9.6.9. OPGAVEN

	Steekproef 1	Steekproef 2	
A	2	5	7
B	7	2	9
	9	7	16

9.12. Toets exact de hypothese, dat nevenstaande steekproeven uit dezelfde dichotomie stammen. Kies $\alpha = 0{,}05$.

	Steekproef 1	Steekproef 2	
+	10	3	13
—	2	5	7
	12	8	20

9.13. Toets de hypothese dat nevenstaande steekproeven uit dezelfde dichotomie afkomstig zijn:

a. exact,
b. met de χ^2-benadering zonder continuïteitscorrectie,
c. idem met deze correctie.

9.14. Uit: N.T.v.G. 98(1954), 2208.

Tabel I

Indirect ingeënt met gele-koorts virus	Hersen-suspensie	Aantal	Over-levende	Over-levende in pct
Dengue-muizen, 3 × intraperitoneaal	5 pct	18	11	61,1
Controle-muizen	5 pct	10	3	30

'Verschil tussen percentage overlevende dengue-muizen en controle-muizen: 31,1%.'
Verschillen deze percentages significant bij een 5% drempel?

9.15. Uit H. J. SLUITER, Cor pulmonale, Diss., Groningen, 1955.
Bij 8 van de 33 emphyseempatiënten zonder decompensatio cordis (24%) en 13 van de 22 emphyseempatiënten met een decompensatie van het rechterhart (59%) werden trommelstokvingers waargenomen. Is dit verschil significant bij een 5% drempel?

9.16. Uit de kursus: Summarizing Experimental Data, Columbia University, U.S.A. The observations in the following table are presented in a paper on

the treatment of erysipelas in children (J. Am. Med. Ass., March 10, 1934) as evidence that ultraviolet therapy gives better results than serum therapy in infants.

Results of treatment of children under one year of age

Treatment	Lived	Died	Total cases	Mortality percentage
Ultraviolet	12	7	19	36,8
Serum	23	32	55	58,2

a. What questions would you ask relative to the comparability of the mortality percentages for the two forms of treatment?

b. Test the statistical significance ($\alpha = 0{,}05$) of the difference in mortality under the two treatments.

9.17. Uit: N.T.v.G., 100 (1956), 2769—2783.

Tabel I. Primaire sterfte bij shunt-operaties wegens portale hypertensie

Schrijver en datum	Aandoening	Aantal gevallen		Letaliteit			
		porta-cava shunt	splenorenale shunt	Aantal		Procent	
				PCS	SRS	PCS	SRS
Blakemore 1952	Cirrose	65	54	13	11	20	20
	Extra-hepatisch block	4	27	0	1	0	4
Linton 1951	Cirrose	11	19	3	1	27	5
	Extra-hepatisch block	4	27	0	6	0	22
Child 1953	Cirrose	31	1	4	0	13	0
	Extra-hepatisch block	—	3	—	0	—	0
Comb. v/d geg. van 3 schrijvers	Cirrose	107	74	20	12	19	16
	Extra-hepatisch block	8	57	0	7	0	12

De auteur merkt op: 'Men ziet uit bovenstaande cijfers, dat onafhankelijk van de toegepaste methode de sterfte bij levercirrose aanzienlijk hoger is dan die bij extra-hepatisch block, een verschil dat voor de hand ligt als men bedenkt, dat de lijder aan levercirrose wel steeds een uiterst dubieus risico vormt voor iedere grote operatie (lever-insufficiëntie!)' Kunt u zich met deze conclusie verenigen?

9.7. De χ^2-toets op de $2 \times r$ tabel

9.7.1. TOEPASSING

Deze methode kan worden gebruikt voor het toetsen van de nulhypothese dat twee aselecte steekproeven uit dezelfde populatie met r categorieën afkomstig zijn.[1]

[1] Zie voor een geordend categorisch systeem ook hoofdstuk 21.

Voorbeeld

9.13. KROKFORS en KINNUN (Br. Med. J. (1954), 1305) geven de links in tabel 9.10 opgenomen verdelingen van de bloedgroepen bij vrouwen met gynaecologisch carcinoom (groep 1) en gezonde controles (groep 2).

Tabel 9.10. Een 2×4 tabel

Bloed-groep	f_{ij} = waarg. freq.		Tot.	e_{ij} = verwachte freq.		$f_{ij} - e_{ij}$		$(f_{ij} - e_{ij})^2/e_{ij}$	
	Groep 1	Groep 2		1	2	1	2	1	2
O	98	113	211	107,2	103,8	— 9,2	+ 9,2	0,7896	0,8154
A	139	110	249	126,6	122,4	+12,4	—12,4	1,2145	1,2562
B	48	57	105	53,3	51,7	— 5,3	+ 5,3	0,5270	0,5433
AB	25	20	45	22,9	22,1	+ 2,1	— 2,1	0,1926	0,1995
Totaal	310	300	610	310,0	300,0	0	0	2,7237	2,8144
								2,8144	
								5,5381	$= G_0(3)$

Wij veronderstellen dat de groepen zo zijn samengesteld, dat zij als aselecte steekproeven uit hun respectieve populaties kunnen worden beschouwd. De nulhypothese luidt dat de bloedgroepenverdelingen van deze populaties overeenstemmen.

9.7.2. DE TOETSINGSGROOTHEID

De waargenomen frequenties vormen een zg. $2 \times r$ tabel, d.w.z. er zijn twee kolommen (voor de twee steekproeven) en r rijen (het categorisch systeem waarop de waarnemingen betrekking hebben). Deze tabel bevat $2r$ velden en evenals bij de 2×2 tabel kan de verwachte frequentie van een veld (d.i. de frequentie onder H_0) worden berekend door het product van de bij dat veld behorende randfrequenties te nemen en dit te delen door n (het totale aantal waarnemingen). Als toetsingsgrootheid kan dan dienen

$$G = \sum_{i=1}^{r} \sum_{j=1}^{2} \frac{(f_{ij} - e_{ij})^2}{e_{ij}}, \tag{9.21.1}$$

waarin f_{ij} de waargenomen en e_{ij} de verwachte frequentie van het veld op rij i en in kolom j is.

De berekening van G_0 is voor het voorbeeld rechts in tabel 9.10 uitgevoerd en verloopt als volgt: 1. Bereken de verwachte frequenties in één kolom. Die van de tweede kolom vindt men door aftrekking van de rij-totalen. 2. Bepaal de verschillen $f - e$. De twee verschillen op één rij zijn gelijk, doch hebben een tegengesteld teken. 3. Kwadrateer ieder verschil en deel het kwadraat door de bijbehorende verwachting (dit is in tabel 9.10 gedaan. Men kan ook per rij het kwadraat van het rij-verschil vermenigvuldigen met de som van de

reciproken van de verwachte frequenties van die rij). 4. G_0 is de som van de bij (3) verkregen uitkomsten. Voor het voorbeeld is $G_0(3)=5{,}54$.

Verkorte berekening. Als men voor de waargenomen frequenties van de $2\times r$ tabel de in tabel 9.11 aangegeven symbolen gebruikt, kan men G_0 ook berekenen volgens

$$G_0 = \frac{n^2}{n_1 n_2}\left[\sum_{i=1} \frac{a_i^2}{a_i + b_i} - \frac{n_1^2}{n}\right]. \tag{9.21.2}$$

Daar n_1, n_2 en n direct bekend zijn, behoeft men dus slechts voor elke rij de fractie $a_i^2/(a_i + b_i)$ te berekenen.

Tabel 9.11. $2 \times r$ tabel

Categorie	Steekproef		Totaal
	1	2	
1	a_1	b_1	$a_1 + b_2$
2	a_2	b_2	$a_2 + b_2$
:	:	:	:
:	:	:	:
r	a_r	b_r	$a_r + b_r$
Totaal	$\Sigma a_i = n_1$	$\Sigma b_i = n_2$	$n_1 + n_2 = n$

Voor voorbeeld 9.13 is de som van deze fracties:

$$\frac{(98)^2}{211} + \frac{(139)^2}{249} + \frac{(48)^2}{105} + \frac{(25)^2}{45} = 158{,}9428,$$

zodat

$$G_0 = \frac{(610)^2}{310 \times 300}\left[158{,}9428 - \frac{(310)^2}{610}\right] = 5{,}61.$$

(Het geringe verschil met de uitkomst in tabel 9.10 is een gevolg van afrondingsfoutjes.)

Het nadeel van deze rekenmethode is echter, dat men niet, zoals in tabel 9.10, categorie na categorie de verschillen tussen de waargenomen en verwachte frequenties en de bijdragen tot G_0 kan inspecteren.

9.7.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

Onder de hypothese dat beide steekproeven uit dezelfde populatie stammen volgt de grootheid $\underline{G}$ bij benadering een χ^2-verdeling met $r - 1$ vrijheidsgraden (bij gegeven randfrequenties kunnen $r - 1$ velden vrij worden ingevuld). De benadering is goed als n niet te klein is en geen verwachtingen kleiner dan 5 voorkomen. Wil men met een wat grovere benadering genoegen nemen, dan kan bij meer dan twee vrijheidsgraden nog een verwachte frequentie van 2 in één veld worden toegelaten, mits de overige verwachtingen elk ten minste 5 zijn. Voor deze en voor de volgende χ^2-toetsen geldt, dat men een iets

betere benadering verkrijgt door als toetsingsgrootheid $\frac{n-1}{n} G$ te nemen. Deze procedure heeft alleen voor kleine waarden van n praktische betekenis.

9.7.4. UITVOERING VAN DE TOETSING

Er dient rechts éénzijdig te worden getoetst. H_0 wordt bij een onbetrouwbaarheidsdrempel α verworpen, indien

$$G_0 \geqslant \chi^2_{1-\alpha} \; (\nu = r - 1).$$

Bij voorbeeld 9.13 is $G_0 = 5{,}61$ (uitkomst van de verkorte methode, die het nauwkeurigst is) en er zijn 3 vrijheidsgraden. De rechtse kans van deze uitkomst ligt blijkens tabel D tussen 0,1 en 0,2, zodat H_0 niet wordt verworpen.

9.7.5. OPGAVEN

9.18. Een intelligentietest, toegepast op twee aselecte steekproeven, levert nevenstaande uitkomsten op. Toets (met $\alpha = 0{,}05$) de hypothese, dat zij uit dezelfde populatie stammen.

		Testresultaat			
		Slecht	Matig	Goed	
Steekproef	1	14	7	9	30
	2	3	11	16	30
		17	18	25	60

9.19. Uit Int. Review of Vit.Res. (1951), 1—12. Vier aselect samengestelde groepen muizen werden ingespoten met een bacteriëel inoculum (Staphylococcus aureus), dat resp. gekweekt was op voedingsbodems met de vitaminen niacinamide (*NA*), folinezuur (*F*), para-aminobenzoëzuur (*PA*) en B_6 als pyridoxine (B_6). De proef leverde de volgende uitkomsten op:

	NA	*F*	*PA*	B_6
Overlevend	10	9	9	13
Gestorven	30	31	41	27
Totaal	40	40	50	40

Toets (met $\alpha = 0{,}05$) de hypothese dat de waargenomen verschillen tussen de vier groepen op toeval berusten.

9.8. De χ^2-toets op de $k \times 2$ tabel

Deze toets kan worden gebruikt voor het toetsen van de nulhypothese dat k aselecte steekproeven uit dezelfde dichotomie afkomstig zijn.[1]

De toetsingsgrootheid is

$$G = \sum_{i=1}^{2} \sum_{j=1}^{k} \frac{(f_{ij} - e_{ij})^2}{e_{ij}}, \tag{9.22}$$

waarin f_{ij} de waargenomen en e_{ij} de verwachte frequentie van het veld op rij i en in kolom j is.

[1] Zie voor een geordend categorisch systeem ook hoofdstuk 21.

Onder H_0 volgt de grootheid $\underline{G}$ voor niet te kleine steekproeven bij benadering een χ^2-verdeling met $k-1$ vrijheidsgraden. De toetsing geschiedt rechts éénzijdig. H_0 wordt bij een drempelwaarde α verworpen, indien

$$G_0 \geqslant \chi^2_{1-\alpha}\,(\nu = k-1).$$

De uitvoering van de toetsing verloopt analoog aan 9.7.

9.9. De χ^2-toets op de $k \times r$ tabel

Deze toets kan men toepassen voor het toetsen van de nulhypothese dat k aselecte steekproeven uit dezelfde populatie met r categorieën stammen. De toetsingsgrootheid is

$$G = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(f_{ij} - e_{ij})^2}{e_{ij}}. \tag{9.23}$$

Onder H_0 volgt de grootheid G bij benadering een χ^2-verdeling met $(k-1)\,(r-1)$ vrijheidsgraden, indien n niet te klein is en geen te lage verwachtingen voorkomen (zie 9.7.3). De toetsing geschiedt rechts éénzijdig en H_0 wordt bij een drempelwaarde α verworpen voor

$$G_0 \geqslant \chi^2_{1-\alpha}\,[\nu = (k-1)\,(r-1)].$$

Formule (9.23) kan als volgt worden herleid:

$$G_0 = \Sigma\Sigma \frac{f_{ij}^2 - 2f_{ij}e_{ij} + e_{ij}^2}{e_{ij}} = \Sigma\Sigma \frac{f_{ij}^2}{e_{ij}} - 2\Sigma f + \Sigma e = \Sigma \frac{f_{ij}^2}{e_{ij}} - n.$$

Noemt men de bij veld (i, j) behorende randfrequenties f_i en n_j, dan is $e_{ij} = f_i n_j / n$, zodat

$$G_0 = \Sigma\Sigma \frac{n f_{ij}^2}{f_i n_j} - n = n \left[\Sigma\Sigma \frac{f_{ij}^2}{f_i n_j} - 1 \right]. \tag{9.24}$$

De som tussen de grote haken kan worden verkregen door van elk veld het kwadraat van de waargenomen frequentie te delen door het product van de bij dat veld behorende randfrequenties en deze quotiënten over alle kr velden te sommeren. Deze berekening kan zeer goed op een rekenmachine worden uitgevoerd. Een nadeel van deze rekenwijze is echter, dat men niet meer veld na veld het verschil tussen de waargenomen en verwachte frequentie en de daaruit resulterende bijdrage tot G_0 kan bestuderen.

Voorbeeld

9.14. Uit: N.T.v.G. 100 (1956), 104. De uitkomsten in tabel 9.12 betreffen de wijze, waarop de baringspijn tijdens de ontsluitingsperiode werd doorstaan bij drie groepen vrouwen, elk met een verschillende voorbereiding. Hierop is de χ^2-toets toegepast via formule (9.23). De berekening verloopt analoog aan die bij de voorgaande χ^2-toetsen. De berekende G_0 is zeer significant ($P \ll 0{,}001$), zodat de hypothese dat de drie steekproeven uit dezelfde populatie afkomstig zijn wordt verworpen. Aan welke voorwaarden dient de proefopzet te voldoen, wil men hieruit de conclusie mogen trekken dat de psychische prophylaxis nut heeft?

Tabel 9.12. Wijze, waarop de ontsluitingsperiode voor de baring is doorstaan bij drie groepen vrouwen

A = Voorlichting en oefening B = Voorlichting C = Geen voorbereiding

	Waargenomen frequenties f_{ij}				Verwachte frequenties e_{ij}			$f_{ij} - e_{ij}$		
	A	B	C	Tot.	A	B	C	A	B	C
Zeer goed	168	41	114	323	135,3	39,2	148,5	32,7	1,8	—34,5
Goed	100	27	107	234	98,0	28,4	107,6	2,0	—1,4	— 0,6
Matig	50	19	81	150	62,8	18,1	69,1	—12,8	0,9	11,9
Slecht	17	10	66	93	38,9	11,3	42,8	—21,9	—1,3	23,2
Totaal	335	97	368	800	335,0	97,0	368,0	0,0	0,0	0,0

	$(f_{ij} - e_{ij})^2$			$(f_{ij} - e_{ij})^2/e_{ij}$		
	A	B	C	A	B	C
Zeer goed	1069,29	3,24	1190,25	7,9031	0,0826	8,0152
Goed	4,00	1,96	0,36	0,0408	0,0690	0,0033
Matig	163,84	0,81	141,61	2,6089	0,0448	2,0493
Slecht	479,61	1,69	538,24	12,3293	0,1496	12,5757
				22,8821	0,3460	22,6435

$G_0 = 22{,}8821 + 0{,}3460 + 22{,}6435 = 45{,}9$ $\quad \chi^2_{0,999}\,(\nu = 6) = 22{,}457$

9.10. Gebruik van de χ^2-toetsen bij toetsing van de hypothese, dat twee categorische systemen stochastisch onafhankelijk zijn

9.10.1. TOEPASSING

Als men aselect een steekproef van n elementen uit een bepaalde populatie heeft getrokken en per element een waarneming aan twee verschillende categorische systemen heeft verricht, verkrijgt men n waarnemingsparen waarvan de uitkomsten in een $k \times r$ tabel zijn samen te vatten. Met behulp van de in 9.9 beschreven χ^2-toets kan men de nulhypothese toetsen dat de twee categorische systemen (stochastisch) onafhankelijk zijn.

Voorbeeld

9.15. Uit: Mens en Onderneming (1952), 408. In dit artikel wordt o.m. opgemerkt: 'Van 281 gevallen van genezen longtuberculose, welke na hun kuur minstens 4 jaar onder controle hebben gestaan, is nage-

gaan hoeveel recidieven er optraden na de genezing, hierbij tevens in rekening brengende het tijdsinterval na de kuur en de ernst van de doorgemaakte ziekte. De volgende tabel geeft hiervan een overzicht.

		Ernst v/d afwijking			
		Licht	Gevorderd	Ernstig	Totaal
Recidieven	Binnen 1 jaar na kuur	11(13%)	27(16%)	9(31%)	47
	1 — 2 jaar na kuur	6 (7%)	16(10%)	2 (7%)	24
	2 — 3 jaar na kuur	3 (3%)	10 (6%)	2 (7%)	15
	3 — 4 jaar na kuur	4 (5%)	8 (5%)	0 (0%)	12
Geen recidieven		62(72%)	105(63%)	16(55%)	183
Totaal		86(100%)	166(100%)	29(100%)	281

Hieruit blijkt dus, dat, hoewel zelfs bij lichte processen nog een grote recidiefkans bestaat, deze recidiefkans ook afhankelijk is van de ernst van het proces . . .'

9.10.2. DE TOETSINGSGROOTHEID

Onder de hypothese, dat beide categorische systemen (i.c. de ernst van het doorgemaakte proces en het optreden van recidieven) stochastisch onafhankelijk zijn, kunnen de verwachte frequenties op de gebruikelijke wijze worden berekend (product van de bij een veld behorende marginale frequenties, gedeeld door n). De toetsingsgrootheid is G volgens (9.23). Daar in voorbeeld 9.15 bij sommige velden zeer lage frequenties optreden, hebben wij de recidiefcategorieën < 1 en 1 tot 2 jaar, resp. 2 tot 3 en 3 tot 4 jaar, gecombineerd. De berekeningen zijn in de volgende tabel uitgevoerd:

		f_{ij} Afwijking			Totaal	e_{ij} Afwijking			$f_{ij} - e_{ij}$		
		L	G	E		L	G	E			
recidieven	tot 2 jaar	17	43	11	71	21,7	42,0	7,3	—4,7	+1,0	+3,7
	2 tot 4 jaar	7	18	2	27	8,3	15,9	2,8	—1,3	+2,1	—0,8
	geen	62	105	16	183	56,0	108,1	18,9	+6,0	—3,1	—2,9
Totaal		86	166	29	281	86,0	166,0	29,0	0	0	0

$$G_0 = (-4{,}7)^2/21{,}7 + (1{,}0)^2/42 + \ldots + (-2{,}9)^2/18{,}9 = 4{,}80$$

Er komt één verwachte frequentie voor die lager is dan 5, maar deze kan, gezien de grootte van de overige verwachtingen, zonder bezwaar worden toegelaten.

9.10.3. VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0 EN UITVOERING VAN DE TOETSING

Onder de nulhypothese volgt de grootheid G bij benadering een χ^2-verdeling met $(k - 1)(r - 1)$ vrijheidsgraden. H_0 wordt dus bij een drempelwaarde α verworpen, indien

$$G_0 \geqslant \chi^2_{1-\alpha}[\nu = (k - 1)(r - 1)].$$

De rechtse kans van de waargenomen $G_0 = 4{,}8$ blijkt ongeveer 0,3 te zijn [tabel D, $\nu = (3 - 1)(3 - 1) = 2 \times 2 = 4$]. Bij deze groepering luidt dus de conclusie dat de waargenomen samenhang nog heel goed toevallig kan zijn.

9.10.4. OPGAVEN

9.20. Uit DIXON en MASSEY (1957). A certain vitamin was thought to increase energy. In an experiment 100 men were given the vitamin. Also, as a control, 100 men were given harmless sugarpills but told that they were being given the vitamin. Their reactions are given in the table. Analyze the data.

	Had more energy	Had less energy	No change
Control group	20	10	70
Treated group	36	8	56

9.21. Uit de kursus: Summarizing Experimental Data, Columbia University, U.S.A.

The American Journal of Public Health for February, 1931, contains a paper on meningitis in New York City in which cases collected over a twenty year period are tabulated as to age and etiology. The following table is a portion of this tabulation:

Meningitis Cases, New York City

Form	Age in years				Total under 5
	Under 1	1	2	3 & 4	
Meningococcus	195	86	51	118	450
Tuberculous	136	183	107	150	576
Pneumococcus	28	13	8	14	63
Streptococcus	25	5	11	11	52
Total	384	287	177	293	1141

Analyze the material of this table so as to evaluate the evidence it presents on the relationship between form of meningits and age of case.

9.22. Uit VAN DER WIELEN, De Huisarts etc., Gez. Org. T.N.O., Van Gorcum & Comp., Assen, 1960. Onderstaande tabel bevat gegevens betreffende het apotheek houden en het aantal werkuren voor een groep huisartsen:

Apotheekhoudend	Aantal werkuren				Totaal
	< 8	8 — 10	10 — 12	> 12	
Wel	12 (60%)	44 (51%)	42 (42%)	11 (22%)	109
Niet	8 (40%)	43 (49%)	58 (58%)	40 (78%)	149
Totaal	20	87	100	51	258

Beschouw de groep als een aselecte steekproef uit de populatie van huisartsen in Nederland ten tijde van het onderzoek en toets (met $\alpha = 0{,}01$) of een significante correlatie aanwezig is.

9.23. Uit Br. Med. J., 11-11-50. Bij een onderzoek werd het effect van PAS (P), streptomycine (S) en een gecombineerde behandeling met beide middelen (PS) vergeleken bij drie aselect samengestelde groepen patiënten met longtuber-

culose. Elke patiënt verschafte na de behandeling een sputum-monster dat in één van drie categorieën werd ingedeeld:

Sputum	Behandeling			Totaal
	P	*S*	*PS*	
Uitstrijk +	56	46	37	139
Uitstrijk — Kweek +	30	18	18	66
Uitstrijk — Kweek —	13	20	35	68
Totaal	99	84	90	273

Verschillen de uitkomsten van de drie behandelingen significant bij een 5% drempel?

HOOFDSTUK 10.

KWANTITATIEVE VERDELINGSVRIJE TOETSEN

10.1. De χ^2-toets voor aanpassing

10.1.1. TOEPASSING

De χ^2-toets voor aanpassing dient voor het toetsen van de hypothese H_0, dat een aselecte steekproef van n waarden stamt uit een populatie met een bepaald verdelingstype (bijv. binomiaal, POISSON, normaal).

10.1.2. UITVOERING VAN DE TOETSING

De toets kan worden uitgevoerd als het aantal waarnemingen niet te klein is, zodat een frequentieverdeling kan worden opgesteld. In de regel wordt onder H_0 wel het verdelingstype gespecificeerd, maar is (zijn) geen parameter(s) van de verdeling gegeven. Men kan deze parameter(s) uit de waargenomen frequentieverdeling schatten en vervolgens de best-passende (binomiale, POISSON, normale) verdeling berekenen.[1] Op de waargenomen en best-passende (= verwachte) verdelingen kan daarna de χ^2-toets uit 9.3 worden toegepast. Het aantal vrijheidsgraden is dan echter

$$\nu = k - 1 - m,$$

waarin: m = het aantal uit de steekproef geschatte parameters,
k = het aantal klassen. Men dient er voor te zorgen dat de verwachte frequenties (in de best-passende verdeling) ten minste gelijk aan 5 zijn; zo nodig kan men daartoe klassen met een lage verwachting combineren. Vanzelfsprekend is k dan gelijk aan het aantal klassen, dat nà groepering aanwezig is.

Voor het berekenen van de best-passende binomiale verdeling en van de best-passende POISSON-verdeling ontleent men één parameter (P, resp. μ) aan de waargenomen verdeling, zodat het aantal vrijheidsgraden bij toepassing van de χ^2-toets $k - 1 - 1 = k - 2$ bedraagt. Voor de best-passende normale verdeling schat men twee parameters (μ en σ), zodat dan met $k - 3$ vrijheidsgraden getoetst wordt. Voorbeeld 10.1 verduidelijkt een en ander.

Toetst men de hypothese dat de steekproef afkomstig is uit een

[1] Zie resp. 4.3.3, 4.4.4 en 5.5.1.

volledig gespecificeerde verdeling (waarvan dus de parameters eveneens gegeven zijn), dan behoeven deze niet uit de steekproef te worden geschat en is $\nu = k - 1$.

Voorbeeld

10.1. In tabel 10.1 is de waargenomen en de best-passende normale verdeling uit tabel 5.2 overgenomen. Bij de berekening van G_0 zijn, om te lage verwachte frequenties te vermijden, de twee hoogste en de drie laagste klassen gecombineerd. De toetsing verloopt als volgt:

a. H_0: De steekproef stamt uit een normaal verdeelde populatie met $\mu = 13{,}4$ en $\sigma = 0{,}95$.
H_1: De steekproef stamt niet uit deze populatie.

b. Kies $\alpha = 0{,}05$.

c. Onder H_0 volgt de grootheid G [volgens (9.3)] bij benadering een χ^2-verdeling met $\nu = 9 - 1 - 2 = 6$ (na de groepering resteren 9 klassen, terwijl twee parameters zijn geschat).

d. Uit tabel D wordt afgelezen, dat H_0 bij een 5% drempel kan worden verworpen voor $G_0 > 12{,}59$ [d.i. $\chi^2_{0,95}(\nu = 6)$].

e. Bereken G_0. Deze berekening is in tabel 10.1 uitgevoerd en levert $G_0 = 3{,}29$ op.

f. Daar deze waarde aanzienlijk lager is dan 12,59 wordt H_0 niet verworpen.

Tabel 10.1. Toepassing van de χ^2-benadering op de waargenomen en best-passende normale verdeling van de hemoglobinewaarden in tabel 5.2

Klassegrenzen	f_i	e_i	$f_i - e_i$	$(f_i - e_i)^2$	$(f_i - e_i)^2/e_i$
15,75—16,25	2	1,3			
15,25—15,75	3 (5)	3,6 (4,9)	0,1	0,01	0,0020
14,75—15,25	11	9,9	1,1	1,21	0,1222
14,25—14,75	19	20,7	—1,7	2,89	0,1396
13,75—14,25	26	32,1	—6,1	37,21	1,1592
13,25—13,75	46	39,5	6,5	42,25	1,0696
12,75—13,25	38	35,7	2,3	5,29	0,1482
12,25—12,75	24	25,7	—1,7	2,89	0,1125
11,75—12,25	15	13,7	1,3	1,69	0,1234
11,25—11,75	4	5,5			
10,75—11,25	1 (6)	1,8 (7,8)	—1,8	3,24	0,4154
10,25—10,75	1	0,5			
Totaal	190	190,0	0	—	$G_0 = 3{,}2921$

Het is duidelijk, dat men op grond van de uitkomst van deze toetsing *niet* mag beweren, dat de hemoglobinewaarden van deze kinderen normaal verdeeld zijn. Vrijwel geen enkele verdeling is precies normaal en als men dus een niet-significante uitkomst vindt, komt dit doordat men niet genoeg waarnemingen verricht heeft om het zeker bestaande verschil statistisch aan te tonen. Als het aantal waarnemingen echter niet te klein is en G_0 is verre van significant, dan is het nut van zo'n uitkomst dat men zich de moeite kan besparen om een ander model voor de betrokken populatieverdeling te zoeken.

10.1.3. OPGAVEN

10.1. Onderzoek via de χ^2-toets ($\alpha = 0{,}05$) de overeenstemming tussen de waargenomen en best-passende verdeling bij:

a. Opgave 4.13 *b.* Opgave 4.14 *c.* Opgave 4.22 *d.* Opgave 5.14.

10.2. De toets van Kolmogorov-Smirnov (K-S toets)

10.2.1. TOEPASSING

Deze toets is in principe bestemd voor het toetsen van de hypothese H_0, dat een aselecte steekproef afkomstig is uit een populatie met een volledig gespecificeerde continue verdeling.

In 10.1 hebben wij gezien, dat men voor het toetsen van deze hypothese ook de χ^2-toets voor aanpassing kan gebruiken. Het verschil tussen de beide toetsen is echter: 1. De χ^2-toets is bruikbaar, als alleen het verdelingstype wordt gespecificeerd en als één of meer parameters uit de steekproef worden geschat. De *K-S* toets kan strikt genomen alleen worden gebruikt, als ook de parameters gegeven zijn. Uit gegevens van MASSEY (1951) blijkt echter, dat de toepassing van deze toets met geschatte parameters waarschijnlijk tot een uitkomst leidt, die conservatief is t.a.v. het verwerpen van H_0, d.w.z. een te grote overschrijdingskans oplevert. 2. Een nadeel van de χ^2-toets vergeleken met de *K-S* toets is, dat bij eerstgenoemde de waarnemingen in klassen moeten worden ingedeeld, zodat de uitkomst van de toetsing tot op zekere hoogte van deze indeling afhangt. Bij de *K-S* toets is een indeling in klassen niet nodig; deze verzwakt zelfs het onderscheidingsvermogen. Deze toets kan daardoor ook worden toegepast, als de steekproefomvang te klein is voor het gebruik van de χ^2-toets. 3. De *K-S* toets geldt strikt genomen alleen voor continu verdeelde waarden. De toets kan wel voor discreet verdeelde grootheden worden gebruikt, maar levert ook dan een te grote overschrijdingskans op.

10.2.2. UITVOERING VAN DE TOETSING

De toetsing wordt meestal grafisch uitgevoerd.

Voorbeeld

10.2. Men wil onderzoeken of een populatie standaardnormaal verdeeld is (d.w.z. een normale verdeling volgt met $\mu = 0$ en $\sigma = 1$) en trekt daartoe aselect een steekproef van 20 elementen, die de volgende waarden oplevert (naar grootte gerangschikt en afgerond op 2 decimalen):

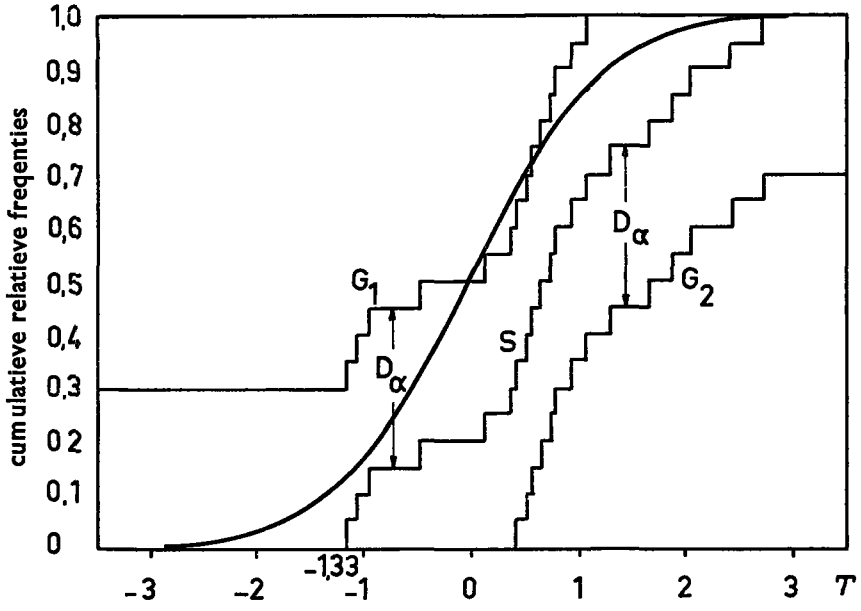
—1,33	—1,06	—0,93	—0,45	0,15	0,40	0,44	0,55	0,59	0,68
0,78	0,79	0,79	1,10	1,32	1,69	1,91	2,06	2,44	2,75.

De toetsing heeft het volgende verloop (zie figuur 10.1):

a. Teken het relatieve somfrequentie-histogram van de verkregen waarnemingen. Voor het voorbeeld is dit de trappenlijn S.

b. Teken boven en onder de trappenlijn S, op een afstand D_α, de parallel verlopende trappenlijnen G_1 en G_2.

In tabel S zijn voor een reeks waarden van $n \leqslant 35$ en een vijftal onbetrouwbaarheidsdrempels α de waarden van D_α gegeven. Voor $n > 35$ kunnen deze met de in tabel S opgenomen formules worden berekend. In het voorbeeld is $n = 20$ en bij $\alpha = 0{,}05$ volgt uit tabel S: $D_{0,05}(20) = 0{,}294$. De trappenlijnen G_1 en G_2 in figuur 10.1 bevinden zich dus op deze afstand boven, resp. beneden, de trappenlijn S.



Figuur 10.1. Grafiek voor de *K-S* toets, zie voorbeeld 10.2.

c. Teken in de grafiek de cumulatieve kansverdeling onder H_0. Bij het voorbeeld luidt H_0: de populatieverdeling is standaardnormaal. Uit tabel A kan men voor een aantal waarden van de standaardnormaal verdeelde grootheid T de linkse kansen aflezen, bijv. voor:

T	—2,0	—1,5	—1,0	—0,5	0	0,5	1,0	1,5	2,0
$P_L(T)$	0,0228	0,0668	0,1587	0,3085	0,5000	0,6915	0,8413	0,9332	0,9772

De punten $[T, P_L(T)]$ kunnen nu in de grafiek worden getekend en door deze punten kan de kromme van de cumulatieve populatieverdeling worden geschetst.

d. De waarde $D_\alpha(n)$ is zodanig bepaald, dat H_0 op grond van een steekproef van n elementen bij een drempelwaarde α kan worden verworpen, indien de kromme ergens buiten de grenslijnen G_1 en G_2 valt.

Daar de kromme in figuur 10.1 (tussen $T = 0$ en $T = 1$) buiten de grenslijnen loopt, wordt H_0 bij een 5% drempel verworpen.[1]

10.2.3. SCHATTING VAN DE POPULATIEVERDELING

De trappenlijnen G_1 en G_2, die met D_α uit S zijn geconstrueerd, sluiten een betrouwbaarheidsgordel voor de populatieverdeling in met een betrouwbaarheidscoëfficiënt $1 - \alpha$, zodat men met een betrouwbaarheid $1 - \alpha$ kan beweren, dat de grenslijnen de cumulatieve relatieve populatieverdeling omsluiten.

Tabel S kan ook worden gebruikt om de steekproefomvang te bepalen die nodig is om bij een bepaalde waarde $1 - \alpha$ een betrouwbaarheidsgordel van bepaalde breedte te verkrijgen. Onderstel bijv. dat men een zodanige steekproef wil trekken, dat een 95% gordel verkregen wordt, waarvan de breedte 0,05 bedraagt. Dan is $D_\alpha = 0{,}05/2 = 0{,}025$, dus (zie tabel S):

$$\frac{1{,}36}{\sqrt{n}} = 0{,}025, \text{ zodat: } n = \left[\frac{1{,}36}{0{,}025}\right]^2 = (54{,}4)^2 = 2960.$$

Wenst men een 99% betrouwbaarheidsgordel van dezelfde breedte, dan dient de steekproefomvang te zijn:

$$n = \left[\frac{1{,}63}{0{,}025}\right]^2 = 4251.$$

10.2.4. TOEPASSING OP DISCREET VERDEELDE WAARNEMINGEN

Bij een discreet verdeelde populatie is de *K-S* toets van belang, als de steekproefomvang zo klein is, dat men bij het gebruik van de χ^2-toets tot een (min of meer willekeurige) groepering van klassen moet overgaan om te lage verwachte frequenties te vermijden (zie opgave 10.3). De toets kan dan meestal eenvoudiger tabellarisch worden uitgevoerd en verloopt op de volgende wijze:

a. Bereken uit de waargenomen frequenties f_i (de verwachte frequenties e_i) de somfrequenties $F_i(E_i)$ en de relatieve somfrequenties F_i/n (E_i/n).

b. Bepaal het maximale verschil $| F_i/n - E_i/n | = D_0$.

[1] De in het voorbeeld gebruikte steekproef werd aselect getrokken uit een normaal verdeelde populatie met $\mu = 0{,}5$ en $\sigma = 1$. Teken in figuur 10.1 ook de hiermede corresponderende cumulatieve verdeling.

BIBLIOTHEEK NEDERLANDS INSTITUUT
VOOR PRAEVENTIEVE GENEESKUNDE TNO
WASSENAARSEWEG 56 - LEIDEN

c. Kies de drempelwaarde α waarbij getoetst wordt en ontleen $D_\alpha(n)$ aan tabel S.

d. Verwerp H_0 bij de gekozen drempelwaarde indien $D_0 \geqslant D_\alpha(n)$ is.

Voorbeeld

10.3. In tabel 10.2 is de tabellarische uitvoering van de *K-S* toets gedemonstreerd voor de uitkomsten van 360 worpen met een dobbelsteen (voorbeeld 9.1). De te toetsen hypothese H_0 luidt dat de dobbelsteen zuiver is en men kiest $\alpha = 0{,}05$.

Tabel 10.2. Toepassing van de *K-S* toets op discreet verdeelde uitkomsten; 360 worpen met een dobbelsteen (voorbeeld 9.3)

	Aantal ogen x_i	1	2	3	4	5	6
Waarnemingen	f_i	67	52	39	72	55	75
	F_i	67	119	158	230	285	360
	F_i/n	0,189	0,330	0,439	0,639	0,792	1,000
Verwachtingen	e_i	60	60	60	60	60	60
	E_i	60	120	180	240	300	360
	E_i/n	0,167	0,333	0,500	0,667	0,833	1,000

$D_0 = \text{max.} \mid F_i/n - E_i/n \mid = \mid 0{,}500 - 0{,}439 \mid = 0{,}061$
Tabel S: $D_{0,05}(360) = 1{,}36/\sqrt{360} = 1{,}36/18{,}97 = 0{,}072$
$D_0 = 0{,}061$ is kleiner dan $D_{0,005}(360) = 0{,}072$, zodat H_0 niet wordt verworpen.

10.2.5. OPGAVEN

10.2. Men heeft aselect de volgende steekproef getrokken:

7,53	9,04	12,02	11,70	9,16	12,26	9,37	11,06	9,19	7,62
7,34	9,07	8,95	10,62	12,22	12,60	10,41	9,06	10,52	9,90

Toets de hypothese dat deze steekproef stamt uit een normaal verdeelde populatie met:

a. $\mu = 9$, $\sigma = 2$, *b.* $\mu = 10$, $\sigma = 3$, *c.* $\mu = 10$, $\sigma = 2$ (Kies $\alpha = 0{,}05$).

10.3. De waargenomen verdeling van een discreet verdeelde grootheid is:

x_i	1	2	3	4	5
f_i	0	1	0	5	4

a. Toets, met $\alpha = 0{,}05$, de hypothese H_0: $P(1) = P(2) = \ldots = P(5) = 1/5$ door middel van de *K-S* toets.

SIEGEL (1956) merkt naar aanleiding van deze gegevens op: 'A reanalysis by the χ^2 test ... will highlight the superior power of the *K-S* test.' Hij past de χ^2-toets toe, na de klassen 1 en 2 resp. de klassen 3, 4 en 5 te hebben gecombineerd en vindt dan $G_0(\nu = 1) = 3{,}75$ (zonder continuïteitscorrectie).

b. Kunt u zich met SIEGEL's conclusie verenigen? Wat toont dit voorbeeld wèl aan?

10.3. De rang-tekentoets[1]

10.3.1. TOEPASSING

Deze toets kan men toepassen als de (relatieve) grootte van de verschillen tussen gepaarde waarnemingen bekend is. De nulhypothese luidt dat de verschillen afkomstig zijn uit een symmetrisch verdeelde populatie met mediaan nul. Bij de praktische toepassing van de toets komt dit neer op de hypothese dat de verschillen tussen de gepaarde waarnemingen toevallig zijn.

Toelichting [2].

Onderstel dat men wil nagaan of een geneesmiddel invloed uitoefent op een kwantitatieve grootheid, zoals bloeddruk, gehalte van één of andere stof in het bloed, e.d. Ook reeds zonder gebruik van het middel bestaat er een individuele variabiliteit, die van persoon tot persoon kan verschillen. Beschouw nu één patiënt en wel voordat het middel is toegediend. De grootheid a die gemeten wordt is een variabele, die voor de beschouwde patiënt een bepaalde verdeling volgt. Men dient nu het geneesmiddel toe. Herhaalde (in gedachten of werkelijk verrichte) waarnemingen leveren dan een grootheid b. Heeft het geneesmiddel geen enkele uitwerking, dan bezitten de grootheden $\underline{a}$ en $\underline{b}$ dezelfde verdeling. Men kan gemakkelijk bewijzen, dat de grootheid $\underline{z} = \underline{a} - \underline{b}$ symmetrisch verdeeld is met gemiddelde = mediaan gelijk aan nul. Heeft het middel echter enigerlei uitwerking, dan zal $\underline{b}$ een enigszins andere verdeling bezitten dan $\underline{a}$ en de grootheid $\underline{z}$ is dan niet meer symmetrisch verdeeld t.o.v. het nulpunt.

Indien het verrichten van waarnemingen niet te lastig en voor de patiënt niet hinderlijk is, kan men aan één patiënt meerdere waarnemingen voor en na de toediening verrichten. Men kan dan, voor iedere patiënt afzonderlijk, met een toets voor twee aselecte steekproeven (bijv. met de toets van WILCOXON) nagaan, of de verdeling van $\underline{a}$ (vóór toediening) dezelfde is als die van $\underline{b}$ (na toediening). Daar men dit niet slechts van één patiënt wil weten, maar van zoveel mogelijk en er ook vaak bezwaren bestaan tegen het verrichten van meerdere waarnemingen per patiënt, zal men overgaan tot het verrichten van één waarneming voor en één na de toediening bij zo veel mogelijk patiënten. Deze opzet van het experiment heeft het grote voordeel, dat de draagwijdte van de verkregen conclusie zo groot mogelijk wordt.

Voor ieder der patiënten vindt men dan één waarde van de grootheid $\underline{z}$, die, als het middel geen uitwerking heeft, symmetrisch verdeeld is t.o.v. het nulpunt. De grootheden $\underline{z}$ zullen echter in het algemeen niet voor alle patiënten dezelfde verdeling bezitten. Noemt men de bij n patiënten bepaalde waarden (= verschillen, die niet gelijk zijn aan 0):

$$z_1, z_2, \ldots, z_i, \ldots, z_n$$

en de bijbehorende stochastische grootheden:

$$\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_i, \ldots, \underline{z}_n$$

[1] De toets is het eerst aangegeven door WILCOXON (1945) en wordt in de uitvoerige handleiding van BENARD & VAN EEDEN (1956) de *symmetrietoets van* WILCOXON genoemd. O.i. is de naam rang-tekentoets (die ook in de buitenlandse literatuur gebruikelijk is) sprekender en geeft deze minder kans op verwarring met *de* toets van WILCOXON voor twee steekproeven.

[2] Wij volgen hier grotendeels het betoog van HEMELRIJK in (1950a).

dan stelt de rangtekentoets ons in staat de hypothese H_0 te toetsen, dat de grootheden $\underline{z}_i$ alle symmetrisch verdeeld zijn met het nulpunt als punt van symmetrie. Bij deze toets behoeft dus niet te worden ondersteld, dat alle $\underline{z}_i$ dezelfde verdeling bezitten.

Men toetst dus de praktische hypothese, dat de toediening van het geneesmiddel bij geen der patiënten invloed op de gemeten grootheid $\underline{a}$ uitoefent, of wel: dat de verschillen tussen de gepaarde metingen toevallig zijn. Bij het onderzoek moet er voor gezorgd worden, dat alle paren van verrichte metingen onafhankelijk zijn. Deze voorwaarde is, daar men verschillende personen onderzoekt, gewoonlijk reeds automatisch vervuld, mits de meettechniek voldoende constant is (zie opgave 10.6). Vanzelfsprekend dient men er ook voor te zorgen, dat gedurende het onderzoek alleen de eventuele uitwerking van het middel wordt gemeten, d.w.z. dat geen andere factoren invloed op de uitkomsten (kunnen) hebben.

Voorbeeld

10.4. Men wil onderzoeken, of het calciumgehalte in het serum van konijnen door intraveneuze toediening van vitamine D_2 wordt beïnvloed. Men verricht daartoe het volgende experiment: bij 5 proefdieren meet men het calciumgehalte. Vervolgens krijgt elk dier een injectie met 200.000 I.E. vitamine D_2 in solubilusator. Na drie dagen worden de calciumgehalten opnieuw bepaald. De uitkomsten van dit experiment staan in tabel 10.3.

Tabel 10.3. Calciumgehalte (mg%) in het serum van konijnen, A = voor, B = drie dagen na toediening van 200.000 I.E. D_2

Dier nr.	A	B	$B-A$	Rangnummers	Rangnummers met teken
1	16,2	17,3	+1,1	2	+ 2
2	12,9	16,5	+3,6	5	+ 5
3	15,8	17,7	+1,9	3	+ 3
4	17,0	19,2	+2,2	4	+ 4
5	14,6	14,4	−0,2	1	−1
Controle: $\Sigma R = \frac{n(n+1)}{2} = \frac{5 \times 6}{2} = 15$				15	$\begin{array}{r} +14 \\ -\ 1 \\ \hline +13 = V_0 \end{array}$

10.3.2. DE TOETSINGSGROOTHEID

De verkregen uitkomsten worden als volgt bewerkt (zie tabel 10.3):

1. Bepaal het verschil tussen elk stel gepaarde waarnemingen (bijv. $B - A$).

2. Schrap alle verschillen die gelijk zijn aan nul. Van de m verschillen blijven er dan n over (in het voorbeeld komen geen verschillen met de waarde 0 voor, zodat $n = 5$).

3. Vervang de *absolute waarden* van de verschillen door rangnummers te beginnen met 1 voor het verschil met de kleinste absolute waarde. De rangnummers worden dus toegekend, zonder op het teken te letten!

4. Voorzie de rangnummers van de negatieve verschillen, $R(-)$, van een minteken en die van de positieve verschillen, $R(+)$, van een plusteken.

5. De toetsingsgrootheid V is gelijk aan de algebraïsche som van de getallen verkregen in 4, d.w.z. de som van de rangnummers met een positief teken, verminderd met de som van de rangnummers met een negatief teken

$$V = \Sigma R(+) - \Sigma R(-).$$

Opmerking: WILCOXON (1945) gebruikt als toetsingsgrootheid $\Sigma R(+)$. Het is duidelijk dat $V = 2\Sigma R(+) - \frac{1}{2}n(n+1)$.

10.3.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

De hypothese H_0 die men bij voorbeeld 10.4 wenst te toetsen luidt dat de toediening van het vitaminepreparaat geen invloed uitoefent op het calciumgehalte van het serum. Als geen gelijken (verschillen met dezelfde absolute waarde) optreden, kan de verdeling van de toetsingsgrootheid $\underline{V}$ onder H_0 worden verkregen, door alle 2^n mogelijke uitkomsten van het experiment systematisch uit te schrijven. In tabel 10.4 is dit voor de $2^5 = 32$ mogelijke uitkomsten van het experiment in voorbeeld 10.4 uitgevoerd.

Tabel 10.4. Verdeling van $\underline{V}$ onder H_0 voor $n = 5$

A. Rangnummers met de mogelijke tekens

Rangnr.	Tekens																															
1	+	−	+	+	+	+	−	−	−	−	+	+	+	+	+	+	−	−	−	−	−	−	+	+	+	+	−	−	−	−	+	−
2	+	+	−	+	+	+	−	+	+	+	−	−	−	+	+	+	−	−	−	+	+	+	−	−	−	+	−	−	−	+	−	−
3	+	+	+	−	+	+	+	−	+	+	−	+	+	−	−	+	−	+	+	−	−	+	−	−	+	−	−	+	−	−	−	−
4	+	+	+	+	−	+	+	+	−	+	+	−	+	−	+	−	+	−	+	−	+	−	−	−	+	−	−	−	+	−	−	−
5	+	+	+	+	+	−	+	+	+	−	+	+	−	+	−	−	+	+	−	+	−	−	+	−	−	−	+	−	−	−	−	−
$\Sigma R(+)$	15	14	13	12	11	10	12	11	10	9	10	9	8	8	7	6	9	8	7	7	6	5	6	5	4	3	5	4	3	2	1	0
$\Sigma R(-)$	0	1	2	3	4	5	3	4	5	6	5	6	7	7	8	9	6	7	8	8	9	10	9	10	11	12	10	11	12	13	14	15
V	15	13	11	9	7	5	9	7	5	3	5	3	1	1	−1	−3	3	1	−1	−1	−3	−5	−3	−5	−7	−9	−5	−7	−9	−11	−13	−15

B. Verdeling van $\underline{V}$ onder H_0

V	−15	−13	−11	−9	−7	−5	−3	−1	1	3	5	7	9	11	13	15	Tot.
f	1	1	1	2	2	3	3	3	3	3	3	2	2	1	1	1	32

Als er geen gelijken zijn bezit de verdeling van $\underline{V}$ onder H_0 de volgende eigenschappen:

a. De verdeling is discreet. V neemt alleen even waarden aan, als $n = 3, 4, 7, 8, 11, 12$, enz., dus als n een 4-voud of 4-voud min 1 is. Voor andere waarden van n neemt V alleen oneven waarden aan.

b. De verdeling is symmetrisch t.o.v. $\mu_v = 0$

c. De variantie van de verdeling is

$$\sigma_v^2 = \frac{n(n+1)(2n+1)}{6}. \tag{10.1}$$

10.3.4. UITVOERING VAN DE TOETSING ZONDER GELIJKEN

Exacte toetsing ($n \leqslant 20$). In BENARD en VAN EEDEN (1956, tabel I) vindt men voor $n = 3$ t/m 20 exacte rechts éénzijdige overschrijdingskansen van V. Wegens de symmetrie van de verdeling wordt voor een waarde van V die negatief is de links éénzijdige overschrijdingskans gevonden door in de tabel de rechtse kans van $-V$ op te zoeken.

Voor het experiment in voorbeeld 10.4 is $V_0 = 13$. Uit de kansverdeling in tabel 10.4B blijkt, dat de tweezijdige overschrijdingskans van deze uitkomst gelijk is aan: $P_L(-13) + P_R(+13) = 2[P(15) + P(13)] = 2 \times {}^2/_{32} = {}^1/_8 =$ $= 0{,}125$. Bij een 5% drempel kan H_0 dus niet worden verworpen. Overigens blijkt, dat de omvang van het experiment te klein is om bij tweezijdige toetsing met $\alpha = 0{,}05$ tot het verwerpen van H_0 te kunnen overgaan; zelfs als alle verschillen positief zijn ($V_0 = 15$) is de tweezijdige overschrijdingskans nog groter dan 0,05, nl. 0,0625. Als echter de tegenhypothese luidt dat de calciumgehalten tengevolge van de toediening van het preparaat stijgen, zodat rechts éénzijdig wordt getoetst, kan bij de uitkomst $V_0 = 15$ de nulhypothese bij een 5% drempel worden verworpen.

Het gebruik van tabel J ($n \leqslant 100$). Tabel J bevat voor $n = 5$ t/m 100 en *tweezijdige toetsing* bij de drempelwaarden 0,01, 0,02, 0,05 en 0,10 de rechter kritieke waarden V_R van $\underline{V}$. Deze zijn zo gekozen dat $P(\underline{V} \geqslant V_R \mid H_0) \leqslant \alpha$ is. De linker kritieke waarden bij deze onbetrouwbaarheidsdrempels zijn: $V_L = -V_R$.

Voor $n = 14$ leest men bijv. in tabel J af:

Tabel J

$n = 14$	α			
	0,01	0,02	0,05	0,10
V_R	81	75	63	55
zodat: V_L	−81	−75	−63	−55

Hieruit volgt bijv. $P_D(-63) = P_D(63) \leqslant 0{,}05$; $0{,}05 > P_D(70) > 0{,}02$. Toetst men tweezijdig met onbetrouwbaarheidsdrempel α dan wordt H_0 verworpen indien

$$| V_0 | \geqslant V_R(\alpha, n).$$

Tabel J kan ook worden gebruikt bij *éénzijdige toetsing*, mits men de *drempelwaarden halveert.* Zo geldt bijv. voor $n = 14$ (zie voorgaande tabel)

$$P_L(-75) \leqslant 0{,}01, \quad P_R(75) \leqslant 0{,}01, \quad P_L(-81) \leqslant 0{,}005.$$

Toetst men links éénzijdig (alternatieve hypothese: de mediaan van de verschillen is kleiner dan 0), dan wordt H_0 met onbetrouwbaarheidsdrempel α verworpen, indien $V_0 \leqslant V_L(2\alpha, n)$ is. Toetst men rechts éénzijdig (alternatief: de mediaan van de verschillen is groter dan 0). dan besluit men met een onbetrouwbaarheid α tot het verwerpen van H_0, indien $V_0 \geqslant V_R(2\alpha, n)$ is.

Voorbeeld

10.5. Onderstaande uitkomsten betreffen het aantal erythrocyten (mill. per mm³) bij 12 patiënten met anemie tengevolge van bloedverlies, (*a*) onmiddellijk voor het begin van de therapie, (*b*) na een intraveneuze ijzertherapie van 15 dagen:

Patiënt nr.	1	2	3	4	5	6	7	8	9	10	11	12
a	2,9	2,5	2,5	3,7	3,0	3,6	3,3	2,4	3,5	3,3	2,9	1,9
b	3,9	3,8	3,6	3,6	3,6	3,4	3,8	4,1	4,2	3,3	3,8	3,7
$b - a$	1,0	1,3	1,1	—0,1	0,6	—0,2	0,5	1,7	0,7	0,0	0,9	1,8
Rang-nummer	7	9	8	1	4	2	3	10	5		6	11
Rang-nummer met teken	7	9	8	—1	4	—2	3	10	5		6	11

$V_0 = 7+9+8+4+3+10+5+6+11 - (1 + 2) = 60 \qquad n = 12 - 1 = 11$

Men toetst de hypothese H_0, dat het aantal erythrocyten niet systematisch verandert, met als alternatief H_1 dat in de onderzoekperiode een verhoging van het aantal erythrocyten optreedt en kiest $\alpha = 0{,}01$. Onder H_1 verwacht men, dat de verschillen $b - a$ in het algemeen positief zullen zijn met grote waarden, zodat rechts éénzijdig wordt getoetst. In tabel J staat bij $n = 11$ en $\alpha = 0{,}02$ als rechter kritieke waarde $V_R = 52$. Daar $V_0 = 60$ deze waarde overschrijdt, wordt H_0 verworpen ten gunste van H_1. (Uit tabel J blijkt verder, dat de rechtse kans van V_0 kleiner is dan 0,005. De exacte rechtse kans blijkt 0,00244 te bedragen.) Mag men nu zonder meer tot een gunstig effect van de therapie concluderen?

De normale benadering ($n > 20$). Bij niet te kleine n kan men gebruik maken van het feit, dat de grootheid $\underline{V}$ bij benadering normaal verdeeld is met gemiddelde 0 en spreiding σ_v(10.1), zodat de grootheid

$$(10.2) \qquad \underline{T} = \frac{\underline{V} - \mu_v}{\sigma_v} = \frac{\underline{V}}{\sigma_v} = \frac{\underline{V}}{\sqrt{n(n+1)(2n+1)/6}}$$

bij benadering standaardnormaal verdeeld is. Daar de discrete verdeling van V benaderd wordt door de continue normale verdeling, is het aanbrengen van een continuïteitscorrectie gewenst. De afstand tussen de opeenvolgende waarden van V bedraagt 2 en de correctie is gelijk aan de helft van deze afstand, dus 1. Bij gebruik van de normale benadering berekent men zodoende

$$(10.3) \qquad T_0 = \frac{|V_0| - 1}{\sigma_v} \text{ en } P_D(T_0) \text{ bij tweezijdige toetsing,}$$

of

(10.4.1) $T_0 = \dfrac{V_0 + 1}{\sigma_v}$ en $P_L(T_0)$ bij links éénzijdige toetsing,

resp.

(10.4.2) $T_0 = \dfrac{V_0 - 1}{\sigma_v}$ en $P_R(T_0)$ bij rechts éénzijdige toetsing.

H_0 wordt verworpen indien de gevonden overschrijdingskans kleiner dan α is. Als men uitsluitend wil vaststellen, of de berekende T_0 significant is bij een drempelwaarde 0,05, 0,01 of 0,001 kan men gebruik maken van tabel 6.10.

Voorbeeld

10.6. Onderstel, dat $n = 20$ en dat men $V_0 = 124$ vindt. Uit tabel J blijkt, dat de rechts éénzijdige overschrijdingskans van deze uitkomst $< 0{,}01$ is. De exacte rechtse kans blijkt 0,009617 te bedragen. De normale benadering levert op:

$$T_0 = \frac{124 - 1}{\sqrt{20 \times 21 \times 41/6}} = \frac{123}{53{,}57} = 2{,}296 \qquad P_R(2{,}296) = 0{,}0108 \text{ (tabel A).}$$

Ondanks de kleine n is de benadering dus reeds zeer bevredigend.

10.3.5. UITVOERING VAN DE TOETSING MET GELIJKEN

Wanneer n klein is en er veel gelijken (verschillen met dezelfde absolute waarde) zijn, kan men strikt genomen noch van tabel J, noch van de normale benadering gebruik maken. Men dient dan de verdeling van $\underline{V}$ exact te berekenen en BENARD en VAN EEDEN (1956) geven hiervan een voorbeeld.

Indien er relatief weinig gelijken voorkomen kan men aan elk verschil binnen een groep gelijken het gemiddelde toekennen van de rangnummers, die de gelijken in deze groepen zouden moeten bezetten, waarna V_0 op de gebruikelijke wijze kan worden bepaald. Is n klein en zijn er slechts enkele groepen met weinig gelijken, dan kan men deze uitkomst nog wel bij benadering interpreteren met tabel J. Is n groter en komen er meerdere groepen gelijken van grotere omvang voor, dan kan de normale benadering worden toegepast. Als de gelijken gesplitst kunnen worden in k groepen van resp. $t_1, t_2, \ldots, t_k$ gelijken (zie tabel 10.5), waarbij men voor elk verschil dat niet tot een groep gelijken behoort $t = 1$ neemt en men stelt

$$D = \sum_{i=1}^{k} t_i^3,$$

dan dient men bij deze benadering in plaats van σ_v volgens (10.1) te nemen

$$\sigma_v'^2 = \sigma_v^2 - \frac{D - n}{12}. \tag{10.5}$$

Het is duidelijk dat ingeval geen gelijken optreden $D = n$.

Bij het optreden van gelijken is de afstand tussen de opeenvolgende waarden van V in het algemeen niet constant, zodat geen algemene regel voor de continuïteitscorrectie te geven is. Met een correctie ter grootte van 1 blijft men aan de veilige kant (d.w.z. corrigeert men te sterk, zodat men tot een te grote overschrijdingskans komt).

Met (10.5) vindt men voor de spreiding van V onder H_0 een kleinere waarde dan met (10.1). Past men dus, ondanks het optreden van gelijken, de normale benadering met (10.1) toe, dan vindt men een (iets) te lage T_0 en dus een (iets) te grote overschrijdingskans. Is deze te grote kans reeds kleiner dan of gelijk aan α, dan kan men H_0 verwerpen zonder dat men tot berekening van D en het aanbrengen van de correctie behoeft over te gaan.

Voorbeeld

10.7. Als voorbeeld beschouwen wij de 16 verschillen tussen gepaarde waarnemingen in tabel 10.5 De te toetsen hypothese H_0 luidt, dat deze verschillen toevallig zijn en men kiest als drempelwaarde $\alpha = 0{,}05$.

Tabel 10.5. Toepassing van de rang-tekentoets als gelijken voorkomen

Waargenomen verschillen	Rangnummers met tekens		Gelijken	
	—	+	t_i	t_i^3
—0,1	2			
—0,1	2		$t_1 = 3$	27
+0,1		2		
—0,2	4½		$t_2 = 2$	8
+0,2		4½		
—0,3	6		$t_3 = 1$	1
—0,4	7		$t_4 = 1$	1
—0,5	9½			
—0,5	9½		$t_5 = 4$	64
—0,5	9½			
+0,5		9½		
+0,7		12	$t_6 = 1$	1
—0,8	13½		$t_7 = 2$	8
—0,8	13½			
—1,2	15		$t_8 = 1$	1
—1,5	16		$t_9 = 1$	1
Totaal	—108	+28		
	+ 28		$\Sigma t_i = n = 16$	$D = 112$
	— 80 = V_0			

De eerste groep gelijken bestaat uit de drie verschillen met absolute waarde 0,1, zodat $t_1 = 3$. Zij dienen de rangnummers 1, 2 en 3 te bezetten en krijgen dus alle het gemiddelde rangnummer $(1 + 2 + 3)/3 = 2$ (met het juiste teken).

De tweede groep gelijken bestaat uit de verschillen — 0,2 en + 0,2, zodat $t_2 = 2$. Zij moeten de plaatsen 4 en 5 bezetten en krijgen dus beide het gemiddelde rangnummer $4\frac{1}{2}$. Voor de verschillen — 0,3 en — 0,4 noteert men resp. $t_3 = 1$ en $t_4 = 1$, enz. De juistheid van de bewerkingen kan worden gecontroleerd, daar de som van de absolute waarden van de positieve en negatieve rangsommen gelijk dient te zijn aan $\frac{1}{2}n(n+1)$ en $\Sigma t_i = n$.

Men vindt $V_0 = -80$. Berekent men de spreiding van de verdeling van V onder H_0 met (10.1) dan is de uitkomst:

$$\sigma_V^2 = \frac{(16)\,(17)\,(32+1)}{6} = 1496, \qquad \sigma_V = \sqrt{1496} = 38{,}68.$$

De normale benadering met (10.3) levert dus op

$$T_0 = \frac{|-80|-1}{38{,}68} = \frac{79}{38{,}68} = 2{,}04.$$

Daar deze uitkomst groter is dan 1,96 (zie tabel 6.10) kan men H_0 bij de gekozen 5% drempel verwerpen en behoeft men feitelijk niet meer over te gaan tot de berekening met de correctie voor gelijken. Voeren wij deze volledigheidshalve uit, dan is volgens (10.5)

$$\sigma_V'^2 = 1496 - \frac{112-16}{12} = 1488, \qquad \sigma_V' = \sqrt{1488} = 38{,}57$$

Met de normale benadering vindt men nu

$$T_0 = \frac{|-80|-1}{38{,}57} = \frac{79}{38{,}57} = 2{,}05.$$

Het verschil met de voorgaande uitkomst is dus van weinig betekenis.

In 10.5.6 behandelen wij een berekeningsmethode voor V_0 die van belang is wanneer het aantal verschillen groot is en veel gelijken optreden.

10.3.6. OPGAVEN

10.4. Toets de hypothese, dat de volgende verschillen tussen gepaarde continu verdeelde waarnemingen toevallig zijn. Kies $\alpha = 0{,}05$.
—11 —10 —8 —3 —1 —1 0 0 0 +2 +4 +5 +8 +9 +14 +17 +20 +25 +30 +46 +67

10.5. D'Amour en Smith (J. Pharmac. Exp. Ther. (1941)) ontwikkelden een methode om de reactietijd te meten van ratten, waarvan de staart werd bestraald door een lichtbundel van zodanige intensiteit dat na ongeveer 5 seconden pijn optrad. Bij een onderzoek naar het effect van morphine op de reactietijd werden 12 ratten getest voor de toediening van morphine en vervolgens 30 minuten na een intraperitoneale injectie van 6 mg morphine. De reactietijden (in sec.) waren:

Rat	1	2	3	4	5	6	7	8	9	10	11	12
Voor	5,0	4,5	4,0	4,4	4,1	5,0	4,6	4,7	5,0	4,4	4,2	4,7
Na	5,0	4,8	9,0	4,7	4,2	5,0	4,2	5,0	4,2	5,8	4,6	4,9

Pas de rangtekentoets toe. Kies $\alpha = 0{,}10$.

10.6. Uit: GREENBERG (1951).

'In an experiment to measure the development of immunity in mice to a particular nematode (Trichinella), young mice are matched according to sex and placed in two groups, one experimental and the other control. After a series of stimulating infections among the experimental group only, the immunity is measured by challenging both groups with a like number of organisms. To challenge each mouse with the infecting parasites, a small amount of solution is prepared containing the necessary larvae. Thus, the solution might be such that 0,05 cc is supposed to contain 200 larvae. Each mouse is then given 0,05 cc by mouth from a syringe. Mice from the two groups are infected alternately, in the hope that this will cancel variations in dosage.

A systematic scheme of this kind formed by an odd-even, odd-even order of injecting the mice during challenge may lead to a serious bias. Following the previously described procedure, instead of inoculating mice with the larvae, thirty dosages were placed on individual slides for the purpose of counting the number of larvae. The odd-numbered slides were considered experimentals, the even-numbered ones controls, exactly as it would have occured in an experiment. The uniformity data for these thirty slides are presented in the table.'

Set	Exp.	Control	*C—E*
1	200	207	+ 7
2	204	206	+ 2
3	211	208	— 3
4	211	213	+ 2
5	212	217	+ 5
6	220	221	+ 1
7	225	235	+10
8	238	238	0
9	238	237	— 1
10	238	246	+ 8
11	243	243	0
12	251	252	+ 1
13	259	261	+ 2
14	275	277	+ 2
15	279	283	+ 4
Mean	233,6	236,3	+2,7

a. Verschillen de beide waarnemingsreeksen significant bij een drempelwaarde 0,05?

b. Wat valt u bij deze reeksen op en wat zou daarvoor de verklaring kunnen zijn?

c. Op welke wijze zou men naar uw mening de uitvoering van de proef en de proefopzet kunnen verbeteren?

10.7 Uit: N.T.v.G. 100 (1956), 1364.

Uit tabel VII: Natrium- en kaliumgehalte in het bloed in mEq. per liter vóór en na behandeling met Diamox bij 15 patiënten. De waarnemingen 1 t.m. 6, 8 en 9 betreffende patiënten met hartklepgebreken, voornamelijk van het ostium mitrale. De 10e en 11e waarneming zijn gedaan bij lijders aan chronisch cor pulmonale en de waarnemingen 14 t.m. 18 betreffen patiënten met arteriosclerotische hartziekten, waaronder ook hypertensie. Alle patiënten waren bij het begin van de toediening gedecompenseerd.

Patiënt no.		1	2	3	4	5	6	8	9	10	11	14	15	16	17	18
Natr.	Voor	135	134	143	143	130	130	128	137	140	139	136	132	140	136	140
	Na	140	135	140	144	138	136	137	132	143	140	141	140	139	146	138
K	Voor	4,7	4,7	4,1	4,3	5,7	4,5	4,6	5,1	5,8	4,6	4,8	4,5	4,3	4,6	5,3
	Na	4,3	4,1	5,3	4,3	6,6	3,3	4,2	3,9	3,7	4,3	4,0	3,7	4,2	4,4	4,3

Analyseer deze uitkomsten. Kies $\alpha = 0{,}05$.

10.8. Uit: BENARD & VAN EEDEN (1956).
Om na te gaan of het aantal trombocyten per mm³ bij konijnen door reserpine wordt beïnvloed, werd dit aantal bij 12 dieren onmiddellijk vóór en 4 uur na toediening van reserpine bepaald:

Na	128	153	115	159	189	211	177	147	126	89	146	221
Voor	90	136	151	115	154	200	157	148	133	80	150	181
N — V	+38	+17	—36	+44	+35	+11	+20	—1	—7	+9	—4	+40

Toets de hypothese, dat de reserpine geen invloed heeft op het aantal trombocyten. Kies $\alpha = 0{,}05$.

10.4. De toets van Friedman [1]

10.4.1. TOEPASSING

Deze toets kan worden toegepast wanneer men beschikt over k verwante waarnemingsreeksen (zie 8.2.1) en de nulhypothese wil toetsen dat voor elk element de waarnemingen een aselecte volgorde vertonen. De toets is gevoelig voor een overeenkomstig verloop van de waarden bij het merendeel der elementen [2].

Voorbeeld

10.8. Men heeft een drietal analysten, onafhankelijk van elkaar, het hemoglobinegehalte van zes bloedmonsters met dezelfde sicca-hemometer laten bepalen. De uitkomsten van dit onderzoekje zijn in tabel 10.6 opgenomen. De te toetsen hypothese H_0 luidt dat de waargenomen verschillen tussen de uitkomsten der analysten toevallig zijn.

Tabel 10.6. Hemoglobinebepalingen (in g%) van drie analysten

Monster	Analyst			Rangorde		
	A	*B*	*C*	*A*	*B*	*C*
1	14,3	14,5	14,2	2	3	1
2	12,4	13,4	12,3	2	3	1
3	13,9	13,8	13,6	3	2	1
4	15,4	15,2	15,1	3	2	1
5	10,1	9,8	9,9	3	1	2
6	14,9	15,0	14,7	2	3	1
Aantal kolommen = $k = 3$				Rangsom	15	14	7
Aantal rijen = $n = 6$					S_1	S_2	S_3

10.4.2. DE TOETSINGSGROOTHEID

Men kent aan de op één rij staande k verwante waarnemingen van laag naar hoog de rangnummers $1, 2, \ldots, k$ toe. Vervolgens bepaalt men van elke kolom de som van de daarin voorkomende rangnummers (= de rangsom) S_j. De som van de op één rij staande rangnummers is

[1] Ook genoemd: de methode van de m rangschikkingen.
[2] Wanneer men wil onderzoeken of een overwegend stijgend of dalend verloop aanwezig is kan men gebruik maken van de toets in 12.2.

$k(k+1)/2$, zodat de som van alle rangnummers $nk(k+1)/2$ bedraagt. Het gemiddelde van de k sommen S_j is

$$\bar{S} = \frac{nk(k+1)/2}{k} = \frac{n(k+1)}{2}.$$

Als toetsingsgrootheid kiest men

$$(10.6)\quad K = \sum_{j=1}^{k} (S_j - \bar{S})^2 = \sum_{j=1}^{k} S_j^2 - k\bar{S}^2 = \sum_{j=1}^{k} S_j^2 - \frac{kn^2(k+1)^2}{4}.$$

Bij voorbeeld 10.8 is $\bar{S} = 6(3+1)/2 = 12$, zodat

$$K_0 = (15-12)^2 + (14-12)^2 + (7-12)^2 = 38.$$

10.4.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

Onder H_0 zijn voor iedere rij (rangschikking) alle $k!$ permutaties van de k rangnummers even waarschijnlijk. De rijen zijn onderling onafhankelijk, zodat bijv. de kansverdeling van $\underline{K}$ onder H_0 voor $k = 3$ en $n = 2$ als volgt kan worden bepaald:

Rij 1	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3	1 2 3
Rij 2	1 2 3	1 3 2	2 1 3	2 3 1	3 1 2	3 2 1
S_j	2 4 6	2 5 5	3 3 6	3 5 4	4 3 5	4 4 4
$\bar{S} = \frac{2\times 4}{2} = 4$	4 4 4	4 4 4	4 4 4	4 4 4	4 4 4	4 4 4
$S_j - \bar{S}$	−2 0 +2	−2 +1 +1	−1 −1 +2	−1 +1 0	0 −1 +1	0 0 0
$K = \Sigma(S_j - \bar{S})^2$	8	6	6	2	2	0

Verdeling van $\underline{K}$ onder H_0 voor $k = 3$ en $n = 2$

K	0	2	6	8	Totaal
$P(K)$	$^1/_6$	$^2/_6$	$^2/_6$	$^1/_6$	1

De verdeling van $\underline{K}$ voor $k = 3$ en $n = 3$ kan nu worden verkregen, door bij *elk* drietal S_j's uit de voorgaande opstelling alle permutaties van de 3e rij rangnummers uit te schrijven. De *eerste* reeks S_j's (2, 4, 6) levert dan op:

S_j, rij 1 en 2	2 4 6	2 4 6	2 4 6	2 4 6	2 4 6	2 4 6
Perm. rij 3	1 2 3	1 3 2	2 1 3	2 3 1	3 1 2	3 2 1
$S_j (n=3)$	3 6 9	3 7 8	4 5 9	4 7 7	5 5 8	5 6 7
$\bar{S} = \frac{3\times 4}{2} = 6$	6 6 6	6 6 6	6 6 6	6 6 6	6 6 6	6 6 6
$S_j - \bar{S}$	−3 0 +3	−3 +1 +2	−2 −1 +3	−2 +1 +1	−1 −1 +2	−1 0 +1
$K = \Sigma(S_j - \bar{S})^2$	18	14	14	6	6	2

Doet men hetzelfde voor de vijf overige reeksen, dan verkrijgt men de volgende verdeling van $\underline{K}$ onder H_0:

K	0	2	6	8	14	18	Totaal
$P(K)$	$^2/_{36}$ 0,056	$^{15}/_{36}$ 0,417	$^6/_{36}$ 0,167	$^6/_{36}$ 0,167	$^6/_{36}$ 0,167	$^1/_{36}$ 0,028	1 1,002

Het is duidelijk dat $\underline{K}$ de waarde 0 aanneemt, als alle S_j gelijk (dus gelijk aan $\bar{S}$) zijn. Naarmate de S_j meer verschillen neemt $\underline{K}$ een hogere waarde aan. De maximale waarde wordt bereikt, als de rangschikkingen per rij een volkomen overeenstemming vertonen. Beschouw bijv. het onderstaande schema, waarin B steeds de hoogste en C steeds de laagste waarnemingen heeft opgeleverd. De rangsommen zijn dan resp. $2n$, $3n$ en n en men vindt $K_0 = (6-6)^2 + (9-6)^2 + (3-6)^2 = 18$.

Rangnummerschema
bij volkomen overeenstemming

		$k=3$		
		A	B	C
$n=3$	1	2	3	1
	2	2	3	1
	3	2	3	1
	S_j	6	9	3

Als echter H_0 onjuist is kan men juist uitkomsten van dit type verwachten. Bij het gebruik van de toets van FRIEDMAN zal men dus uitsluitend bij hoge waarden van K tot het verwerpen van H_0 willen overgaan, zodat rechts éénzijdig wordt getoetst.

10.4.4. UITVOERING VAN DE TOETSING ZONDER GELIJKEN

Exacte toetsing, tabel K-1. Voor $k = 3$ en $n = 2$ t/m 10, voor $k = 4$ en $n = 2$ t/m 6 en voor $k = 5$, $n = 3$ kan men voor een aantal waarden van K de exacte rechtse overschrijdingskans, indien deze kleiner is dan ongeveer 0,15, in tabel K-1 aflezen.

Bij voorbeeld 10.8 is $k = 3$, $n = 6$ en $K_0 = 38$. Uit tabel K-1 blijkt, dat de exacte kans in dit geval 0,052 bedraagt.

Kritieke waarden, tabel K-2. In tabel K-2 zijn voor $k \leqslant 15$ en $n \leqslant 15$ de rechter kritieke waarden K_R van K bij een onbetrouwbaarheidsdrempel 0,05 opgenomen. Men kan dus bij deze drempel tot het verwerpen van H_0 besluiten indien $K_0 \geqslant K_R$ is.

χ^2-benadering. Voor grotere waarden van k en n en bij toetsing met andere drempelwaarden kan men gebruik maken van het feit, dat de grootheid

$$(10.7) \qquad \underline{G} = \frac{12\underline{K}}{nk(k+1)} = \frac{12}{nk(k+1)} \sum_{j=1}^{k} \underline{S_j}^2 - 3n(k+1)$$

bij benadering een χ^2-verdeling volgt met $k-1$ vrijheidsgraden. Bij het gebruik van deze benadering kan dus H_0 bij een drempelwaarde α worden verworpen indien

$$G_0 \geqslant \chi^2_{1-\alpha}(\nu = k-1).$$

Bij toepassing van de χ^2-benadering op voorbeeld 10.8 vindt men: $G_0 = 12 \times 38/6 \times 3 \times 4 = 6{,}33$, met een overschrijdingskans die iets kleiner is dan 0,05 [$\chi^2_{0,95}(2) = 5{,}991$]. Ondanks de kleine waarden van k en n is de benadering reeds bevredigend.

Opmerking

Er is voor kleine k en n een nauwkeuriger benadering mogelijk via de F-verdeling (die in hoofdstuk 14 behandeld wordt). Deze wordt o.m. gegeven door KENDALL (1948), VAN ELTEREN (1951) en MORONEY (1953). De χ^2-benadering blijkt dan nl. in het algemeen een iets te hoge overschrijdingskans op te leveren, d.w.z. zij blijft, wat fouten van de eerste soort betreft, aan de veilige kant. Als n relatief groot is t.o.v. k (een situatie die in de praktijk veel zal voorkomen) is de χ^2-benadering in het algemeen beter dan deze F-benadering [zie VAN ELTEREN (1951)].

10.4.5. UITVOERING VAN DE TOETSING MET GELIJKEN

Als bij een rangschikking gelijken optreden kan men deze het gemiddelde toekennen van de rangnummers, die zij zouden moeten bezetten. Als er niet te veel gelijken zijn blijft men bij gebruik van tabel K t.a.v. het verwerpen van H_0 aan de veilige kant. Past men de

Tabel 10.7. Toepassing van de toets van FRIEDMAN als veel gelijken optreden

Patiënt	Bereidingswijze			
	A	*B*	*C*	*D*
1	3,5	3,5	1,5	1,5
2	4	2	2	2
3	1,5	3,5	3,5	1,5
4	3,5	3,5	1,5	1,5
5	3	3	3	1
6	3	3	1	3
7	3,5	1,5	3,5	1,5
8	2,5	2,5	2,5	2,5
9	3	3	1	3
10	2,5	2,5	2,5	2,5
11	3,5	1,5	3,5	1,5
12	4	2	2	2
13	2,5	2,5	2,5	2,5
14	3	3	3	1
15	4	2	2	2
S_j	47	39	35	29
$\bar{S}$	37,5	37,5	37,5	37,5
$S_j - \bar{S}$	9,5	1,5	−2,5	−8,5

Gelijken

Omvang v/d groep t_i	f_i	Controle $f_i t_i$	$f_i t_i^3$
1	7	7	7
2	10	20	80
3	7	21	189
4	3	12	192
	$kn = 60$		$D = 468$

$$K = \Sigma(S_j - \bar{S})^2 = (9{,}5)^2 + (1{,}5)^2 + \ldots + (8{,}5)^2 = 171.$$

Formule 10.8:

$$G_{(3)} = \frac{12(4-1)171}{15(4)^3 - 468} = 12{,}51.$$

(Men kan bij de berekening de 3 onbesliste gevallen (uitsluitend nullen of enen) ook weglaten, maar dient dan in formule (10.8) voor n te nemen $15 - 3 = 12$.)

De χ^2-benadering van de toets van FRIEDMAN volgens formule (10.8) is —wanneer de waarnemingen een dichotomie betreffen — identiek aan de χ^2-benadering van de toets van COCHRAN volgens formule (9.11).

χ^2-benadering toe, dan kan een correctie voor gelijken worden aangebracht door in plaats van (10.7) te nemen

$$G = \frac{12(k-1)K}{nk^3 - D} \tag{10.8}$$

waarin $D = \Sigma t_i^3$ en t_i de omvang is van de i-de groep gelijken (hierbij is weer $t = 1$ voor elke waarneming die niet tot een groep gelijken behoort). Ook in dit geval is de betekenis van de correctie doorgaans gering, zodat deze zonder bezwaar achterwege kan worden gelaten als de verkregen uitkomst zonder correctie significant of duidelijk insignificant is.

Voor een beschouwing betreffende de exacte verdeling van K bij het voorkomen van gelijken verwijzen wij naar VAN ELTEREN (1951).

Voorbeeld

10.9. De waarnemingen in tabel 9.6 zijn kwalitatief, zodat per patiënt betrekkelijk veel gelijken optreden. Het rangnummerschema dat men bij toepassing van de toets van FRIEDMAN verkrijgt is in tabel 10.7 opgenomen. Voor de berekening van D kan men het best een afzonderlijk tabelletje opstellen.

10.4.6. OPGAVEN

10.9. Uit: N.T.v.G. 102 (1958), 274.

De adrenalineproef geeft een inzicht in de kracht (uitgedrukt in mm Hg) waarmee de capillairen van de huid zich kunnen samentrekken. Bij vier personen in syncope kon aanstonds nadat zij waren gaan liggen, deze proef worden verricht. De uitkomsten waren:

Proefpersoon	Vóór syncope	Aanstonds na syncope	2 uur later	4 uur later
1	65	20	60	70
2	75	20	30	55
3	80	25	45	60
4	70	30	65	70

Toets de hypothese dat de waargenomen verschillen toevallig zijn. Kies $\alpha = 0{,}05$.

10.10. Uit: N.T.v.G. 100 (1956), 3065.

Nevenstaande tabel geeft de aantallen bloedplaatjes (volgens Fonio) bij voldragen kinderen op de 1e, 3e en 5e levensdag. Welke hypothese kan worden getoetst? Voer de toetsing uit met $\alpha = 0{,}05$.

	1e dag	3e dag	5e dag
1.	195.860	150.000	152.600
2.	219.000	212.000	244.800
3.	235.600	232.600	228.800
4.	148.980	183.600	161.280
5.	173.900	186.200	173.900
6.	149.300	173.800	152.800
7.	171.440	169.860	179.580
8.	284.480	206.400	231.840
9.	280.000	244.800	248.160
10.	161.000	169.740	194.000

10.11. Uit N.T.v.G. 98(1954), 1745. In onderstaande tabel staan de penicillinespiegels per cm^3 plasma van 8 patiënten op 7 tijdstippen. De eerste inspuiting van 300.000 I.E. almocilline werd gegeven terstond na de bloedwinning op dag 0 (waarop alle spiegels 0 waren), de tweede na de bloedwinning op dag 1, enz. Toets de hypothese, dat de verschillen tussen de dagen op toeval berusten (kies hierbij $\alpha = 0{,}05$).

Pat.	Spiegels op dag						
	1	2	3	4	5	6	7
39	0,20	0,11	0,31	0,18	0,26	0,16	0,15
40	0,09	0,12	0,15	0,17	0,16	0,19	0,19
41	0,38	0,14	> 1	0,20	0,14	0,19	0,22
42	0,07	0,11	> 1	0,12	0,12	0,16	0,22
43	0,00	0,12	0,25	0,29	0,25	0,31	0,36
44	0,00	0,11	0,22	0,35	0,18	0,25	0,21
45	0,00	0,05	0,10	0,11	0,16	0,18	0,15
46	0,26	1,00	0,52	0,27	0,39	0,37	0,36

10.12. Uit ARMITAGE, Course in the Application of Statistical Methods to Medicine and Epidemiology, London School of Hygiene and Tropical Medicine.

'In a series of assays of pertussis vaccines, three vaccines were tested on each of ten days. The responses given are estimates of the logarithms of the doses of vaccine (in millions of organisms) required to protect 50% of mice against a subsequent infection with pertussis organisms. Test (with $\alpha = 0{,}05$) whether the differences between days and between vaccines are significant.'

Day	Vaccine		
	A	*B*	*C*
1	2,64	2,93	2,93
2	2,00	2,52	2,56
3	3,04	3,05	3,35
4	2,07	2,97	2,55
5	2,54	2,44	2,45
6	2,76	3,18	3,25
7	2,03	2,30	2,17
8	2,20	2,56	2,18
9	2,38	2,99	2,74
10	2,42	3,20	3,14

10.5. De toets van Wilcoxon

10.5.1. TOEPASSING

Deze toets kan worden toegepast als men beschikt over twee aselecte steekproeven, resp. twee aselect samengestelde groepen, en de nulhypothese wil toetsen dat zij afkomstig zijn uit populaties met dezelfde verdeling, resp. uit dezelfde populatie. De toets is voornamelijk gevoelig voor verschillen in niveau van de vergeleken verdelingen.

Voorbeeld

10.10. Een steekproef van 8 ratten (van hetzelfde geslacht, van de zelfde leeftijd en met ongeveer hetzelfde gewicht) wordt aselect in twee even grote subgroepen P en C gesplitst. Groep P ondergaat een zekere behandeling, groep C dient als controle. De dieren ontvangen hierna gedurende zes weken dezelfde voeding, daarna worden zij gedood.

Na verwijdering van kop en staart worden zij gevild, waarna karkasanalyse wordt verricht. De percentages vet in het karkas zijn in tabel 10.8 vermeld. De te toetsen hypothese luidt dat deze percentages toevallig verschillen. Deze proef is bedoeld als een oriënterend onderzoek, dat men op grotere schaal wil voortzetten indien de behandeling tot een vermeerdering van het vetgehalte blijkt te leiden (er wordt dus rechts éénzijdig getoetst). Men kiest $\alpha = 0{,}05$.

Tabel 10.8. Het vetgehalte van het karkas bij behandelde ratten (P) en controle-ratten (C)

Groep P		Groep C	
1	Rangorde R_1	2	Rangorde R_2
35	7	20	3
28	6	16	2
21	4	22	5
49	8	14	1
$n_1 = 4$	$\Sigma R_1 = 25$	$n_2 = 4$	$\Sigma R_2 = 11$

10.5.2. DE TOETSINGSGROOTHEID

Men kent aan de waarnemingen, onafhankelijk van de steekproef waarin zij voorkomen, van laag naar hoog de rangnummers 1, 2, . . ., n toe ($n = n_1 + n_2$). Bij voorbeeld 10.10 zijn dit dus de rangnummers 1 t/m 8, die in de kolommen R_1 en R_2 zijn opgenomen. Als controle kan dienen dat $\Sigma R_1 + \Sigma R_2 = \frac{1}{2}n(n + 1)$.

Als toetsingsgrootheid kiezen wij de som van de rangnummers in steekproef 1, $\Sigma R_1 = S$.

Opmerking

In verschillende publicaties betreffende de toets van WILCOXON wordt de door MANN & WHITNEY (1947) ingevoerde toetsingsgrootheid U of de door WABEKE en VAN EEDEN (1955) ingevoerde grootheid $W = 2U$ gebruikt. De met S_0 corresponderende waarde U_0 kan voor een bepaald steekproevenpaar als volgt worden bepaald: rangschik de waarnemingen van links naar rechts van laag naar hoog, onderstreep de waarnemingen van steekproef 2 en bepaal vervolgens voor iedere waarneming in steekproef 1 hoeveel waarnemingen van steekproef 2 rechts daarvan staan (dus: hoger zijn). De som van de op deze wijze verkregen getallen is dan U_0. Voorbeeld 10.10 levert op:

Steekproef	2	2	2	1	2	1	1	1	$n_1 = n_2 = 4$
Gerangschikte waarnemingen	14	16	20	21	22	28	35	49	
Bijdragen tot U				1	+	0 +	0 +	0	$= 1 = U_0$

Hieruit volgt: $W_0 = 2U_0 = 2$. Wil men W_0 rechtstreeks uit de waarnemingen berekenen, dan dient men bij het voorgaande procédé bij iedere 'één' *tweemaal* het aantal rechts daarvan staande 'tweeën' te nemen. Tussen de grootheden W en S bestaat de relatie

(10.9) $$W = 2n_1n_2 + n_1(n_1 + 1) - 2S.$$

10.5.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

De kansverdeling van de grootheid $\underline{S}$ onder H_0 kan worden bepaald op de wijze, die in voorbeeld 3.17 is aangegeven. Er zijn, bij het trekken van 4 uit de 8 rangnummers, in totaal $C_4^8 = 70$ verschillende steekproeven mogelijk. De laagste waarde wordt verkregen bij trekking van de rangnummers 1, 2, 3 en 4 ($S = 10$) en deze uitkomst kan slechts op één manier tot stand komen. De uitkomst $S = 11$ kan eveneens slechts op één manier worden bereikt, nl. bij trekking van 1, 2, 3 en 5. De uitkomst $S = 12$ kan op twee manier ontstaan ($1 + 2 + 3 + 6$ en $1 + 2 + 4 + 5$), enz. De kansverdeling die men zo verkrijgt staat in tabel 10.9.

Tabel 10.9. Verdeling van $\underline{S}$ onder H_0 voor $n_1 = n_2 = 4$

S	f	$P(S = S)$
26	1	0,0143
25	1	0,0143
24	2	0,0286
23	3	0,0429
22	5	0,0714
21	5	0,0714
20	7	0,1000
19	7	0,1000
18	8	0,1143
17	7	0,1000
16	7	0,1000
15	5	0,0714
14	5	0,0714
13	3	0,0429
12	2	0,0286
11	1	0,0143
10	1	0,0143

De rechts éénzijdige overschrijdingskans van de waargenomen rangsom in voorbeeld 10.10, $S_0 = 25$, blijkt $2(0{,}0143) = 0{,}0286$ te bedragen. H_0 kan dus bij de gekozen 5% drempel worden verworpen.

Als geen gelijken optreden bezit de verdeling van $\underline{S}$ onder H_0 de volgende eigenschappen:

a. De verdeling is discreet en neemt met 1 opklimmende de waarden van $n_1(n_1 + 1)/2$ tot en met $n_1n_2 + n_1(n_1 + 1)/2$ aan.

b. De verdeling is symmetrisch t.o.v.

$$\mu_S = \frac{n_1(n + 1)}{2}. \tag{10.10}$$

c. De variantie van de verdeling is:

$$\sigma_S^2 = \frac{n_1n_2(n + 1)}{12} \qquad (n = n_1 + n_2). \tag{10.11}$$

10.5.4. UITVOERING VAN DE TOETSING ZONDER GELIJKEN

Het gebruik van tabel L. In tabel L vindt men voor $n_1 \leqslant n_2 \leqslant 25$ en de drempelwaarden 0,01, 0,02, 0,05 en 0,10 de linker en rechter kritieke waarden van S bij *tweezijdige toetsing*. Voert men een tweezijdige toetsing uit met onbetrouwbaarheid α, dan kan H_0 worden verworpen indien

$$S_0 \leqslant S_L(\alpha, n) \text{ of } S_0 \geqslant S_R(\alpha, n).$$

Voor $n_1 = 8$ en $n_2 = 6$ leest men bijv. uit tabel L – 2 af, dat bij tweezijdige toetsing met $\alpha = 0{,}05$ voor $S_0 \leqslant 21$ of $S_0 \geqslant 59$ tot het verwerpen van H_0 kan worden besloten.

Bij *éénzijdige toetsing* dient men de onbetrouwbaarheidsdrempels in tabel L te halveren. Toetst men H_0 met als alternatieve hypothese dat de waarden in steekproef 1 systematisch lager zijn dan die in steekproef 2 $[H_1 : P(\underset{\sim}{x}_1 > \underset{\sim}{x}_2) < \frac{1}{2}]$ dan dient links éénzijdig te worden getoetst, zodat men bij een drempelwaarde α tot het verwerpen van H_0 ten gunste van H_1 zal overgaan voor

$$S_0 \leqslant S_L(2\alpha, n).$$

Luidt de alternatieve hypothese dat de waarden in steekproef 1 systematisch hoger zijn dan die in steekproef 2 $[H_1 : P(\underset{\sim}{x}_1 > \underset{\sim}{x}_2) > \frac{1}{2}]$ dan wordt rechts éénzijdig getoetst en besluit men tot het verwerpen van H_0 voor

$$S_0 \geqslant S_R(2\alpha, n).$$

Voorbeeld

10.11. Men wil met $\alpha = 0{,}05$ de hypothese toetsen dat twee slaapmiddelen, X en Y, indien daarvan een bepaalde dosis bij muizen wordt ingespoten, de slaaptijd van deze dieren op dezelfde wijze beïnvloeden. Men splitst een groep van 16 muizen aselect in twee even grote subgroepen en spuit deze resp. met X en Y in. De slaaptijden der dieren (in minuten) blijken te zijn:

Middel X,	5	$6\frac{1}{2}$	10			$11\frac{1}{2}$	13		17			$20\frac{1}{2}$	21				
Middel Y,					$10\frac{1}{2}$			16			$17\frac{1}{2}$	20			24	25	$28\frac{1}{2}$	$35\frac{1}{2}$
Rangnummers I	1	2	3			5	6		8			11	12				
Rangnummers II				4				7		9	10			13	14	15	16

De som van de rangnummers in groep I is $S_0 = 48$. Er dient tweezijdig te worden getoetst en uit tabel L – 2 blijkt dat voor $n_1 = n_2 = 8$ en $\alpha = 0{,}05$ de linker en rechter kritieke waarden van S resp. gelijk zijn aan 49 en 87. Daar S_0 beneden de linker kritieke waarde ligt wordt de nulhypothese verworpen, ten gunste van de alternatieve hypothese, dat middel X tot slaaptijden leidt die systematisch lager zijn dan die van middel Y.

De normale benadering. Voor grotere waarden van n kan, mits n_1 en n_2 niet te veel verschillen, gebruik worden gemaakt van het feit dat de grootheid

$$(10.12) \qquad \underset{\sim}{T} = \frac{\underset{\sim}{S} - \mu_S}{\sigma_S} = \frac{\underset{\sim}{S} - n_1(n+1)/2}{\sqrt{\frac{n_1 n_2 (n+1)}{12}}}$$

bij benadering standaardnormaal verdeeld is. Daar de afstand tussen de opeenvolgende waarden van deze toetsingsgrootheid gelijk is aan 1, dient de aan te brengen continuïteitscorrectie $\frac{1}{2}$ te bedragen. Bij gebruik van de normale benadering berekent men dus

(10.13) $T_0 = \dfrac{|\,S_0 - \mu_S\,| - \frac{1}{2}}{\sigma_S}$ en $P_D(T_0)$ bij tweezijdige toetsing,

of

(10.14.1) $T_0 = \dfrac{S_0 - \mu_S + \frac{1}{2}}{\sigma_S}$ en $P_L(T_0)$ bij links éénzijdige toetsing,

resp.

(10.14.2) $T_0 = \dfrac{S_0 - \mu_S - \frac{1}{2}}{\sigma_S}$ en $P_R(T_0)$ bij rechts éénzijdige toetsing.

H_0 wordt verworpen indien de gevonden overschrijdingskans kleiner is dan α.

Voorbeelden

10.12. Past men ondanks de kleine waarden van $n_1 = n_2 = 8$ de normale benadering toe op voorbeeld 10.11 dan vindt men

$$T_0 = \frac{|\,48 - 8(17)/2\,| - \frac{1}{2}}{\sqrt{8 \times 8 \times 17/12}} = -2{,}05, \quad P_D = 0{,}04 \qquad \text{(tabel A)}.$$

Deze toetsing leidt dus eveneens tot het verwerpen van H_0 bij een 5% drempel. De benadering is reeds zeer bevredigend, daar de exacte tweezijdige overschrijdingskans 0,038 blijkt te zijn.

10.13. Stel dat men voor $n_1 = 9$ en $n_2 = 10$ vindt dat $S_0 = 115$. Men toetst de hypothese H_0 met als alternatieve hypothese $P(\bar{x}_1 > \bar{x}_2) > \frac{1}{2}$, zodat rechts éénzijdige toetsing dient plaats te vinden. De normale benadering geeft

$$T_0 = \frac{|\,115 - 9(20)/2\,| - \frac{1}{2}}{\sqrt{9 \times 10 \times 20/12}} = 2{,}00, \quad P_R = 0{,}0228 \qquad \text{(tabel A)}.$$

Ook hier hebben wij kleine steekproeven genomen om de benaderende uitkomst met de exacte kans te kunnen vergelijken. Daar laatstgenoemde 0,0217 is blijkt de benadering uitstekend te zijn.

10.5.5. UITVOERING VAN DE TOETSING MET GELIJKEN

Wanneer n klein is en er veel gelijken voorkomen, kan men noch van tabel L noch van de normale benadering gebruik maken. In WABEKE en VAN EEDEN (1955) is aangegeven, hoe men dan de exacte verdeling van S kan berekenen. Treden er slechts twee groepen gelijke waarnemingen op, dan verkrijgt men een 2×2 tabel en gaat de toets van WILCOXON over in de toets van FISHER.

Als n_1 en n_2 niet te klein zijn en onderling niet te veel verschillen en als bovendien de groepen gelijke waarnemingen niet te veel in omvang verschillen kan men aan de gelijken binnen een groep weer het gemiddelde toekennen van de rangnummers die zij moeten bezetten. Men kan dan de normale benadering toepassen met in plaats van σ_S volgens (10.11) σ_S' volgens

(10.15) $$\sigma_S'^2 = \frac{n_1 n_2 (n^3 - D)}{12n(n-1)},$$

waarin $D = \Sigma t_i^3$, en $t_1, t_2, \ldots, t_k$ de omvang is van de 1e, 2e, ..., ke groep gelijken (voor elke uitkomst die niet tot een groep gelijken behoort neemt men $t = 1$). Bevatten beide steekproeven geen gelijke waarnemingen dan wordt $D = n$, zodat $n^3 - D = n^3 - n = n(n^2 - 1)$ en (10.15) in (10.11) overgaat. De afstand tussen de opeenvolgende waarden van $\underline{S}$ is bij het optreden van gelijken in het algemeen niet constant, maar neemt men een continuïteitscorrectie ter grootte van 1 dan blijft men aan de veilige kant t.a.v. het verwerpen van H_0.

Met (10.15) vindt men voor de spreiding van $\underline{S}$ een kleinere waarde dan met (10.11). Houdt men dus bij de toepassing van de normale benadering geen rekening met de gelijken, dan vindt men een te grote overschrijdingskans. Is deze kleiner dan of gelijk aan α, dan wordt H_0 verworpen en behoeft men niet tot berekening van D over te gaan.

10.5.6. REKENSCHEMA

Wanneer het aantal gelijken (vrij) groot is kan men S_0 berekenen via een rekenschema, waarbij wordt uitgegaan van de niet-gegroepeerde frequentieverdelingen van de twee steekproeven. Wij bespreken dit schema aan de hand van het voorbeeld in tabel 10.10.

Tabel 10.10. Berekeningsschema voor de toets van WILCOXON bij het optreden van veel gelijken. Gegevens: de frequentieverdelingen van de *cd*-tijd (aantal sec × 100) bij mannen: A = beneden 40 jaar, B = 40 jaar en ouder. Uit H. ELSBACH, diss., Leiden, 1954

x	A f_{i1}	B f_{i2}	Totaal f_i	Somfreq. F_i	$\overline{R}_i$	$f_{i1}\overline{R}_i$	$f_{i2}\overline{R}_i$	$f_i^3 = t^3$
7		1	1	1	1		1	1
8		2	2	3	2,5		5	8
9		2	2	5	4,5		9	8
10	2	2	4	9	7,5	15	15	64
11	3	1	4	13	11,5	34,5	11,5	64
12	10	3	13	26	20	200	60	2197
13	5	7	12	38	32,5	162,5	227,5	1728
14	7	1	8	46	42,5	297,5	42,5	512
15	2	1	3	49	48	96	48	27
16	1		1	50	50	50		1
Tot.	30 n_1	20 n_2	50 n	—	—	855,5 $S_1 = S$	419,5 S_2	4610 D

De berekening verloopt in de volgende stappen:

1. Bepaal de totaalfrequenties f_i en vervolgens de somfrequenties F_i.

2. Bereken uit de somfrequenties de gemiddelde rangnummers door te nemen $\overline{R}_i = (F_{i-1} + F_i + 1)/2$. Zo is bijv. voor $x = 12$: $\overline{R} = (13 + 26 + 1)/2 = 20$.

3. Bereken de som van de rangnummers in de eerste steekproef

$S_1 = S = \Sigma f_{i1}\overline{R}_i$ (= 855,5). Wanneer men ook de som van de rangnummers van de tweede steekproef, S_2, berekent, kan ter controle worden gebruikt dat $S_1 + S_2 = n(n + 1)/2$.

4. Daar de totaalkolom de aantallen gelijken aangeeft, is $D = \Sigma f_i^3$ (= 4610).

5. Pas de normale benadering toe, met σ_S' volgens (10.15) in plaats van σ_S:

$$T_0 = \frac{S - n_1(n+1)/2 \pm \frac{1}{2}}{\sqrt{\frac{n_1 n_2(n^3 - D)}{12n(n-1)}}} = \frac{855{,}5 - 30\,(51)/2 - 0{,}5}{\sqrt{\frac{(30)\,(20)\,[(50)^3 - 4610]}{12(50)\,(49)}}} = 1{,}82,$$

$$P_D(1{,}82) \approx 0{,}07 \text{ (tabel A)}.$$

Gebruik van het schema bij de rangtekentoets. Wanneer veel verschillen met dezelfde absolute waarde optreden, kan men de waarde V_0 van de toetsingsgrootheid V van de rangtekentoets eveneens eenvoudig door middel van het zo juist behandelde schema berekenen. Wanneer men het schema toepast met als eerste steekproef de absolute waarden van de negatieve verschillen en als tweede steekproef de positieve verschillen, dan volgt V_0 uit de rangsommen S_1 en S_2 door te nemen

(10.16) $$V_0 = S_2 - S_1.$$

In tabel 10.11 is een voorbeeld van deze berekeningsmethode gegeven.

Tabel 10.11. Toepassing van het rekenschema bij de rangtekentoets

Waargenomen verdeling van verschillen

v	6	5	4	3	2	1	0	—1	—2	—3	—4	—5	Totaal	Totaal excl. 0
f	1	2	5	5	8	12	9	6	5	3	3	1	60	51

Rekenschema

$\lvert v \rvert$	Negatief f_{i1}	Positief f_{i2}	Totaal f_i	F_i	$\overline{R}_i$	$f_{i1}\overline{R}_i$	$f_{i2}\overline{R}_i$	f_i^3
1	6	12	18	18	9,5	57	114	5832
2	5	8	13	31	25	125	200	2197
3	3	5	8	39	35,5	106,5	177,5	512
4	3	5	8	47	43,5	130,5	217,5	512
5	1	2	3	50	49	49	98	27
6		1	1	51	51		51	1
Tot.	18	33	51 n	—	—	468 S_1	858 S_2	9081 D

$$V_0 = S_2 - S_1 = 858 - 468 = 390$$

Formule (10.15):

$$\sigma_V'^2 = \frac{n(n+1)(2n+1)}{6} - \frac{D-n}{12} = \frac{51(52)(103)}{6} - \frac{9081-51}{12}$$

$$= 44773{,}50, \qquad \sigma_V' = 211{,}6$$

Formule (10.3):

$$T_0 = \frac{390-1}{211{,}6} = 1{,}84, \qquad P_D \approx 0{,}066.$$

10.5.7. OPGAVEN

10.13. ELLIOT en CRISMON (Am. J. Physiol. (1947), 366) vergeleken de calciumgehalten (in milli-equiv. per liter plasma) van twee aselect samengestelde groepen ratten, A: afgekoeld tot de rectale temperatuur 25°C bedroeg, B: niet afgekoeld. De uitkomsten waren:

A: 5,76 5,74 5,23 6,19 5,83
B: 5,46 4,91 5,18 4,91 4,82.

Pas de toets van WILCOXON toe met $\alpha = 0{,}05$.

10.14. CORKHILL (Biochem. J. (1930) 779) geeft de volgende glycogeenpercentages in de spieren van controledieren (A) en dieren ingespoten met insuline (B):

A. 0,19 0,18 0,21 0,30 0,66 0,42 0,08 0,12 0,30 0,27
B. 0,15 0,13 0,00 0,07 0,27 0,24 0,19 0,04 0,08 0,20 0,12.

Pas de toets van WILCOXON toe met $\alpha = 0{,}05$.

10.15. Pas de toets van WILCOXON toe op onderstaande aselecte steekproeven ($\alpha = 0{,}05$):

Steekproef 1: 1,23 1,94 1,15 2,84 0,96 0,25 1,71 1,90 1,14 0,30 3,90
0,18 0,88 2,67 1,49 0,43 0,57 2,93
Steekproef 2: 0,59 0,92 1,56 1,34 0,72 0,22 0,35 1,44 1,70 0,55 0,36
1,15 1,78 0,94 1,09 0,86 0,65 0,52 0,88 0,90

10.16. Uit de kursus 'Summarizing experimental data', Columbia University, U.S.A.

In an article on uric acid clearance in the Journal of Clinical Investigation 22, 1943, SCHAFFER et al. give the following data on plasma uric acid (mgm per cent) in 10 antepartum cases with a diagnosis of pre-eclampsia and 12 who were clinically normal:

Normal pregnancy: 4,40 4,43 3,82 3,93 3,51 4,62 2,89 3,91 3,84 2,67
4,68 3,23
Pre-eclampsia: 3,92 5,54 6,92 4,23 4,63 3,95 5,39 5,93 7,12 5,27

Toets de hypothese dat het gehalte bij lijders niet systematisch verschilt van dat bij gezonden.

10.17. SIMONSON c.s. (J. Biol. Chem. (1947), 169) bepaalden de volgende magnesiumgehalten in het serum (mg/100 ml) bij gezonde mannen en vrouwen:

Mannen: 2,0 2,0 2,0 2,1 2,1 2,2 1,9 2,3 2,1 1,9 1,8
2,2 2,0 2,2 2,3 2,0 2,0 2,2 2,0 2,0 1,9
Vrouwen: 1,8 2,1 2,1 2,0 2,2 1,8 1,9 2,2 2,1 2,0 1,8
2,3 2,0 2,0 1,8 1,8 2,1 1,9 1,7 2,1 1,9

Stel de frequentieverdelingen op en pas de toets van WILCOXON toe met behulp van het rekenschema. Kies $\alpha = 0{,}05$.

10.6. De toets van Kruskal & Wallis

10.6.1. TOEPASSING

Deze toets kan worden toegepast wanneer men de nulhypothese wil toetsen, dat k aselecte steekproeven uit populaties met dezelfde verdeling, resp. uit dezelfde populatie, afkomstig zijn. De toets is voornamelijk gevoelig voor verschillen in niveau tussen de te vergelijken verdelingen.

Voorbeeld

10.14. Een onderzoeker meent dat hij in staat is een groep proefdieren door een 'lukrake' keuze aselect in sub-groepen te verdelen. Men voert nu een proef uit, waarbij hij 15 ratten in drie groepen van 5 splitst en bepaalt daarna de gewichten (in gr) van de dieren. De uitkomsten zijn in tabel 10.12 opgenomen. Mag men concluderen dat ten aanzien van de gewichten een aselecte verdeling over de groepen is bereikt? Wegens de kleine omvang van de proef kiest men $\alpha = 0{,}10$.

Tabel 10.12. Gewichten van drie 'lukraak' samengestelde groepen ratten

	1		2		3		
	Gewicht	Rangnr.	Gewicht	Rangnr.	Gewicht	Rangnr.	
	235	1	266	2	326	8	
	296	4	270	3	330	9	
	304	6	298	5	358	13	
	332	10	302	7	360	14	
	342	12	338	11	434	15	
S_j		33		28		59	$\Sigma S_j =$ 120
n_j		5		5		5	$\Sigma n_j = n =$ 15
S_j^2/n_j	$\frac{(33)^2}{5} = 217{,}8$		$\frac{(28)^2}{5} = 156{,}8$		$\frac{(59)^2}{5} = 696{,}2$		$\Sigma S_j^2/n_j = 1070{,}8$

$$H_0 = \frac{12 \times 1070{,}8}{15 \times 16} - 3 \times 16 = 5{,}54$$

10.6.2. DE TOETSINGSGROOTHEID

Evenals bij de toets van WILCOXON kent men aan alle waarnemingen, onafhankelijk van de steekproef waarin zij voorkomen, een rangnummer toe, te beginnen met 1 voor de laagste waarneming t/m $n = \Sigma n_j$ voor de hoogste. In tabel 10.12 is deze rangnummering voor het voorbeeld uitgevoerd. Daarna bepaalt men per steekproef de rangsom S_j. Als toetsingsgrootheid is door KRUSKAL & WALLIS ingevoerd[1]

$$H = \frac{12}{n(n+1)} \sum_{j=1}^{k} \frac{S_j^2}{n_j} - 3(n+1). \tag{10.17}$$

[1] De toets wordt ook de H-toets genoemd.

10.6.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

Als H_0 juist is kan men uitgaan van k steekproeven, die verkregen zijn door de rangnummers 1, 2, . . ., k aselect in k groepen van respectievelijk de omvang $n_1, n_2, \ldots, n_k$ te splitsen. Men kan dan de kansverdeling van de toetsingsgrootheid berekenen door alle $n!/n_1!n_2! \ldots n_k!$ manieren waarop dit kan gebeuren op te schrijven en voor elke manier de waarde van $\underline{H}$ te berekenen. Zo zijn er bijv. voor $n = 5$ en $n_1 = n_2 = 2$, $n_3 = 1$ de volgende $5!/2!2!1! = 30$ mogelijkheden, waarbij wij de rangsommen en de waarden van $\underline{H}$ hebben aangegeven:

	1 3 5	1 3 4	1 4 3	1 2 5	1 2 4	1 4 2	1 2 5	1 2 3	1 3 2
	2 4	2 5	2 5	3 4	3 5	3 5	4 3	4 5	4 5
S_j	3 7 5	3 8 4	3 9 3	4 6 5	4 7 4	4 9 2	5 5 5	5 7 3	5 8 2
H	3,6	3,0	3,6	2,4	1,4	3,0	2,0	0,4	1,4
	1 2 4	1 2 3	1 3 2	2 1 5	2 1 4	2 4 1	2 1 5	2 1 3	2 3 1
	5 3	5 4	5 4	3 4	3 5	3 5	4 3	4 5	4 5
S_j	6 5 4	6 6 3	6 7 2	5 5 5	5 6 4	5 9 1	6 4 5	6 6 3	6 8 1
H	0,6	0	0,6	2,0	0,6	3,6	2,4	0	2,4
	2 1 4	2 1 3	2 3 1	3 1 5	3 1 2	3 2 1	3 1 4	3 1 2	3 2 1
	5 3	5 4	5 4	4 2	4 5	4 5	5 2	5 4	5 4
S_j	7 4 4	7 5 3	7 7 1	7 3 5	7 6 2	7 7 1	8 3 4	8 5 2	8 6 1
H	1,4	0,4	2,0	3,6	0,6	2,0	3,0	1,4	2,4
	4 1 3	4 1 2	4 2 1						
	5 2	5 3	5 3						
S_j	9 3 3	9 4 2	9 5 1						
H	3,6	3,0	3,6						

Voor $n = 5$ en $n_1 = n_2$, $n_3 = 1$ luidt dus de kansverdeling van de grootheid $\underline{H}$ onder H_0:

H	0	0,4	0,6	1,4	2,0	2,4	3,0	3,6
$P(\underline{H} = H)$	2/30	2/30	4/30	4/30	4/30	4/30	4/30	6/30

De grootheid $\underline{H}$ blijkt de waarde 0 aan te nemen, wanneer de rangsommen zich verhouden als de steekproefomvangen (6—6—3). Naarmate de S_j verder van deze verhouding verschillen, neemt $\underline{H}$ een grotere waarde aan. Hieruit volgt dat rechts éénzijdig wordt getoetst.

Voor grotere waarden van n wordt het berekenen van de exacte kansverdeling van $\underline{H}$ zeer tijdrovend. De toetsingsgrootheid is echter zo gekozen dat weer met een χ^2-benadering kan worden gewerkt.

10.6.4. UITVOERING VAN DE TOETSING ZONDER GELIJKEN

Exacte toetsing. In tabel M zijn voor $k = 3$ en elke $n_j \leqslant 5$ de exacte rechtse overschrijdingskansen van een aantal waarden van $\underline{H}$ onder H_0

opgenomen. Bij voorbeeld 10.14 vonden wij dat $H_0 = 5{,}54$ en uit tabel M blijkt dat de rechtse kans van deze uitkomst tussen 0,051 en 0,100 ligt (voor $n_1 = n_2 = n_3 = 5$ heeft immers de waarde 5,66 een rechtse kans van 0,051 en de waarde 4,56 een rechtse kans van 0,100).

χ^2-benadering. Voor $n_j > 5$ blijkt de grootheid $\underline{H}$ onder H_0 bij benadering een χ^2-verdeling met $k - 1$ vrijheidsgraden te volgen. Deze benadering is beter naarmate de omvang der steekproeven groter is. Men kan de nulhypothese bij een drempelwaarde α verwerpen, indien

$$H \geqslant \chi^2_{1-\alpha}(\nu = k - 1).$$

10.6.5. UITVOERING VAN DE TOETSING MET GELIJKEN

Als gelijke waarnemingen voorkomen krijgen deze zoals gebruikelijk per groep het gemiddelde van de rangnummers die zij moeten bezetten. Voor $k = 3$ en $n_j \leqslant 5$ kan men, als het aantal gelijken gering is, tabel M nog wel bij benadering gebruiken. Voor $k \geqslant 3$ en kleine steekproeven met veel gelijken zou men feitelijk de exacte kansverdeling van $\underline{H}$ onder H_0 moeten uitrekenen. Als echter de steekproeven niet te klein zijn en er niet te veel groepen gelijken van niet te sterk verschillende omvang voorkomen kan men weer een χ^2-benadering gebruiken, waarbij dan in plaats van formule (10.7) wordt genomen

$$H = \frac{3(n-1)\,[4\Sigma S_j^2/n_j - n(n+1)^2]}{n^3 - D}, \tag{10.18}$$

waarin $D = \Sigma t_i^3$ en $t_i =$ de omvang van de i-de groep gelijken. Treden geen gelijken op, dan is $D = n$ en gaat (10.18) over in (10.17).

Wanneer gelijken voorkomen is de met (10.18) berekende waarde van $\underline{H}$ steeds groter dan de waarde die men met (10.17) vindt. Als H_0 reeds op grond van laatstgenoemde waarde kan worden verworpen, kan men de berekening van D en die van H_0 via (10.18) zonder bezwaar achterwege laten.

Voorbeeld

10.15. In tabel 10.13 is een voorbeeld met vrij veel gelijken uitgewerkt, waarbij een rekenschema analoog aan dat voor de toets van WILCOXON (zie 10.5.6) is gebruikt. De gegevens zijn ontleend aan W.A. HERRMANN (Diss., Leiden, 1956) en zij betreffen de werkzaamheid van 'barrier creams' op de huid van de handen van arbeiders, die frequent in contact komen met een zuuroplossing. De grootheid x betreft de som van de verbeteringen en verslechteringen van de toestand van de huid voor 5 criteria. De drie verschillende creams ($A =$ Lanoline, Vaseline $\overline{aa}$, $B =$ Kerodex 71, $C = W/o$ basis met 50% siliconolie) werden aan de arbeiders van drie ploegen verstrekt.

Tabel 10.13. Toepassing van de toets van KRUSKAL en WALLIS op gegevens met vrij veel gelijken

x	A	B	C	Totaal	F_i	$\overline{R_i}$	$f_{i1}\overline{R_i}$	$f_{i2}\overline{R_i}$	$f_{i3}\overline{R_i}$	f_i^3
	f_{i1}	f_{i2}	f_{i3}	f_i						
—2	1			1	1	1	1			1
—1,5		1		1	2	2		2		1
—1	2	2	1	5	7	5	10	10	5	125
—0,5	2		1	3	10	9	18		9	27
0	4	2	2	8	18	14,5	58	29	29	512
+0,5	3	5	3	11	29	24	72	120	72	1331
+1	4	3	2	9	38	34	136	102	68	729
+1,5		3	1	4	42	40,5		121,5	40,5	64
+2			1	1	43	43			43	1
+2,5		1		1	44	44		44		1
+3		1	3	4	48	46,5		46,5	139,5	64
+3,5			1	1	49	49			49	1
Totaal	16	18	15	49	—	—	295	475	455	2857
	n_1	n_2	n_3	n			S_1	S_2	S_3	D

$\Sigma S_j^2/n_j = (295)^2/16 + (475)^2/18 + (455)^2/15 = 31775{,}4514$

Formule (10.18):

$$H_0 = \frac{3(49-1)\,[4(31775{,}4514) - 49(50)^2]}{(49)^3 - 2857} = 5{,}77$$

De waarde $H_0 = 5{,}77$ kan bij benadering worden beschouwd als een waarde van χ^2 met 2 vrijheidsgraden en uit tabel D blijkt dat $\chi^2_{0,95}(\nu = 2) = 5{,}991$. De rechtse kans van H_0 is dus iets groter dan 0,05. Berekent met H_0 met formule (10.17) dan vindt men een iets lagere uitkomst, n.l. $H_0 = 5{,}63$.

10.6.6. $k = 2$: OVERGANG IN DE TOETS VAN WILCOXON

Past men de toets toe met $k = 2$ dan is H_0 volgens (10.17), resp. (10.18) gelijk aan de T_0^2 die men verkrijgt bij het gebruik van de normale benadering van de toets van WILCOXON zonder continuïteitscorrectie.

10.6.7. OPGAVEN

10.18. Uit de kursus Biostatistics 2, John Hopkins University, U.S.A.
De onderstaande tabel geeft de vleugelbreedten (in mm) naar vindplaats van een op bomen levende sprinkhaansoort (katydid). Toets de hypothese, dat de vleugelbreedte onafhankelijk is van de vindplaats. Kies $\alpha = 0{,}05$.

Vindplaats														
Eik					*Esdoorn*					*Populier*				
8,8	6,6	9,1	9,9	9,2	9,4	5,2	10,7	12,6	8,8	14,1	9,3	12,1	12,5	12,4

10.19. De volgende gegevens betreffen het fosforgehalte (in mg%) in het serum van drie aselect gekozen groepen konijnen, die een verschillende behandeling hebben ondergaan. Toets de hypothese dat deze behandeling geen invloed op het fosforgehalte heeft uitgeoefend. Kies $\alpha = 0{,}05$.

Groep A:	42,3	36,3	42,9	37,3	38,6	47,6	32,0	34,6	38,9	31,7	37,5	41,4
Groep B:	46,9	41,0	54,3	50,3	49,3	46,3	36,5	43,2	44,1	39,2	39,4	42,5
Groep C:	41,8	36,5	37,0	31,8	43,1	47,1	39,0	34,7	38,2	40,3	40,0	35,9

10.20. Uit WINSEMIUS c.s., Vitamine C en geestelijke arbeid, Verh. N.I.P.G. XXV, 1954.

Bij drie aselect samengestelde groepen proefpersonen zijn de volgende vitamine C gehalten (omgerekend per g creatinine) waargenomen:

A	14,2	3,6	11,1	11,3	9,5	12,3	7,8	2,7	31,5	5,0	15,7	
B	4,5	5,0	21,9	16,7	5,1	4,0	4,0	14,1	9,5	12,0	13,1	4,5
C	8,3	6,8	7,2	0,0	6,7	24,0	1,3	20,2	12,8	9,2	12,2	139,7

Onderzoek (met $\alpha = 0{,}10$) of deze groepen t.a.v. het vitamine C gehalte als aselect kunnen worden beschouwd.

10.7. De rangcorrelatietoets van Spearman

10.7.1. TOEPASSING

Deze toets kan worden gebruikt als men beschikt over de aselect verkregen waarnemingsparen (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) en de nulhypothese wil toetsen, dat de grootheden $\underset{\sim}{x}$ en $\underset{\sim}{y}$ (stochastisch) onafhankelijk zijn.

Voorbeeld

10.16. Van 8 konijnen, die na herhaalde intraveneuze toediening van een vitamine-D preparaat zijn gestorven, werden de calciumpercentages in de gedroogde stof van de aorta en de nieren bepaald. Deze zijn in tabel 10.14 vermeld. De te toetsen hypothese H_0 luidt dat de percentages in aorta en nieren onafhankelijk zijn. De tegenhypothese H_1 is dat er een positieve co(r)relatie tussen déze percentages bestaat, d.w.z. dat in het algemeen lage (hoge) percentages in de aorta samengaan met lage (hoge) percentages in de nieren [men spreekt van een negatieve correlatie, indien in het algemeen lage (hoge) waarden van de ene grootheid samengaan met hoge (lage) waarden van de andere]. Men kiest $\alpha = 0{,}05$.

Tabel 10.14. Calciumpercentages in de gedroogde stof van aorta (A) en nieren (N) van konijnen na herhaalde injecties met een vitamine-D preparaat

Konijn nr.	% Ca		Rangnummers		$d = R_x - R_y$	d^2
	A x	N y	A R_x	N R_y		
2.08	10,3	5,2	6	7	−1	1
2.09	12,7	7,1	7	8	−1	1
2.15	5,4	0,8	5	2	+3	9
2.21	4,6	1,6	4	4	0	0
2.25	4,1	3,3	3	5	−2	4
2.33	0,7	0,0	1	1	0	0
2.34	13,4	3,5	8	6	+2	4
2.37	3,6	1,2	2	3	−1	1
	$n = 8$				$\Sigma d = 0$	$R_0 = \Sigma d^2 = 20$

10.7.2. DE TOETSINGSGROOTHEID

Men kent aan alle waargenomen x-waarden van laag naar hoog de rangnummers 1, 2, . . ., n toe; men verkrijgt dan de rangnummers R_x. Vervolgens herhaalt men deze produre voor de y-waarden; dit levert de rangnummers R_y op. Per waarnemingspaar bepaalt men $d=R_x-R_y$. Als toetsingsgrootheid wordt gekozen:

$$R = \Sigma(R_x - R_y)^2 = \Sigma d^2. \tag{10.19}$$

Voor de waarnemingen in tabel 10.14 vindt men $R_0 = 20$.

10.7.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

Onder de nulhypothese kan de kansverdeling van $\underline{R}$ worden bepaald door alle $n!$ permutaties van de rangnummers 1 t/m n uit te schrijven en voor elke permutatie en de reeks 1, 2, . . . n de waarde van $\underline{R}$ te berekenen (tabel 10.18 bevat o.m. de op deze wijze bepaalde kansverdeling van $\underline{R}$ voor $n = 6$). Wanneer de rangorden van de x- en y-waarden precies overeenstemmen is elke d gelijk aan nul, zodat $R = 0$. Als de rangorden precies tegengesteld zijn, zodat de rangorde-paren $(1, n)$, $(2, n-1)$, . . . $(n-1, 2)$, $(n, 1)$ zijn, blijkt te gelden dat $R = \frac{1}{3}n(n^2-1)$. De verdeling van $\underline{R}$ is symmetrisch t.o.v.

$$\mu_R = \frac{n(n^2-1)}{6} \tag{10.20}$$

en heeft als variantie

$$\sigma_R{}^2 = \frac{n^2(n^2-1)(n+1)}{36} = \frac{n^2(n+1)^2(n-1)}{36}. \tag{10.21}$$

10.7.4. UITVOERING VAN DE TOETSING ZONDER GELIJKEN

Tabel N. In tabel N zijn voor $n = 5(1)50$ en de drempelwaarden 0,001, 0,01, 0,02, 0,05 en 0,10 de linker en rechter kritieke waarden van $\underline{R}$ bij *tweezijdige* toetsing opgenomen. Deze tabel kan gebruikt worden voor éénzijdige toetsing door de drempelwaarden te halveren.

Voor $n = 5(1)8$ zijn de kritieke waarden exact bepaald. Voor grotere waarden van n is de exacte berekening van de kansverdeling van $\underline{R}$ praktisch onuitvoerbaar (bij $n = 9$ zou men dan reeds $9! = 362880$ permutaties moeten uitschrijven). De kritieke waarden zijn hier dan ook bij benadering berekend.

Bij voorbeeld 10.16 vonden wij $R_0 = 20$. De nulhypothese wordt getoetst met als alternatieve hypothese, dat een positieve relatie aanwezig is. Indien H_1 juist is zullen de rangnummers per paar in de regel min of meer overeenstemmen, zodat de verschillen tussen de rangnummers klein zullen zijn en de grootheid $\underline{R}$ een lage waarde zal aannemen. Men zal H_0 dus uitsluitend willen verwerpen voor te lage

waarden van R, zodat links éénzijdig wordt getoetst. Uit tabel N blijkt dat voor $n = 8$ en $\alpha = 0{,}05$ de linker kritieke waarde bij éénzijdige toetsing gelijk is aan 30 (deze waarde staat dus onder $\alpha = 0{,}10$). Daar R_0 aanzienlijker kleiner is (een linkse kans bezit, die tussen 0,025 en 0,01 ligt) wordt H_0 verworpen ten gunste van H_1.

Het is duidelijk dat rechts éénzijdig wordt getoetst, wanneer de alternatieve hypothese luidt dat tussen de twee grootheden een negatieve relatie bestaat. Tweezijdige toetsing vindt plaats bij de alternatieve hypothese dat er een relatie aanwezig is, zonder dat daarvan a priori de 'richting' kan worden aangegeven.

Normale benadering. Voor grotere waarden van n kan men gebruik maken van het feit dat de grootheid

$$(10.22) \qquad \underline{T} = \frac{\underline{R} - \mu_R}{\sigma_R} = \frac{\underline{R} - n(n^2-1)/6}{\sqrt{\frac{n^2(n+1)^2(n-1)}{36}}} = \frac{6\underline{R} - n(n^2-1)}{n(n+1)\sqrt{n-1}}$$

bij benadering standaardnormaal verdeeld is. Het is gewenst een continuïteitscorrectie van 1 aan te brengen. Men berekent dus:

$$(10.23) \qquad T_0 = \frac{|\,6R_0 - n(n^2-1)\,| - 6}{n(n+1)\sqrt{n-1}}$$ en $P_D(T_0)$ bij tweezijdige toetsing,

of

$$(10.24.1) \qquad T_0 = \frac{6R_0 + 6 - n(n^2-1)}{n(n+1)\sqrt{n-1}}$$ en $P_L(T_0)$ bij links eenzijdige toetsing (H_1: positieve correlatie),

resp.

$$(10.24.2) \qquad T_0 = \frac{6R_0 - 6 - n(n^2-1)}{n(n+1)\sqrt{n-1}}$$ en $P_R(T_0)$ bij rechts éénzijdige toetsing (H_1: negatieve correlatie).

Past men — ondanks het kleine aantal waarnemingsparen — de normale benadering toe op voorbeeld 10.16, dan vindt men met (10.24.1)

$$T_0 = \frac{6 \times 20 + 6 - 8(8^2-1)}{8 \times 9\sqrt{8-1}} = \frac{-378}{190{,}49} = -1{,}98, \quad P_L \approx 0{,}024.$$

Deze uitkomst stemt reeds bevredigend overeen met de exacte kans die tabel N oplevert.

In 10.7.5 geven wij nog een tweede voorbeeld van het gebruik van de normale benadering.

10.7.5. UITVOERING VAN DE TOETSING MET GELIJKEN

Komen er in de waarden van x gelijken voor dan ontvangen deze per groep het gemiddelde van de hen toekomende rangnummers. Evenzo handelt men met gelijken onder de waarden van y. Het verschil d per paar en $R = \Sigma d^2$ worden daarna op de gewone wijze berekend.

Is het aantal gelijken betrekkelijk klein en verschillen de groepen gelijken niet te veel in omvang, dan kan men tabel N nog wel gebruiken en dan kan ook de normale benadering nog worden toegepast zonder dat een correctie voor gelijken wordt aangebracht. Bij grote aantallen gelijken kan men, als het aantal waarnemingsparen niet te klein is, beter de volgende normale benadering nemen:

$$\underline{T} = \frac{12R - (n^3 - D_x) - (n^3 - D_y)}{2\sqrt{(n^3 - D_x)\,(n^3 - D_y)}/(n-1)}. \tag{10.25}$$

In deze formule is $D_x = \Sigma t^3_{xi}$, waarin t_{xi} de omvang is van de i-de groep gelijken in de waarden van x; $D_y = \Sigma t^3_{yj}$, waarin t_{yj} de omvang is van de j-de groep gelijken in de waarden van y (in beide gevallen noteert men voor elke waarde die niet tot een groep gelijken behoort $t = 1$).

Tabel 10.15. Gewicht (x) en reactietijd op de hot plate (y) van 15 muizen

Muis	Waarnemingen x	Waarnemingen y	Rangnummers x	Rangnummers y	d +	d —	d^2
1	23	10,5	13	9,5	3,5		12,25
2	19	14,5	2,5	15		12,5	156,25
3	21	8,5	6,5	4	2,5		6,25
4	20	11	4,5	11,5		7	49
5	22	10	9,5	7,5	2		4
6	22	12,5	9,5	13		3,5	12,25
7	23	13,5	13	14		1	1
8	23	7	13	3	10		100
9	17	10,5	1	9,5		8,5	72,25
10	25	9	15	5	10		100
11	21	11	6,5	11,5		5	25
12	22	6	9,5	1	8,5		72,25
13	19	9,5	2,5	6		3,5	12,25
14	22	10	9,5	7,5	2		4
15	20	6,5	4,5	2	2,5		6,25
					+41,0	—41,0	633
					$\Sigma d = 0$		R_0

Gelijken

Omvang t	Omvang t^3	x f	x ft^3	y f	y ft^3
1	1	2	2	9	9
2	8	3	24	3	24
3	27	1	27		
4	64	1	64		
		$D_x =$	117	$D_y =$	33

Voorbeeld

10.17. Gegevens van Dr. CHR. L. RÜMKE. Tabel 10.15 bevat voor ieder van 15 muizen: x = het gewicht in gr, y = de reactietijd in seconden op de hot plate (de muis wordt op een zinken plaat gebracht, waarvan de temperatuur ongeveer 56°C is; de reactietijd is de tijd die verloopt tussen het ogenblik waarop het dier op de plaat wordt gebracht en het ogenblik waarop hij de voorpootjes van de plaat aftrekt). Men wil onderzoeken of er een correlatie is tussen gewicht en reactietijd en kiest $\alpha = 0{,}05$.

Voor de berekening van D_x en D_y is voor x en y vastgesteld, hoeveel groepen gelijken (f) van de omvang (t) voorkomen.
Formule (10.25) levert op

$$T = \frac{12(633) - (15^3 - 117) - (15^3 - 33)}{2\sqrt{(15^3 - 117)\,(15^3 - 33)/(15 - 1)}} = \frac{996}{1763{,}8} = 0{,}56, \quad P_D \approx 0{,}58$$

De hypothese H_0 dat het gewicht en de reactietijd op de hot plate onderling onafhankelijk zijn kan dus niet worden verworpen.

10.7.6. DE RANGCORRELATIECOËFFICIËNT VAN SPEARMAN

Als de grootheden $\underset{\sim}{x}$ en $\underset{\sim}{y}$ afhankelijk zijn spreekt men van *correlatie*. Als maat voor deze correlatie kan worden gebruikt de rangcorrelatiecoëfficiënt

$$r_R = 1 - \frac{6R}{n(n^2 - 1)}. \tag{10.26}$$

Het is duidelijk dat het volgende geldt: *a*) $r_R = 1$, als $R = 0$, d.i. als de beide rangorde-reeksen precies overeenstemmen (men spreekt dan van een volkomen positieve correlatie), *b*) $r_R = 0$, als $R = \mu_R$, d.i. als de x- en y-waarden geheel onafhankelijk zijn, en *c*) $r_R = -1$, als $R = n(n^2 - 1)/3$, d.i. als de twee rangorde-reeksen geheel tegengesteld verlopen (men spreekt dan van een volkomen negatieve correlatie).

Berekent men r_R uit een aselecte steekproef van n waarnemingsparen (x, y), dan is de waargenomen r_R een schatting van ϱ_R, de rangcorrelatiecoëfficiënt in de populatie.[1] De toets voor de hypothese, dat $\underset{\sim}{x}$ en $\underset{\sim}{y}$ stochastisch onafhankelijk zijn kan ook worden gezien als toets voor de hypothese $\varrho_R = 0$. Onder deze hypothese geldt voor de kansverdeling van $\underset{\sim}{r}_R$: $\mu_r = 0$ en $\sigma_r^2 = \frac{1}{n-1}$ en voor niet te kleine n is de grootheid

$$\underset{\sim}{T} = \frac{\underset{\sim}{r}_R - \mu_r}{\sigma_r} = \underset{\sim}{r}_R\sqrt{n-1} \tag{10.27}$$

bij benadering standaardnormaal verdeeld. (10.27) levert dezelfde uitkomst op als (10.23).

[1] ϱ = de Griekse letter rho.

10.7.7. OPGAVEN

10.21. Bij 15 vrouwtjesratten vindt men de volgende waarden voor het begingewicht (x) en de gewichtstoeneming van de 28e tot aan de 84e dag met een proteïnerijk dieet (y) (beide gewichten in gr):

Rat.	1	2	3	4	5	6	7	8	9	10	11	12
x	50	64	76	63	74	60	69	68	56	48	57	59
y	128	159	158	119	133	112	82	126	132	118	107	106

Rat.	13	14	15
x	46	45	65
y	96	103	104

Toets de hypothese dat begingewicht en gewichtstoeneming onderling onafhankelijk zijn ($\alpha = 0{,}05$).

10.22. Uit DE FRAITURE, Diss., Groningen, 1960. Onderstaande gegevens betreffen 13 patiënten met afwijkingen aan de lever. x = ingespoten dosis broom sulfaleïne (mg), y = uitscheiding van broomsulfaleïne met de urine gedurende het eerste uur na de injectie (mg):

Pat.	*KJ*	*GL*	*SN*	*MS*	*SS*	*HV*	*BB*	*GB*	*LB*	*LP*	*WM*
x	154	300	336	312	285	200	324	270	280	287	360
y	5,0	10,2	12,4	15,0	10,4	4,1	4,4	2,2	5,3	7,2	2,6

Pat.	*RM*	*SL*
x	380	400
y	4,5	7,0

Toets (met $\alpha = 0{,}10$) de hypothese dat de grootheden x en y onafhankelijk zijn.

10.23. Uit: N.T.v.G. 100 (1956), 557.
De navolgende gegevens zijn ontleend aan tabel II (patiënten met verschillende ziekten) en betreffen de 24 van de 32 gevallen, waarvan waarnemingen beschikbaar waren betreffende x = calcium in de urine (mg per 24 uur) en y = calcium in het serum (mg pct):

Pat.	1	2	3	4	5	7	10	11	13	15	17	19
x	860	775	450	360	330	280	180	170	140	130	110	95
y	13,8	17,4	13,4	9,9	14,4	10,0	9,7	9,7	8,7	9,1	9,7	10,0

Pat.	20	21	22	23	25	26	27	28	29	30	31	32
x	95	90	90	85	70	60	50	50	40	40	35	20
y	10,3	10,0	10,0	11,0	10,0	11,5	8,8	10,0	9,7	10,0	10,3	9,7

Toets de hypothese, dat de grootheden $\underline{x}$ en $\underline{y}$ stochastisch onafhankelijk zijn, met als alternatieve hypothese dat zij een positieve correlatie vertonen. Kies $\alpha = 0{,}05$.

10.8. De rangcorrelatietoets van Kendall

10.8.1. TOEPASSING

Deze toets kan eveneens worden gebruikt om de juistheid te onderzoeken van de hypothese H_0, dat de grootheden $\underline{x}$ en $\underline{y}$ stochastisch onafhankelijk zijn. In 10.9 vergelijken wij de toets van KENDALL met die van SPEARMAN.

10.8.2. DE TOETSINGSGROOTHEID

Ook de toets van KENDALL berust op de rangnummers van de x- en

y-waarden in de steekproef. Nu beschouwt men echter de grootheid P, d.i. het aantal der paren rangnummers in de y-reeks, waarvan de volgorde naar grootte overeenstemt met de volgorde naar grootte van de overeenkomstige rangnummerparen in de x-reeks.

Beschouw bijv. de volgende reeksen rangnummers, waarbij de x-reeks van links naar rechts in opklimmende grootte-orde staat:

x	1	2	3	4	5	6
y_1	1	2	3	4	5	6
y_2	3	1	2	5	4	6
y_3	6	5	4	3	2	1

In de x-reeks kunnen naar grootte ($x_i < x_j$) de rangnummerparen (1,2), (1,3), (1,4), . . ., (2,3), (2,4), . . ., (3,4), . . ., (5,6) worden gevormd, dat zijn dus in het algemeen $(n-1) + (n-2) + \ldots + 1 = n(n-1)/2$ paren. Indien de y-reeks volledig met de x-reeks overeenstemt, zoals bij y_1, is dus $P = n(n-1)/2$. Verloopt de y-reeks volledig tegengesteld aan de x-reeks, zoals bij y_3, dan stemt geen enkel rangnummerpaar in de y-reeks naar grootte overeen met het overeenkomstige paar in de x-reeks, zodat $P = 0$.

Nu bekijken wij de reeks y_2. Het paar (1,2) staat in de juiste volgorde, het paar (1,3) echter niet, de paren (1,4), (1,5) en (1,6) wel. Het paar (2,3) niet, de paren (2,4), (2,5) en (2,6) wel, enz. Eenvoudig is in te zien dat men voor twee rangnummerreeksen de waarde van P kan bepalen door voor ieder y-rangnummer vast te stellen, hoeveel *hogere* y-rangnummers *rechts* er van staan, dus bijv.:

Rangnummerreeksen	x	1	2	3	4	5	6
	y	3	1	2	5	4	6
Aantal *grotere* y's rechts van y_i		$3+4+3+1+1+0 = 12 = P_0$					
Aantal *kleinere* y's rechts van y_i		$2+0+0+1+0+0 = 3$					
		Totaal $= \frac{1}{2}n(n-1) = 15$					

De grootheid P varieert van 0 t/m $\frac{1}{2}n(n-1)$. KENDALL voert nu als toetsingsgrootheid in

$$S = 2P - \frac{n(n-1)}{2}. \tag{10.28}$$

Deze grootheid varieert van $-\frac{1}{2}n(n-1)$ tot en met $\frac{1}{2}n(n-1)$. Uit het voorgaande volgt, dat S_0 gelijk is aan het aantal paren rangnummers in de y-reeks, waarvan de volgorde naar grootte overeenstemt met de volgorde naar grootte van de overeenkomstige rangnummers in de x-reeks, *verminderd* met het aantal paren waarvoor dit niet het geval

is. Voor de y-reeks in het voorgaande voorbeeld is dus volgens (10.28): $S_0 = 2(12) - 6(5)/2 = 9$, of rechtstreeks uit de tabel: $S_0 = 12-3=9$.

In tabel 10.16 is de berekening van S_0 voor de gegevens in voorbeeld 10.16 uitgevoerd. De waarden van $\underline{x}$ zijn zo geplaatst dat hun rangnummers in de volgorde 1, 2, . . ., n staan.

Tabel 10.16. Calciumpercentages in de gedroogde stof van aorta (A) en nieren (N) bij konijnen (voorbeeld 10.16)

Konijn nr.	% Ca		Rangnummers		S-bijdragen	
	A x	N y	R_x	R_y	+	—
2.33	0,7	0,0	1	1	7	0
2.37	3,6	1,2	2	3	5	1
2.25	4,1	3,3	3	5	3	2
2.21	4,6	1,6	4	4	3	1
2.15	5,4	0,8	5	2	3	0
2.08	10,3	5,2	6	7	1	1
2.09	12,7	7,1	7	8	0	1
2.34	13,4	3,5	8	6	0	0
	$S_0 = 22 - 6 = 16$				22	6

Bij elke R_y wordt dus de *positieve* bijdrage tot S_0 (+ kolom) verkregen door te tellen hoeveel *hogere* rangnummers er onder staan. De *negatieve* bijdrage (—kolom) wordt per R_y verkregen door te tellen hoeveel *lagere* rangnummers er onder staan. S_0 volgt overigens ook uit de positieve bijdragen: $S_0 = 2(22) - 8(7)/2 = 16$.

10.8.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

Onder de hypothese dat de grootheden $\underline{x}$ en $\underline{y}$ stochastisch onafhankelijk zijn, is bij een gegeven x-reeks elke permutatie van de rangnummers van de y-reeks even waarschijnlijk, zodat de kansverdeling van $\underline{S}$ kan worden berekend. Hierbij kan gebruik worden gemaakt van een zg. recursieformule, d.w.z. de verdeling voor $n + 1$ waarnemingsparen kan uit die voor n worden afgeleid.

Voor de verdeling van $\underline{S}$ onder H_0 geldt, als geen gelijken in de waargenomen reeksen voorkomen:

a. Als $\frac{1}{2}n(n-1)$ even (oneven) is, neemt $\underline{S}$ slechts even (oneven) waarden aan.

b. De verdeling is symmetrisch t.o.v.

$$\mu_s = 0 \tag{10.29}$$

en heeft als variantie

$$\sigma_s^2 = \frac{n(n-1)(2n+5)}{18}. \tag{10.30}$$

10.8.4. UITVOERING VAN DE TOETSING ZONDER GELIJKEN

Exacte toetsing, tabel O. Voor $n = 4$ t/m 40 zijn de exacte kansverdelingen van $\underline{S}$ onder H_0 berekend en in de tabellen van KAARSEMAKER & VAN WIJNGAARDEN (1952) opgenomen. In tabel O zijn voor deze waarden van n en de drempelwaarden 0,01, 0,02, 0,05 en 0,10 de rechter kritieke waarden S_R van $\underline{S}$ bij *tweezijdige toetsing* gegeven. Wegens de symmetrie van de verdeling t.o.v. 0 geldt: $S_L = -S_R$. De tabel kan voor *éénzijdige toetsing* worden gebruikt door de drempelwaarden te *halveren.*

In tabel 10.16 vinden wij voor voorbeeld 10.16 met $n = 8$: $S_0 = 16$. H_0 wordt getoetst met als alternatief H_1: tussen de grootheden $\underline{x}$ en $\underline{y}$ bestaat een positieve correlatie. Onder H_1 neemt de grootheid $\underline{S}$ overwegend positieve waarden aan, zodat rechts éénzijdig wordt getoetst. De rechter kritieke waarde bij éénzijdige toetsing met $\alpha = 0{,}05$ en $n = 8$ dient te worden opgezocht in de kolom met $\alpha = 0{,}10$ en men vindt dan $S_R = 16$. In dit geval is $S_0 = S_R$, zodat H_0 wordt verworpen t.g.v. H_1.[1]

Normale benadering. Voor $n > 40$ kan men gebruik maken van het feit, dat de grootheid

$$\underline{T} = \frac{\underline{S} - \mu_S}{\sigma_S} = \frac{\underline{S}}{\sqrt{n(n-1)(2n+5)/18}} \tag{10.31}$$

bij benadering standaardnormaal verdeeld is. Deze benadering is zeer goed. Een continuïteitscorrectie kan worden aangebracht door de absolute waarde van S_0 met 1 te verminderen.

10.8.5. UITVOERING VAN DE TOETSING MET GELIJKEN

Aan de gelijken binnen de x- en binnen de y-reeks wordt weer per groep het gemiddelde rangnummer toegekend. Voor de bepaling van de bijdragen tot S_0 wordt dan de volgende regel aangehouden: als van de waarnemingsparen (x_i, y_i) en (x_j, y_j) met $i < j$, de rangnummers van x_i en x_j, of van y_i en y_j, of van x_i en x_j èn y_i en y_j gelijk zijn, is de bijdrage tot S_0 gelijk aan nul. Dit is gebaseerd op de gedachte dat men beide mogelijke rangschikkingen van de twee gelijken beschouwt en het gemiddelde van de daarbij verkregen S-bijdragen als bijdrage neemt: dit gemiddelde is steeds gelijk aan 0. Als voorbeeld is in tabel 10.17 de berekening van S_0 uitgevoerd voor twee rangnummerreeksen met veel gelijken.

[1] De uitkomst van de toets van KENDALL stemt t.a.v. de uitspraak (H_0 verwerpen) overeen met die van de toets van SPEARMAN (zie 10.7.4). Via eerstgenoemde toets komt men echter tot een overschrijdingskans tussen 0,05 en 0,025 (exact: 0,031) en met laatstgenoemde tot een overschrijdingskans tussen 0,025 en 0,01. Wij komen hier in 10.9 op terug.

Tabel 10.17. Berekening van S_0 bij het optreden van gelijken

Rangnummers		S-bijdragen		
R_x	R_y	+	—	0
1	1	9	0	0
$2\frac{1}{2}$	2	7	0	1
$2\frac{1}{2}$	$4\frac{1}{2}$	4	0	3
$4\frac{1}{2}$	$4\frac{1}{2}$	4	0	2
$4\frac{1}{2}$	$4\frac{1}{2}$	4	0	1
$6\frac{1}{2}$	$4\frac{1}{2}$	3	0	1
$6\frac{1}{2}$	8	1	0	2
8	8	1	0	1
$9\frac{1}{2}$	8	0	0	1
$9\frac{1}{2}$	10	0	0	0
		$33 = S_0$		

Als het aantal gelijken betrekkelijk klein is en de groepen gelijken niet te veel in omvang verschillen, kan men tabel O nog wel bij benadering gebruiken, terwijl ook de normale benadering nog kan worden toegepast zonder dat men een fout van betekenis maakt. Bij grote aantallen gelijken dient een correctie te worden aangebracht. Men dient dan in formule (10.23) σ_S te vervangen door σ_S', waarvoor geldt

$$(10.32) \qquad \sigma_S'^2 = \frac{2[n^3 - D_x - 3(n^2 - C_x)]\,[n^3 - D_y - 3(n^2 - C_y)] + 9(n-2)\,(n^2 - C_x)\,(n^2 - C_y).}{18n(n-1)\,(n-2)}$$

D_x en D_y hebben dezelfde betekenis als in formule (10.25). Verder is $C_x = \Sigma t^2_{xi}$ en $C_y = \Sigma t^2_{yi}$.

10.8.6. DE RANGCORRELATIECOËFFICIËNT VAN KENDALL

Deze coëfficiënt is

$$(10.33) \qquad r_K = \frac{2S}{n(n-1)}$$

en varieert dus evenals de rangcorrelatiecoëfficiënt van SPEARMAN tussen +1 (volkomen positieve correlatie) en —1 (volkomen negatieve correlatie). Bij berekening van r_K uit een aselecte steekproef van n waarnemingsparen is deze een schatting van τ_K, de rangcorrelatiecoëfficiënt in de populatie.[1]

10.8.7. OPGAVEN

10.24. Pas bij opgave 10.21 de toets van KENDALL toe.

10.25. Pas bij opgave 10.22 de toets van KENDALL toe.

[1] τ = de Griekse letter tau.

10.9. Keuze van een rangcorrelatietoets

In tabel 10.18 is voor $n = 6$ de simultane verdeling van de grootheden $\underline{R}(r_R)$ van de toets van SPEARMAN en $\underline{S}(r_K)$ van de toets van KENDALL opgenomen. Deze grootheden worden verschillend gedefinieerd, zodat hun kansverdelingen verschillen. De grootheid $\underline{R}$ neemt in totaal $n^2 = 36$ en de grootheid $\underline{S}$ neemt in totaal $1 + \frac{1}{2}n(n-1) = 16$ verschillende waarden aan. De verdeling van $\underline{S}$ is regelmatiger van vorm dan de verdeling van $\underline{R}$ en kan beter door een normale verdeling worden benaderd. Het verband tussen de grootheden blijkt niet rechtlijnig te zijn, maar vertoont een enigszins gebogen (S-vormig) verloop. Dit blijkt uit het feit dat de rangcorrelatiecoëfficiënt van SPEARMAN in het gebied tussen + 1 en 0 en tussen 0 en — 1 systematisch hogere absolute waarden aanneemt dan de rangcorrelatiecoëfficiënt van KENDALL (ga dit na aan de hand van de tabel).

De rangcorrelatiemethoden zijn tegenwoordig vooral van betekenis, omdat zij gebruikt kunnen worden om *verdelingsvrij* de hypothese te toetsen, dat twee grootheden stochastisch onafhankelijk zijn.[1] Welke methode kan men nu in de praktijk het best gebruiken? VAN DER WAERDEN (1957) toont aan, dat de toets van SPEARMAN een groter onderscheidingsvermogen bezit en concludeert hieruit dat deze toets te prefereren is. Men kan hier tegenover stellen, dat bij deze toets voor $n = 9$ t/m 40 slechts benaderende kritieke waarden ter beschikking staan, terwijl men dan bij de toets van KENDALL over exacte kritieke waarden kan beschikken. Dit bezwaar geldt echter alleen als men bij het gebruik van de toets van SPEARMAN een waarde van $\underline{R}$ vindt, die dicht bij de kritieke waarde ligt. Voor $n > 40$ dient men voor beide toetsen een normale benadering te gebruiken en deze zal dan ook voor de toets van SPEARMAN wel voldoend nauwkeurig zijn.

Een definitieve uitspraak omtrent de keuze tussen de twee toetsen is dus moeilijk te geven. Onze voorkeur gaat echter uit naar de toets van SPEARMAN, mede omdat deze bij grotere n en vrij veel gelijken eenvoudiger te hanteren is.

1 Voorheen was men voornamelijk geïnteresseerd in de rangcorrelatiecoëfficiënten als benadering van de lineaire correlatiecoëfficiënt (die in hoofdstuk 20 besproken wordt).

Tabel 10.18. Simultane verdeling van $r_r(R)$ en $r_k(S)$ voor $n = 6$

R	r_r	S / r_k	−15 / −1	−13 / −0,867	−11 / −0,733	−9 / −0,600	−7 / −0,467	−5 / −0,333	−3 / −0,200	−1 / −0,067	+1 / +0,[illegible]	+5 / +0,333	+7 / +0,467	+9 / +0,600	+11 / +0,733	+13 / +0,867	+15 / +1	Totaal
0	1																1	1
2	0,943															5		5
4	0,886														6			6
6	0,829													1	8			9
8	0,771													16				16
10	0,714												6	6				12
12	0,657												8	6				14
14	0,600											4	20					24
16	0,543										1	8	11					20
18	0,486											21						21
20	0,428										9	10	4					23
22	0,370									2	4	22						28
24	0,314										20	4						24
26	0,257									2	32							34
28	0,200									12	8							20
30	0,144									22	8	2						32
32	0,086								4	30	8							42
34	0,029								16	13								29
36	−0,029								13	16								29
38	−0,086							8	30	4								42
40	−0,144						2	8	22									32
42	−0,200							8	12									20
44	−0,257							32	2									34
46	−0,314						4	20										24
48	−0,370						22	4	2									28
50	−0,428					4	10	9										23
52	−0,486						21											21
54	−0,543					11	8	1										20
56	−0,600					20	4											24
58	−0,657				6	8												14
60	−0,714				6	6												12
62	−0,771				16													16
64	−0,829			8	1													9
66	−0,886			6														6
68	−0,943		5															5
70	−1	1																1

HOOFDSTUK 11

VERDELINGSVRIJE TOETSEN TEGEN VERLOOP IN EEN WAARNEMINGSREEKS

11.1 Inleiding

Alle statistische toetsings- en schattingstechnieken berusten op de onderstelling, dat de elementen van een steekproef, waaraan de waarnemingen verricht zijn, aselect getrokken zijn uit de bestudeerde populatie. Er zijn gevallen, waarin het van belang is te onderzoeken of aan deze onderstelling voldaan is, bijv. als men beschikt over een reeks waarnemingen, waarvan de aselectiviteit op z'n minst twijfelachtig is, of als men een steekproef heeft getrokken op een wijze waarvan men verwacht dat zij aselect is. Het kan ook voorkomen, dat men beschikt over een *tijdreeks*, d.i. een in chronologische volgorde verzamelde reeks waarnemingen en wil onderzoeken of daarin één of ander verloop in de tijd optreedt. In beide gevallen kan men gebruik maken van een aantal toetsen voor de hypothese H_0, dat de beschouwde waarnemingen aselect, d.i. dus onderling onafhankelijk zijn. Onder deze hypothese vertonen de waarnemingen in de regel een onregelmatig verloop. Onder de alternatieve hypothese H_1, dat de waarnemingen niet aselect gespreid zijn, kunnen echter verschillende typen van verloop (Eng. 'trend') optreden.

Tabel 11.1 geeft hier enkele voorbeelden van. Deze tabel bevat vijf reeksen waarnemingen in chronologische volgorde. Wij zullen deze (fictieve) gegevens beschouwen als de systolische bloeddrukken (in mm Hg) van 5 patiënten, A t/m E, die op 20 achtereenvolgende dagen gemeten zijn (elke dag op hetzelfde tijdstip, door dezelfde waarnemer met hetzelfde apparaat). Bestudeer deze reeksen voordat u verder leest en tracht uit de tabel aan te geven of zij naar uw mening aselect zijn, c.q. of zij een bepaald verloop vertonen. Bekijk eerst daarna hun grafische voorstellingen in figuur 11.1. Het blijkt dan, dat de grafieken een veel beter inzicht verschaffen dan de getallenreeksen.

Aan de hand van figuur 11.1 valt nu omtrent de vijf tijdreeksen in tabel 11.1 het volgende op te merken:

A. Deze reeks vertoont een systematisch *stijgend verloop*, d.w.z. de stijging wordt soms door een (geringe) daling onderbroken, maar dit is begrijpelijk indien men bedenkt, dat er in de waarnemingen ook een zekere toevallige variatie zal optreden, die soms de stijging zal tegenwerken of in een daling zal doen omslaan (bij werkelijke gegevens zal

Tabel 11.1. Systolische bloeddruk van 5 patiënten op 20 achtereenvolgende dagen (fictieve gegevens)

Dag	Patiënt				
	A	*B*	*C*	*D*	*E*
1	113	98	142	122	112
2	115	101	137	105	126
3	114	110	136	121	118
4	117	105	132	114	124
5	118	99	126	118	116
6	120	106	119	115	109
7	124	104	116	125	111
8	125	109	112	112	125
9	126	100	113	123	110
10	131	102	110	109	121
11	128	119	111	120	122
12	134	123	115	111	123
13	141	118	117	127	129
14	143	116	123	117	115
15	147	122	130	124	114
16	148	130	124	108	127
17	145	115	131	116	113
18	149	124	135	113	120
19	153	127	139	128	119
20	157	114	140	110	108

dit effect in de regel sterker zijn, dan bij deze fictieve waarnemingen).

B. Deze waarnemingen geven, na de 10de dag, een *plotselinge stijging* te zien. De waarnemingen van de eerste èn die van de laatste tien dagen afzonderlijk maken echter een 'aselecte' indruk.

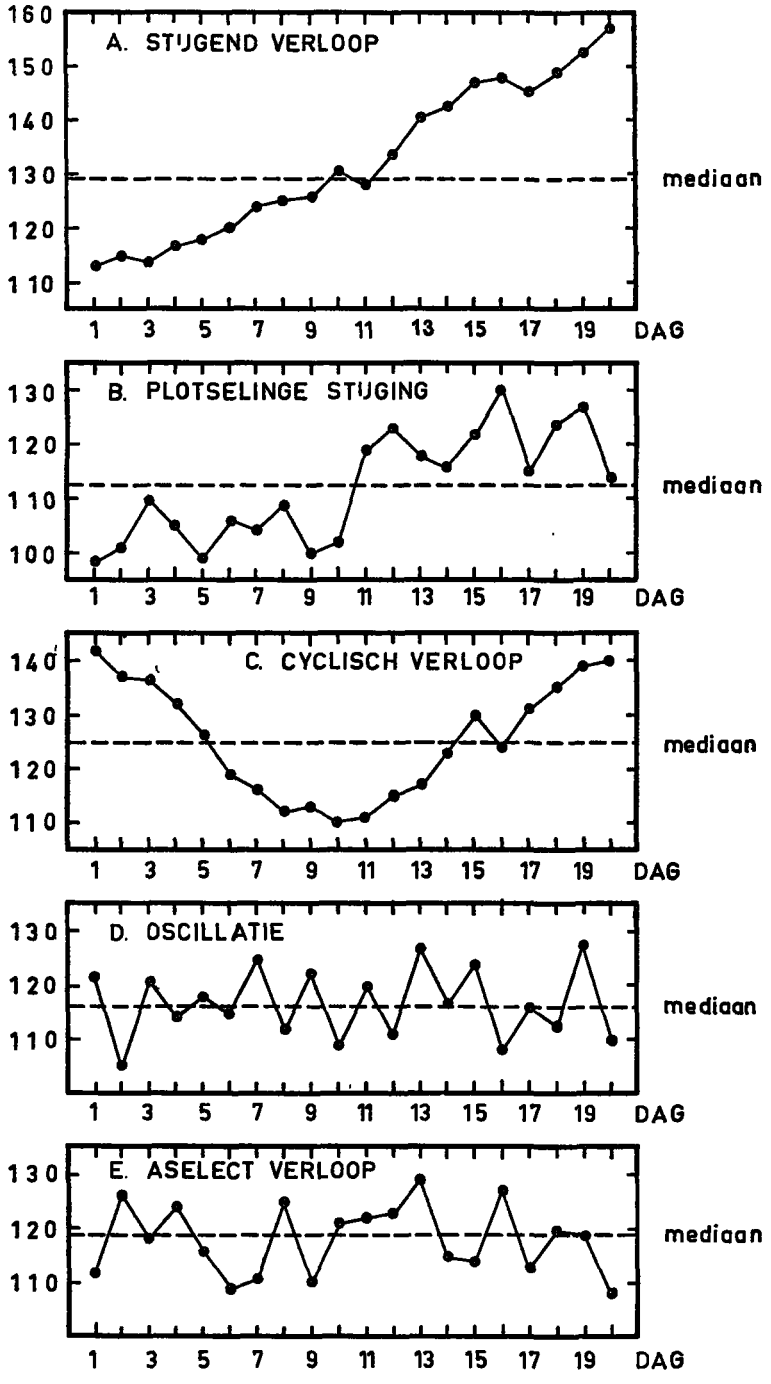
C. Hier ziet men (een deel van) een *cyclisch verloop*, d.w.z. een geleidelijke daling, gevolgd door een geleidelijke stijging, die weer in een daling zal overgaan, enz. Op deze wijze zou zich bijv. een seizoenseffect kunnen manifesteren.

D. Deze waarnemingen vertonen geen niveauverschuiving, maar een *oscillatie*, d.i. een snelle periodieke fluctuatie. De door de waarnemingspunten getekende polygoon bezit een typisch zaagtandpatroon.

E. Slechts deze waarnemingsreeks maakt in zijn geheel een 'aselecte' indruk. Er zijn op het oog geen bijzonderheden in het verloop te onderkennen.

In deze voorbeelden zijn de voornaamste verlooppatronen aangegeven. Uiteraard kunnen zij ook in combinatie voorkomen. Zo kan een stijgend of dalend verloop samengaan met oscillatie of met een plotselinge niveauverandering (bijv. een verlaging tengevolge van het toedienen van een geneesmiddel).

Er zijn verschillende toetsen tegen verloop in een waarnemingsreeks mogelijk. Bij alle luidt de te toetsen hypothese H_0, dat de waar-



Figuur 11.1. Grafieken van de tijdreeksen in tabel 11.1.

nemingen in aselecte volgorde staan. De alternatieve hypothese H_1, waarvoor deze toetsen gevoelig zijn, is echter niet steeds identiek, terwijl de toetsen tegen hetzelfde alternatief een verschillend onderscheidingsvermogen blijken te bezitten. Wij komen hierop in 11.8 terug, nadat wij eerst de toetsen afzonderlijk hebben besproken.

11.2 De rangcorrelatietoetsen als toetsen tegen verloop

De resp. in 10.7 en 10.8 behandelde rangcorrelatietoetsen van SPEARMAN en KENDALL kunnen ook worden gebruikt als men de hypothese H_0 wil toetsen, dat in een waarnemingsreeks geen verloop optreedt, met als alternatieve hypothese H_1, dat er een stijgend of dalend verloop aanwezig is.

De toetsing van deze nulhypothese komt dan neer op de toetsing van de hypothese, dat tussen de waarnemingen en de tijdstippen waarop deze betrekking hebben geen correlatie aanwezig is, met als alternatief dat er een positieve correlatie (stijgend verloop) of een negatieve

Tabel 11.2. Toepassing van de rangcorrelatietoets van SPEARMAN op de tijdreeksen in tabel 11.1

Dag	A				B				C				D				E			
	v	R	d	d^2	w	R	d	d^2	x	R	d	d^2	y	R	d	d^2	z	R	d	d^2
1	113	1	0	0	98	1	0	0	142	20	−19	361	122	15	−14	196	112	5	− 4	16
2	115	3	−1	1	101	4	−2	4	137	17	−15	225	105	1	+ 1	1	126	18	−16	196
3	114	2	+1	1	110	10	−7	49	136	16	−13	169	121	14	−11	121	118	10	− 7	49
4	117	4	0	0	105	7	−3	9	132	14	−10	100	114	8	− 4	16	124	16	−12	144
5	118	5	0	0	99	2	+3	9	126	11	− 6	36	118	12	− 7	49	116	9	− 4	16
6	120	6	0	0	106	8	−2	4	119	8	− 2	4	115	9	− 3	9	109	2	+ 4	16
7	124	7	0	0	104	6	+1	1	116	6	+ 1	1	125	18	−11	121	111	4	+ 3	9
8	125	8	0	0	109	9	−1	1	112	3	+ 5	25	112	6	+ 2	4	125	17	− 9	81
9	126	9	0	0	100	3	+6	36	113	4	+ 5	25	123	16	− 7	49	110	3	+ 6	36
10	131	11	−1	1	102	5	+5	25	110	1	+ 9	81	109	3	+ 7	49	121	13	− 3	9
11	128	10	+1	1	119	15	−4	16	111	2	+ 9	81	120	13	− 2	4	122	14	− 3	9
12	134	12	0	0	123	17	−5	25	115	5	+ 7	49	111	5	+ 7	49	123	15	− 3	9
13	141	13	0	0	118	14	−1	1	117	7	+ 6	36	127	19	− 6	36	129	20	− 7	49
14	143	14	0	0	116	13	+1	1	123	9	+ 5	25	117	11	+ 3	9	115	8	+ 6	36
15	147	16	−1	1	122	16	−1	1	130	12	+ 3	9	124	17	− 2	4	114	7	+ 8	64
16	148	17	−1	1	130	20	−4	16	124	10	+ 6	36	108	2	+14	196	127	19	− 3	9
17	145	15	+2	4	115	12	+5	25	131	13	+ 4	16	116	10	+ 7	49	113	6	+11	121
18	149	18	0	0	124	18	0	0	135	15	+ 3	9	113	7	+11	121	120	12	+ 6	36
19	153	19	0	0	127	19	0	0	139	18	+ 1	1	128	20	− 1	1	119	11	+ 8	64
20	157	20	0	0	114	11	+9	81	140	19	+ 1	1	110	4	+16	256	108	1	+19	361
R_0			10				304				1290				1340				1330	
r_R		+0,992				+0,771				+0,030				−0,008				0		
T_0*		4,32				3,36				0,13				0,03				—		
P_D		$\ll$0,001				<0,001				>0,85				>0 95				—		

* volgens (10.22)

correlatie (dalend verloop) bestaat. De toets is in 't bijzonder gevoelig voor een min of meer regelmatig stijgend (dalend) verloop en in mindere mate voor een sprongsgewijze stijging (daling). Merk op, dat bij deze toepassing van de rangcorrelatietoetsen de tijdstippen van waarneming een niet-stochastische grootheid vormen.

Bij bestudering van de uitkomsten in tabel 11.2 bedenke men, dat voor $n = 20$: $\mu_R = 20(20^2 - 1)/6 = 1330$ en $\sigma_R = \frac{1}{6} \times 20 \times 21 \sqrt{19} = 305{,}12$. Voor elke reeks is de waargenomen waarde van r_R opgenomen. Om een indruk te geven van de grootte-orde van de tweezijdige overschrijdingskans van R_0 is deze via de normale benadering berekend.

Duidelijk blijkt, dat de rangcorrelatietoets het stijgende verloop bij A en B signaleert, maar ongevoelig is voor het cyclische verloop bij C en de oscillatie van D. Bij reeks E met het toevalspatroon is zelfs $R_0 = \mu_R$. Merk voorts op, dat bij B het eerste en het laatste tiental waarnemingen geen stijgend of dalend verloop vertonen. R_0 kan voor elk dezer reeksen met $n = 10$ eveneens uit tabel 11.2 worden afgelezen en blijkt 138, resp. 166 te zijn.

11.3. De serietoets op de waarnemingen

11.3.1. TOEPASSING

Deze toets kan worden gebruikt als men de hypothese H_0 wil toetsen, dat een reeks waarnemingen een aselecte volgorde bezit (aselect getrokken is) en de alternatieve hypothese H_1 luidt, dat dit niet het geval is. De toets is gevoelig voor elk soort verloop (zie echter 11.6).

11.3.2. DE TOETSINGSGROOTHEID

Onderstel dat men 20 maal met drie verschillende geldstukken werpt en daarbij de volgende uitkomsten verkrijgt:

Geldst. I: K K K M M M M K K K K K K M M M M M M M

Geldst. II: K M K M K M K M K M K M K M K M K M K M

Geldst. III: K K M K M M K M M M K K M M K M K K K K

Bij reeks I zien wij betrekkelijk lange series van opeenvolgende gelijke uitkomsten, zodat het totaal aantal waargenomen series (4) vrij klein is. Bij reeks II valt de regelmaat op, waarmee de uitkomsten K en M elkaar afwisselen: elke serie bestaat hier uit slechts één keer K of M, zodat er in totaal 20 series zijn. Bij reeks III ziet men een onregelmatige afwisseling van korte en langere series en deze steekproef met 11 series maakt intuïtief een aselecte indruk.

Bij de serietoets kiest men als toetsingsgrootheid het aantal *series* u (men spreekt ook van: kettingen, iteraties, Eng. 'runs'), d.i. het aantal groepen van opeenvolgende gelijke kenmerken. Bij kwantitatieve waarnemingen kan men deze grootheid eveneens gebruiken en wel door de waarnemingen te klassificeren als gelijk aan of groter dan de mediaan (H) en kleiner dan de mediaan (L). Het aantal waarnemingen met ken-

merk H noemen wij dan m_1, het aantal met kenmerk L m_2, zodat $m_1 + + m_2 = n$. Voor de gegevens in tabel 11.1 verkrijgt men op deze wijze de reeksen met H's en L's, die in tabel 11.3 zijn vermeld (men kan deze klassificatie eenvoudig uitvoeren met de grafieken in figuur 11.1; als men reeds rangnummers heeft toegekend, kan men ook daarvan uitgaan).

Tabel 11.3. Toepassing van de serietoets op de tijdreeksen in tabel 11.1

Reeks	Dag																				u_0
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
A	L	L	L	L	L	L	L	L	L	H	L	H	H	H	H	H	H	H	H	H	4
B	L	L	L	L	L	L	L	L	L	L	H	H	H	H	H	H	H	H	H	H	2
C	H	H	H	H	H	L	L	L	L	L	L	L	L	L	H	L	H	H	H	H	5
D	H	L	H	L	H	L	H	L	H	L	H	L	H	H	H	L	L	L	H	L	16
E	L	H	L	H	L	L	L	H	L	H	H	H	H	L	L	H	L	H	H	L	13

11.3.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

De toetsingsgrootheid $\underline{u}$, het aantal series, kan de waarden 2, 3, . . ., $2m_1$ ($m_1 \leqslant m_2$) aannemen. Onder H_0 zijn alle mogelijke permutaties van de n waarnemingen in de steekproef even waarschijnlijk, zodat men de kansverdeling van $\underline{u}$ kan berekenen. Als $m_1 = m_2 = m = \frac{1}{2}n$ blijkt deze grootheid onder H_0 de volgende kansverdeling te volgen:

$$(11.1) \qquad P(\underline{u} = u) \begin{cases} = \dfrac{2\left(C^{m-1}_{\frac{u}{2}-1}\right)^2}{C^n_m} & \text{voor } \textit{even} \text{ waarden van } \underline{u}, \\ = \dfrac{2\left(C^{m-1}_{\frac{u+1}{2}-1}\right)\left(C^{m-1}_{\frac{u+1}{2}-2}\right)}{C^n_m} & \text{voor } \textit{oneven} \text{ waarden van } \underline{u}. \end{cases}$$

De afleiding van deze formule is te vinden in MOOD (1950) en BENNETT & FRANKLIN (1954). Voor het geval $m_1 \neq m_2$ geldt een iets gecompliceerder formule, die wij hier buiten beschouwing laten. Als voorbeeld is in tabel 11.4 de verdeling van $\underline{u}$ onder H_0 voor $m_1 = m_2 = 5$ berekend met (11.1).

Tabel 11.4 Verdeling van $\underline{u}$ onder H_0 voor $m_1 = m_2 = 5$

u	2	3	4	5	6	7	8	9	10
$P(u) \cdot C_{10}^5$	$2(C_5^4)^2$	$2(C_5^4)(C_5^4)$	$2(C_5^4)^2$	$2(C_5^4)(C_5^4)$	$2(C_5^4)^2$	$2(C_5^4)(C_5^4)$	$2(C_5^4)^2$	$2(C_5^4)(C_5^4)$	$2(C_5^4)^2$
$P(u) \cdot 252$	2	8	32	48	72	48	32	8	2

Voor de verdeling van $\underline{u}$ onder H_0 blijkt verder te gelden:

$$\mu_u = \frac{2m_1m_2}{n} + 1 \tag{11.2}$$

en

$$\sigma_u^2 = \frac{2m_1m_2\,(2m_1m_2 - n)}{n^2(n-1)}. \tag{11.3}$$

Voor $m_1 = m_2 = m = \frac{1}{2}n$ is deze verdeling symmetrisch t.o.v. μ en gaan de voorgaande formules resp. over in

$$\mu_u = \tfrac{1}{2}n + 1 = m + 1 \tag{11.4}$$

en

$$\sigma_u^2 = \frac{n(n-2)}{4(n-1)}. \tag{11.5}$$

Als de alternatieve hypothese H_1 luidt, dat de waarnemingen niet aselect zijn, wordt tweezijdig getoetst. Wil men alleen onderzoeken of de waarnemingen een (plotseling of geleidelijk) stijgend of dalend, of wel een cyclisch verloop vertonen, dan vindt links éénzijdige toetsing plaats, omdat men dan alleen tot het verwerpen van H_0 zal willen overgaan als te weinig (lange) series optreden. Wil men nagaan, of er snelle periodieke fluctuaties (zonder ander verloop) zijn, dan toetst men rechts éénzijdig: men zal H_0 immers alleen willen verwerpen, als te veel (korte) series voorkomen. De toets is echter betrekkelijk ongevoelig voor oscillaties, die tezamen met een stijgende, dalende of cyclische trend voorkomen; men kan dan echter de serietoets op de opeenvolgende verschillen in 11.4 gebruiken.

11.3.4. UITVOERING VAN DE TOETSING

Exacte toetsing, tabel P. In tabel P–1 zijn voor $n \leqslant 39$ en $m_1 < m_2 \leqslant 20$ ($m_1 \neq m_2$) de linker kritieke waarden u_L en de rechter kritieke waarden u_R van $\underline{u}$, die gelden voor *tweezijdige* toetsing met $\alpha = 0{,}01$, 0,02, 0,05 en 0,10 opgenomen (bij gebruik van deze tabel zijn m_1 en m_2 verwisselbaar). In tabel P–2 vindt men deze kritieke waarden voor $m_1 = m_2 = \frac{1}{2}n$ en $n = 10(2)90$ voor *tweezijdige* toetsing met $\alpha =$ 0,01, 0,02, 0,05 en 0,10.

Toetst men tweezijdig met een drempelwaarde α, dan kan men tot het verwerpen van H_0 overgaan indien

$$u_0 \leqslant u_L(\alpha,n) \text{ of } u_0 \geqslant u_R(\alpha,n).$$

Bij *éénzijdige* toetsing dient men de drempelwaarden in de tabellen

P–1 en P–2 te *halveren*. Toetst men links éénzijdig, dan wordt H_0 dus bij een drempelwaarde α verworpen als $u_0 \leqslant u_L(2\alpha,n)$ is. Toetst men rechts éénzijdig, dan besluit men tot het verwerpen van H_0 als $u_0 \geqslant u_R(2\alpha,n)$ is.

Voor de reeksen A t/m E in tabel 11.1 is $n = 20$ en $m_1 = m_2$. Toetst men voor elk dezer reeksen bij een 5% drempel de hypothese H_0, dat een aselecte volgorde aanwezig is, dan zijn volgens tabel P–2 de kritieke waarden $u_L = 6$ en $u_R = 16$. Op grond van de uitkomsten in tabel 11.3 komt men dan tot de volgende conclusies:

A, B en C: H_0 verwerpen, het aantal waargenomen series is significant te laag,

D: H_0 verwerpen, het aantal waargenomen series is significant te hoog,

E: H_0 niet verwerpen.

Voorbeeld

11.1. Uit: Statistical notes for malaria workers, WHO/HS/58, 1956.

Tabel 11.5. Annual malaria morbidity per 1000 population, 1936—1951, Ceylon

Jaar	Rangnr.	Morbiditeit	Rangnr.	d	d^2
1936	1	522	12	—11	121
1937	2	403	10	— 8	64
1938	3	352	8	— 5	25
1939	4	543	15	—11	121
1940	5	572	16	—11	121
1941	6	533	13	— 7	49
1942	7	534	14	— 7	49
1943	8	348	7	+ 1	1
1944	9	265	6	+ 3	9
1945	10	390	9	+ 1	1
1946	11	412	11	0	0
1947	12	196	5	+ 7	49
1948	13	109	4	+ 9	81
1949	14	99	3	+11	121
1950	15	81	2	+13	169
1951	16	58	1	+15	225
		$u_0 = 6$		$\Sigma d = 0$	$R_0 = 1206$

Onderstel, dat men de hypothese H_0 toets dat in deze tijdreeks geen verloop optreedt, met als alternatief H_1 dat een dalend verloop aanwezig is. Men kiest $\alpha = 0{,}05$. Past men de serietoets toe, dan is $u_0 = 6$ met $n = 16$ en $m_1 = m_2 = 8$. Er dient links éénzijdig te worden getoetst. Uit tabel P–2 blijkt, dat H_0 juist niet kan worden verworpen, daar $u_L\,(2\alpha = 0{,}10) = 5$.

Past men echter de rangcorrelatietoets van SPEARMAN toe, dan is $R_0 = 1206$ ($r_R = -0{,}77$) met $n = 16$. Er dient rechts éénzijdig getoetst te worden (zie 10.7) Tabel N geeft bij $\alpha = 0{,}01$ als rechter kritieke waarde $R_R = 1089$ (deze tabel geldt voor tweezijdige toetsing, zodat deze waarde onder 0,02 wordt afgelezen). R_0 is groter dan R_R, zodat H_0 bij een 1% drempel kan worden verworpen. Dit voorbeeld toont aan, dat de rangcorrelatietoets bij dit dalende verloop een veel groter onderscheidingsvermogen bezit dan de serietoets.

Normale benadering. Voor niet te kleine en niet te veel uiteenlopende m_1 en m_2 kan men gebruik maken van het feit, dat de grootheid

$$\underline{T} = \frac{\underline{u} - \mu_u}{\sigma_u} \tag{11.6}$$

bij benadering standaardnormaal verdeeld is. Ook in dit geval is het gewenst een continuïteitscorrectie ter grootte van $\frac{1}{2}$ aan te brengen. Bij *tweezijdige toetsing* neemt men dus:

$$T_0 = \frac{|\, u_0 - \mu_u \,| - \frac{1}{2}}{\sigma_u}, \text{ en } P_D(T_0). \tag{11.7}$$

Bij *links éénzijdige toetsing*:

$$T_0 = \frac{u_0 - \mu_u + \frac{1}{2}}{\sigma_u}, \text{ en } P_L(T_0), \tag{11.8.1}$$

en bij *rechts éénzijdige toetsing*:

$$T_0 = \frac{u_0 - \mu_u - \frac{1}{2}}{\sigma_u}, \text{ en } P_R(T_0). \tag{11.8.2}$$

11.4. Toets tegen verloop van een kans

Deze toets kan worden gebruikt als men beschikt over een reeks waarnemingen betreffende een dichotomie (met de kenmerken A en B), die bijv. op verschillende tijdstippen zijn verricht en de hypothese H_0 wil toetsen dat de kans op kenmerk A in de tijd constant is, met als alternatieve hypothese dat deze kans toe- of afneemt.

Zo betreft bijv. de volgende reeks van veertien waarnemingen, die achtereenvolgens bij één patient zijn verricht, de uitkomst van een bepaalde reactie (1 = positief, 0 = negatief):

Tijdstippen = rangnummers	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Reactie	1	1	1	0	1	1	1	1	0	0	0	1	0	0

Men wil bij een 5% drempel de hypothese H_0 toetsen dat de kans op een positieve reactie in de tijd constant is, met als alternatieve hypothese dat deze kans stijgt of daalt.

Men kan de toetsing uitvoeren door middel van de toets van WILCOXON, door de enen als een steekproef van $n_1 = 8$ en de nullen als een tweede steekproef van $n_2 = 6$ elementen te beschouwen. Kent men aan de opeenvolgende waarnemingen de rangnummers $1, 2, \ldots, n$ toe, dan is de toetsingsgrootheid S_1, de som van de bij de eerste steekproef (de enen) behorende rangnummers. Bij tweezijdige toetsing met

onbetrouwbaarheid α kan men tot het verwerpen van H_0 overgaan, indien $S_1 \leqslant S_L(\alpha, n)$ of $S_1 \geqslant S_R(\alpha, n)$ is. S_L en S_R kunnen voor $\alpha = 0{,}10$, 0,05, 0,02 en 0,01 in tabel L worden afgelezen. Toetst men links éénzijdig (afnemende kans), dan verwerpt men H_0 voor $S_1 \leqslant S_L(2\alpha, n)$; bij rechts éénzijdige toetsing (toenemende kans) verwerpt men H_0 voor $S_1 \geqslant S_R(2\alpha, n)$. Voor grote waarden van n kan men de normale benadering van de toets van WILCOXON volgens formule (10.12) toepassen.

Bij het voorbeeld is

$$S_1 = 1 + 2 + 3 + 5 + 6 + 7 + 8 + 12 = 44;$$

bij tweezijdige toetsing met $\alpha = 0{,}05$ vindt men in tabel L-2 voor $n_1 = 8$ en $n_2 = 6$ de kritieke waarden 44 en 76. Daar de waargenomen som gelijk is aan de linker kritieke waarde wordt H_0 verworpen en tot een dalend verloop van de kans op een positieve reactie geconcludeerd.

Opmerking. Deze toets is aan onze aandacht ontsnapt en zodoende pas later toegevoegd. Hij kan daardoor niet meer in de beschouwing in 11.8 en tabel 11.7 worden opgenomen. Het is duidelijk dat hij in deze tabel achter (A) thuishoort, als kwalitatieve tegenhanger van de rangcorrelatietoetsen.

11.5. De serietoets op de verschillen tussen opeenvolgende waarnemingen

11.5.1. TOEPASSING

Deze toets kan worden gebruikt om de hypothese H_0 te toetsen dat een reeks waarnemingen een aselecte volgorde vertoont. In de regel zal men de toets toepassen met als alternatieve hypothese $H_1(a)$: de waarnemingsreeks vertoont een geleidelijk stijgend (dalend) of een cyclisch verloop, dan wel met $H_1(b)$: de waarnemingsreeks vertoont oscillatie.

In tegenstelling tot de in 11.3 behandelde serietoets op de waarnemingen zelf is deze toets: (*a*) *niet* gevoelig voor een sprongsgewijze stijging (daling), en (*b*) ook gevoelig voor oscillatie, als deze tezamen met een regelmatig stijgend (dalend) of cyclisch verloop voorkomt.

11.5.2. DE TOETSINGSGROOTHEID

Men kent aan de $n - 1$ verschillen tussen de waarnemingen 2 en 1, 3 en 2, . . ., $i + 1$ en i, . . ., n en $n - 1$ een plusteken toe als de $(i + 1)$-de waarneming groter is dan de i-de en een minteken als het tegendeel geldt. De toetsingsgrootheid v is het aantal series van opeenvolgende plussen of minnen. In tabel 11.6 is v_0 bepaald voor de tijdreeksen in tabel 11.1.

Tabel 11.6. Toepassing van de serietoets op verschillen tussen opeenvolgende waarnemingen op de gegevens in tabel 11.1

Reeks	Dag																				v_0
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
A	+	−	+	+	+	+	+	+	+	−	+	+	+	+	+	−	+	+	+		7
B	+	+	−	−	+	−	+	−	+	+	+	−	−	+	+	−	+	+	−		12
C	−	−	−	−	−	−	−	+	−	+	+	+	+	+	−	+	+	+	+		6
D	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−		19
E	+	−	+	−	−	+	+	−	+	+	+	+	−	−	+	−	+	−	−		12

11.5.3. DE VERDELING VAN DE TOETSINGSGROOTHEID ONDER H_0

Onder H_0 blijkt de grootheid $\underline{v}$ een verdeling te volgen met

$$\mu_v = \frac{2n-1}{3} \tag{11.9}$$

en

$$\sigma_v^2 = \frac{16n-29}{90}. \tag{11.10}$$

11.5.4. UITVOERING VAN DE TOETSING

Exacte toetsing, tabel Q. In tabel Q zijn voor $n = 5$ t/m 40 de linker- en rechter kritieke waarden van $\underline{v}$ bij *éénzijdige toetsing* met $\alpha = 0{,}01$ en 0,05 gegeven. Toetst men H_0 met als alternatief $H_1(a)$ dan wordt links éénzijdig getoetst en besluit men tot het verwerpen van H_0, indien $v_0 \leqslant v_L(\alpha,n)$ is. Is het alternatief $H_1(b)$, dan vindt rechts éénzijdige toetsing plaats, zodat H_0 verworpen wordt indien $v_0 \geqslant v_R(\alpha,n)$ is. Wij laten de interpretatie van de uitkomsten in tabel 11.6 aan de lezer over (merk echter op, dat de toets niet gevoelig is voor de sprongsgewijze stijging bij reeks *B*).

Normale benadering. Voor $n > 40$ is de grootheid

$$\underline{T} = \frac{\underline{v} - \mu_v}{\sigma_v} \tag{11.11}$$

bij benadering standaardnormaal verdeeld. Bij het gebruik van deze benadering kan weer een continuïteitscorrectie ter grootte van $\frac{1}{2}$ worden aangebracht.

Het optreden van gelijken. Als twee opeenvolgende waarnemingen gelijk zijn, is hun verschil nul. Het aantal series kan dan niet meer

ondubbelzinnig worden vastgesteld. Is echter het aantal nullen klein, dan kan men ze eenmaal zodanig door plus- of mintekens vervangen, dat het aantal series maximaal wordt en een tweede maal zo, dat dit aantal minimaal wordt. Verkrijgt men beide keren een significante uitkomst, dan wordt H_0 verworpen. Zijn beide uitkomsten insignificant dan wordt H_0 niet verworpen. Is de een het wel en de ander niet, dan kan men niet tot een uitspraak komen. Men weet dan echter wel, dat men op zijn ongunstigst met een uitkomst te doen heeft die bijna significant is.

11.6. De toets op de tekens van de opeenvolgende waarnemingen

Deze toets (die wij kortweg de *plus-min* toets zullen noemen) kan worden gebruikt als men de hypothese H_0 wil toetsen dat de reeks waarnemingen in aselecte volgorde staat, met als alternatieve hypothese H_1 dat hij een *regelmatig* stijgend of dalend verloop vertoont. Deze toets is voor andere vormen van verloop niet gevoelig.

Naar onze mening is de plus-min toets (waarbij alleen een normale benadering ter beschikking staat) alleen van belang, als men voor een betrekkelijk lange reeks waarnemingen snel wil onderzoeken, of van een geleidelijk stijgende of dalende trend sprake is.

De toetsingsgrootheid is het aantal plussen s in de reeks van verschillen tussen de opeenvolgende waarnemingen. Voor de tijdreeksen in tabel 11.1 lezen wij uit tabel 11.6 af:

Reeks	A	B	C	D	E
s_0	16	11	10	9	10

Onder H_0 volgt de grootheid $\underline{s}$ een symmetrische kansverdeling met

$$\mu_s = \frac{n-1}{2} \tag{11.12}$$

en

$$\sigma_s^2 = \frac{n+1}{12} \tag{11.13}$$

(hierin is n het aantal waarnemingen). Voor niet te kleine n volgt onder H_0 de grootheid

$$\underline{T} = \frac{\underline{s} - \mu_s}{\sigma_s} \tag{11.14}$$

bij benadering een standaardnormale verdeling. Deze benadering kan worden verbeterd door het aanbrengen van een continuïteitscorrectie ter grootte van $\frac{1}{2}$.

Bij een stijgend verloop zal men overwegend plussen, bij een dalend verloop overwegend minnen aantreffen. In de regel zal bij het gebruik van deze toets éénzijdige toetsing plaatsvinden. Luidt de alternatieve hypothese, dat een geleidelijk dalend verloop optreedt, dan vindt links éénzijdige toetsing plaats. Houdt het alternatief een geleidelijk stijgend verloop in, dan dient rechts éénzijdig te worden getoetst.

Voor alle reeksen uit tabel 11.1 is $n = 20$, zodat $\mu_s = (20-1)/2 = 9{,}5$ en $\sigma_s^2 = (20+1)/12 = 1{,}75$, $\sigma_s = 1{,}3229$. Het is zonder meer duidelijk, dat H_0 niet kan worden verworpen voor de reeksen B, C, D en E. Voor reeks A levert de normale benadering op

$$T_0 = \frac{16 - 9{,}5 - 0{,}5}{1{,}3229} = 4{,}5.$$

Ondanks de betrekkelijk kleine n kan men dus tot het verwerpen van H_0 overgaan.

11.7. Toetsen gebaseerd op de serielengte

Zoals reeds in 11.3 is opgemerkt komen afwijkingen van het toevalspatroon ook tot uitdrukking in de lengten van de series. Er zijn dan ook toetsen mogelijk waarbij men van de serielengten uitgaat. Eén der mogelijkheden is, dat men de waargenomen frequentieverdeling van de serielengten opstelt en deze door middel van de χ^2-toets voor aanpassing met de verdeling onder H_0 vergelijkt. Deze methode kan alleen worden toegepast als het aantal waarnemingen vrij groot is (daar er anders te lage verwachte frequenties optreden) terwijl zij vrij omslachtig is (bij het berekenen van de verwachte frequenties moet de onderlinge afhankelijkheid van de aantallen runs in rekening worden gebracht).

Een eenvoudiger mogelijkheid is, dat men alleen de lengte van de *langste* serie van de waarnemingsreeks beschouwt. VAN STRIK (1959) bespreekt o.m. de toets waarbij deze lengte als toetsingsgrootheid fungeert en geeft daarbij ook tabellen met kritieke waarden.

Tenslotte vermelden wij dat in GEIGY (1960, blz. 130) nog tabellen met kritieke waarden voor een viertal andere op serielengten gebaseerde toetsen zijn opgenomen.

11.8. Keuze van een toets tegen verloop

De toetsen tegen verloop in een waarnemingsreeks, die in de paragrafen 11.2, 11.3, 11.5 en 11.6 zijn besproken, hebben alle als nulhypothese dat de waarnemingen een aselecte volgorde vertonen. Het is echter reeds gebleken dat de alternatieve hypothesen min of meer verschillend zijn, d.w.z. de toetsen zijn niet gevoelig voor hetzelfde soort verloop, hoewel zij ook niet geheel onafhankelijk zijn daar zij elkaar gedeeltelijk ‚overlappen.' Dit komt nog eens tot uitdrukking in tabel 11.7, waarin voor elke toets is aangegeven welk alternatief [welke soort(en) verloop] hij kan signaleren en op welke wijze dit geschiedt.

Bij de bespreking van deze toetsen hebben wij ze alle op de vijf waarnemingsreeksen in tabel 11.1 toegepast. Wij konden daardoor met deze voorbeelden de verschillen tussen de toetsen demonstreren. In werkelijkheid kan men de volgende situaties onderscheiden:

Tabel 11.7. Overzicht van de verdelingsvrije toetsen tegen verloop in een waarnemingsreeks. H_0: de waarnemingen staan in aselecte volgorde

H_1: de waarnemingen vertonen verloop	Toets			
	Rangcorrelatietoetsen	Serietoets op waarnemingen	Serietoets op verschillen	Plus-min toets
	11.2	11.3	11.5	11.6
A. Stijgend	Pos. corr.	Te weinig series	Te weinig series	Te veel plus
A. Dalend	Neg. corr.			Te weinig plus
B. Sprongsgewijs omhoog	Pos. corr.	Te weinig series		
B. Sprongsgewijs omlaag	Neg. corr.			
C. Cyclisch		Te weinig series	Te weinig series	
D. Oscillatie		Te veel series	Te veel series	

1. Men toetst H_0 met als alternatieve hypothese H_1, dat de waarnemingen niet in aselecte volgorde staan, zonder dat men daarbij een specifiek alternatief kan aangeven. In dit geval verdient (tweezijdige) toepassing van de serietoets op de waarnemingen de voorkeur.

2. Men toetst H_0 met een specifiek alternatief H_1, voortvloeiende uit theoretische overwegingen (*niet* uit bestudering van de waarnemingsreeks zelf). Bijv.: Men verwacht, dat bij toediening van een bloeddrukverlagend middel een geleidelijke daling van de bloeddruk optreedt. Of: Men verwacht, dat op het geven van een anti-bioticum een plotselinge daling van de lichaamstemperatuur volgt. Bij de navolgende H_1 kieze men dan de daarbij vermelde toets(en):

a. Een geleidelijk stijgend (resp. dalend) verloop: Een der rangcorrelatietoetsen (éénzijdig). Bij grote n kan men gemakshalve de plus-min toets gebruiken (geen rangnummering, geen rekenwerk). Men realisere zich echter, dat deze toets een kleiner onderscheidingsvermogen bezit dan de rangcorrelatietoetsen.

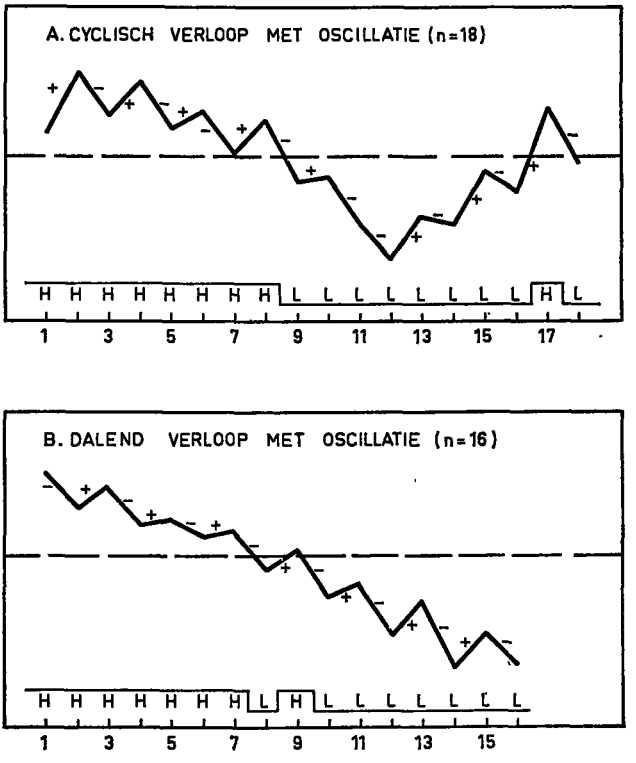
b. Een sprongsgewijze stijging (resp. daling): De serietoets op de waarnemingen (links éénzijdig) of een rangcorrelatietoets (éénzijdig).

c. Een cyclisch verloop: De serietoets op de waarnemingen (links éénzijdig).

d. Oscillatie: De serietoets op de verschillen (rechts éénzijdig).

Bij 2.*c* verdient de serietoets op de waarnemingen de voorkeur boven de serietoets op de verschillen, omdat bij het optreden van een cyclische trend mèt een zekere oscillatie de eerstgenoemde toets vrij ongevoelig

is voor deze oscillatie (zie figuur 11.2A). Bij 2.*d* is precies het omgekeerde van kracht, daar de serietoets op de verschillen de oscillatie signaleert, ondanks het optreden van een stijgende, dalende of cyclische trend (zie figuur 11.2B).



Figuur 11.2. Oscillatie en cyclisch, resp. dalend verloop:
A: Serietoets op de waarnemingen: $u_0 = 4$,
Serietoets op de verschillen: $v_0 = 16$,
B: Serietoets op de waarnemingen: $u_0 = 4$,
Serietoets op de verschillen: $v_0 = 15$.

Indien men op grond van vaktechnische of andere overwegingen reeds vóór het beschikbaar komen van de waarnemingen de aard van het eventuele verloop kan voorzien, kan men aan de hand van tabel 11.7 tevoren de 'meest doeltreffende' toets kiezen.

In de praktijk krijgt men vaak te maken met een typisch detectieprobleem, dat wil in dit geval zeggen: met een waarnemingsreeks, waarvan men wil onderzoeken of er enig soort verloop (en zo ja, welk verloop) optreedt. Als men dan de toets kiest op grond van het verloop, dat men waarneemt (of meent waar te nemen) en deze uitvoert bij

een onbetrouwbaarheidsdrempel α, is de werkelijke onbetrouwbaarheid, door de selectie die reeds in deze keuze verborgen zit, niet meer gelijk aan α. Men kan in deze situatie het best steeds dezelfde *reeks* van toetsen toepassen (bijv. bestaande uit één der rangcorrelatietoetsen in 11.2 en de twee serietoetsen, resp. in 11.3 en 11.5), ieder bij dezelfde onbetrouwbaarheidsdrempel α. Tussen deze drie toetsen bestaat een zekere — onbekende — mate van afhankelijkheid, zodat de simultane onbetrouwbaarheid niet precies kan worden bepaald. Men kan echter wel stellen dat deze ten minste gelijk zal zijn aan α, maar niet groter zal zijn dan 3α.

11.9. Opgaven

11.1. Betreft opgave 2.1. Onderstel, dat de waarnemingen regel voor regel verkregen zijn. Toets de hypothese, dat de steekproef aselect is en kies $\alpha = 0{,}05$.

11.2. Als 11.1 maar voor opgave 2.2.

11.3. Betreft opgave 10.6. Toets voor de beide waarnemingsreeksen in deze opgave de hypothese, dat zij een aselecte volgorde vertonen met als alternatief, dat een stijgend verloop optreedt. Kies $\alpha = 0{,}01$.

11.4. Een steekproef heeft, in volgorde van trekking, elementen met de volgende waarden opgeleverd:

65,8 66,6 78,0 53,9 62,3 62,2 69,0 61,8 51,6 74,8 54,7 51,2 39,2 39,8 48,8 57,6 62,3 39,8 41,0 42,4

Toets de hypothese, dat de waargenomen volgorde aselect is, met als alternatieve hypothese, dat te weinig series voorkomen. Kies $\alpha = 0{,}05$.

11.5. Gegeven zijn de twee volgende aselecte steekproeven:

Steekproef 1: 55 64 67 68 68 70 72 74
Steekproef 2: 56 58 60 61 62 63 64 66

Toets de hypothese, dat deze uit dezelfde populatie stammen (kies daarbij $\alpha = 0{,}05$):

a. Met de serietoets, *b*. Met de toets van WILCOXON.
Welk verschil bestaat er tussen deze toetsen?

11.6. Onderstaande waarnemingen betreffen het hemoglobinegehalte (g%) van een arbeider, gemeten gedurende een jaar met tussenpozen van twee weken. Onderzoek, of deze waarnemingsreeks een cyclisch verloop vertoont ($\alpha = 0{,}01$):

Datum	4/4	18/4	2/5	16/5	30/5	13/6	27/6	11/7	25/7	8/8	22/8	5/9	19/9
Hgb.	12,2	12,6	12,7	12,9	12,9	13,3	13,2	13,4	13,6	13,5	13,3	13,4	13,9
Datum	3/10	17/10	31/10	14/11	28/11	12/12	24/12	10/1	24/1	7/2	21/2	7/3	21/3
Hgb.	14,3	14,6	14,3	13,8	13,9	14,0	13,3	13,5	13,6	13,2	13,1	13,0	12,7

HOOFDSTUK 12

VERDELINGSVRIJE TOETSEN TEGEN VERLOOP IN MEERDERE WAARNEMINGSREEKSEN

12.1. Inleiding

In tabel 12.1 zijn de verdelingsvrije toetsen tegen stijgend of dalend verloop vermeld, die in dit hoofdstuk worden behandeld. Zij betreffen de volgende proefopzetten:

1. *k geordende verwante waarnemingsreeksen.* Dit is proefopzet (*b*) uit 8.2.1, waarbij de k verwante waarnemingsreeksen van 1 t/m k een bepaalde volgorde bezitten, bijv. omdat zij op k opeenvolgende tijdstippen betrekking hebben (zie de voorbeelden in de tabellen 12.2 en 12.3).

2. *k geordende aselecte steekproeven of groepen.* Dit is proefopzet (*c*) uit 8.2.1, waarbij de k aselecte steekproeven of groepen in de volgorde 1 t/m k zijn gerangschikt, bijv. naar leeftijd, naar welstand, naar de toegediende hoeveelheid van een stof, e.d. (zie de voorbeelden in de tabellen 12.4 en 12.7).

Wij bespreken achtereenvolgens de toepassing van deze toetsen op kwantitatieve waarnemingen (waarden of rangnummers) en op waarnemingen betreffende een dichotomie (met de kenmerken A en B). De verlooptoetsen voor een geordend kategorisch systeem komen in hoofdstuk 21 aan de orde.

Tabel 12.1. Toetsen tegen verloop in k waarnemingsreeksen

Proefopzet	Aard van de waarnemingen	
	Kwantitatief: rangnummers waarden	Kwalitatief: dichotomie
k geordende verwante waarnemingsreeksen	§ 12.2	§ 12.3
k geordende aselecte waarnemingsreeksen	§ 12.4	§ 12.5

Bij de door ons behandelde verlooptoetsen dient men aan elke waarnemingsreeks een waarde g_j ($j = 1, 2, \ldots, k$) toe te kennen. Hierbij kan men twee situaties onderscheiden:

a. De ordening vindt plaats op grond van een *grootheid* x (zoals leeftijd, dosis of expositieduur). Zijn de waarden van x equidistant dan neme men

(12.1.1) $g_j = 2j - (k + 1)$ (indien k even is),

(12.1.2) $g_j = j - (k + 1)/2$ (indien k oneven is).

Betreffen vijf aselecte steekproeven bijv. kinderen van 5, 6, 7, 8 en 9 jaar, dan neemt men volgens (12.1.2) voor de eerste steekproef (5 jaar): $g_1 = 1 - (5 + 1)/2 = -2$, voor de tweede (6 jaar): $g_2 = -1$, voor de derde (7 jaar) $g_3 = 0$, voor de vierde (8 jaar) $g_4 = +1$ en voor de vijfde (9 jaar) $g_5 = +2$.

Zijn de waarden van x niet equidistant en wil men hiermede rekening houden, dan neme men

(12.2) $g_j = x_j - \bar{x}$ $(\bar{x} = \Sigma x_j/k)$.

Heeft men bijv. vier steekproeven van kinderen die resp. 4, 6, 7 en 11 jaar zijn, dan is $\bar{x} = (4 + 6 + 7 + 11)/4 = 7$, zodat de waarden van g resp. gelijk zijn aan -3, -1, 0 en $+4$. Merk op dat zowel bij gebruik van (12.1) als bij gebruik van (12.2) steeds $\Sigma g_j = 0$ wordt genomen.

b. De ordening berust op een *kwalitatief criterium* (zoals bijv. de indeling van kinderen naar de 'welstand' van hun ouders, de indeling van personen naar de 'hoogte' van de genoten opleiding, of de indeling van patiënten naar het klinisch stadium ten tijde van het onderzoek. Gewoonlijk kan men dan aan de waarnemingsreeksen de waarden van g het best volgens (12.1) toekennen. Meent men echter dat de 'afstanden' tussen de categorieën verschillend zijn en wil men hiermee rekening houden, dan kan men de waarden van g met (12.2) bepalen.

Heeft men bijv. te maken met steekproeven die onderscheiden zijn met de kwalificaties: G(oed), V(oldoende), M(atig), O(nvoldoende), S(lecht) en bevat de met G gekwalificeerde steekproef uitsluitend 'uitblinkers' en die met S alleen evident ongunstige gevallen, dan zou men als waarden van x kunnen kiezen G = 8, V = 6, M = 5, O = 4 en S = 2. Met (12.2) volgen hieruit als waarden van g resp. $+3$, $+1$, 0, -1 en -3.

12.2. Toets tegen verloop in k verwante waarnemingsreeksen

12.2.1. TOEPASSING

De toets kan worden toegepast als men beschikt over k geordende verwante waarnemingsreeksen, waarvan de k waarnemingen per element uit rangschikkingen of waarden van een grootheid bestaan. De nulhypothese luidt dat de k waarnemingen per element beschouwd kunnen worden als aselecte trekkingen uit dezelfde populatie, met als alternatieve hypothese dat voor het merendeel der elementen de waarnemingen in de volgorde 1, 2, ..., k een gemeenschappelijk (stijgend of dalend) verloop vertonen.

Voorbeeld

12.1. Uit WINSEMIUS c.s., Vitamine C en geestelijke arbeid, Verh. N.I.P.G. 25 (1954). De gegevens links in tabel 12.2 betreffen de vitamine C uitscheiding in de urine (per g creatinine) bij 10 proefpersonen op vier opeenvolgende tijdstippen, met steeds 14 dagen tussenruimte. De nulhypothese is dat per persoon de uitscheiding op de vier tijdstippen van waarneming slechts toevallige verschillen te zien geeft. De alternatieve hypothese luidt dat voor het merendeel der proefpersonen de gehalten in de volgorde 1, 2, 3, 4 een overwegend stijgend of dalend verloop vertonen (zodat tweezijdig wordt getoetst). Men kiest $\alpha = 0{,}01$.

Tabel 12.2. Vier geordende verwante waarnemingsreeksen

No. proefpersoon	Tijdstip van onderzoek				Waarden van g				R_i
	1	2	3	4	—3	—1	+1	+3	
1	14,2	14,0	11,4	14,3	3	2	1	4	8
2	3,6	9,3	23,8	15,0	1	2	4	3	2
3	11,1	9,7	27,3	7,3	3	2	4	1	14
4	11,3	13,5	4,3	14,6	2	3	1	4	6
5	9,5	11,6	19,6	57,0	1	2	3	4	0
6	12,3	6,3	3,0	16,5	3	2	1	4	8
7	7,8	12,9	19,2	12,1	1	3	4	2	6
8	2,7	6,6	10,6	13,9	1	2	3	4	0
9	31,5	67,7	44,0	139,2	1	3	2	4	2
10	5,0	13,0	6,9	13,6	1	3	2	4	2
$k = 4$	$n = 10$				17	24	25	34	$48 = R_0'$

12.2.2. DE TOETSINGSGROOTHEID

De $k = 4$ onderzoektijdstippen zijn equidistant, zodat de vier waarnemingsreeksen volgens (12.1.1) als waarden van g resp. —3, —1, +1 en +3 toegemeten krijgen. Vervolgens vervangt men evenals bij de toets van FRIEDMAN de vier waarden per element van laag naar hoog door rangnummers en bepaalt per waarnemingsreeks de rangsom S_j (zoals gebruikelijk krijgen hierbij gelijke waarden het gemiddelde van de rangnummers die zij moeten bezetten). Als relatief weinig gelijken optreden kan men als toetsingsgrootheid nemen

$$\underline{T} = \frac{\Sigma \underline{S}_j g_j}{\sqrt{\frac{nk(k+1)}{12} \Sigma g_j^2}}. \tag{12.3.1}$$

Zijn er betrekkelijk veel gelijken dan kan men beter nemen

$$\underline{T} = \frac{\Sigma \underline{S}_j g_j}{\sqrt{\frac{nk^3 - D}{12(k-1)} \Sigma g_j^2}}. \tag{12.3.2}$$

waarin D wordt berekend op de wijze die in 10.4.5 is aangegeven. Onder H_0 volgt deze grootheid bij benadering een standaardnormale verdeling.

De toetsing kan ook worden uitgevoerd door per element (per rij) de toetsingsgrootheid R_i van de toets van SPEARMAN te berekenen en bij de verlooptoets $R' = \Sigma R_i$ als toetsingsgrootheid te nemen. De R_i kunnen worden berekend door de rangnummers voor het i-de element met de reeks 1, 2, 3, . . ., k te vergelijken. Men vindt dan bijv. bij persoon 1 van tabel 12.2:

Rangnummers	3	2	1	4
Volgnummers	1	2	3	4
Verschil $= d$	2	0	−2	0
d^2	4	0	4	0

$R_1 = \Sigma d^2 = 8$

Uit de rechts in tabel 12.2 opgenomen kolom blijkt dat $R_0' = 48$.

Onder H_0 bezitten de waarnemingen per element een aselecte volgorde en volgens 10.7.3 is de grootheid $\underline{R}_i$ dan bij benadering normaal verdeeld met gemiddelde $\mu_i = k(k^2 - 1)/6$ en variantie $\sigma_i^2 = k^2 (k + 1)^2 (k - 1)/36$. Hieruit volgt dat onder H_0 de grootheid $\underline{R}'$ eveneens bij benadering normaal verdeeld is met gemiddelde

$$\mu_{R'} = \Sigma\mu_i = n\mu_i = kn\,(k^2 - 1)/6$$

en variantie

$$\sigma_{R'}^2 = \Sigma\sigma_i^2 = n\sigma_i^2 = nk^2\,(k + 1)^2\,(k - 1)/36,$$

zodat de grootheid

$$\underline{T} = \frac{\underline{R}' - \mu_{R'}}{\sigma_{R'}} = \frac{6\underline{R}' - kn\,(k^2 - 1)}{k\,(k + 1)\,\sqrt{n\,(k - 1)}} \tag{12.4}$$

bij benadering standaardnormaal verdeeld is. Deze toetsingsgrootheid is identiek aan (12.3.1), maar heeft een tegengesteld teken; dit is nl. negatief bij een stijgend en positief bij een dalend verloop. Vanzelfspreken kan men ook bij (12.4) een correctie voor gelijken aanbrengen, maar wij laten deze buiten beschouwing omdat men de toetsing dan eenvoudiger via (12.3.2) kan uitvoeren.

12.2.3. UITVOERING VAN DE TOETSING

Bij tweezijdige toetsing (H_1: stijgend of dalend verloop) kan men H_0 bij een onbetrouwbaarheidsdrempel α verwerpen als $P_D(T_0) \leqslant \alpha$ is. Voert men de toetsing uit met als alternatieve hypothese dat een dalend verloop optreedt, dan vindt linkszijdige toetsing plaats en kan men tot verwerpen overgaan als $P_L(T_0) \leqslant \alpha$ is. Luidt de alternatieve hypothese dat een stijgend verloop optreedt, dan wordt rechtszijdig getoetst en H_0 verworpen als $P_R(T_0) \leqslant \alpha$ is.

Voor de gegevens in tabel 12.2 is de berekening van T_0 met formule (12.3.1) in tabel 12.3 uitgevoerd. Er wordt tweezijdig getoetst met $\alpha = 0{,}01$ en daar $P_D(T_0) = P_D(2{,}85) = 0{,}0044$ wordt H_0 verworpen. De vitamine C uitscheiding vertoont dus een systematisch stijgend verloop.

Tabel 12.3. Berekeningen bij het voorbeeld in tabel 12.2

g_j	−3	−1	+1	+3	$\Sigma g_j = 0$
g_j^2	9	1	1	9	$\Sigma g_j^2 = 20$
S_j	17	24	25	34	$\Sigma S_j = 100$
$S_j g_j$	−51	−24	25	102	$\Sigma S_j g_j = 52$

$$T_0 = \frac{52}{\sqrt{\frac{10 \times 4 \times 5}{12} 20}} = 2{,}85.$$

Ten behoeve van deze en de volgende toets zijn in tabel 12.4 voor $k = 3(1)26$ de waarden van Σg_j^2 opgenomen, die gelden indien de waarden van j equidistant volgens formule (12.1) worden gekozen.

Tabel 12.4. Waarden van Σg_j^2 ($j = 1, 2, \ldots, k$) voor $k = 3(1)26$ bij volgens (12.1) gekozen waarden van g

Oneven				Even			
k	Σg^2	k	Σg^2	k	Σg^2	k	Σg^2
3	2	15	280	4	20	16	1360
5	10	17	408	6	70	18	1938
7	28	19	570	8	168	20	2660
9	60	21	770	10	330	22	3542
11	110	23	1012	12	572	24	4600
13	182	25	1300	14	910	26	5850

12.3. Toets tegen verloop in k verwante waarnemingsreeksen met dichotome waarnemingen

12.3.1. TOEPASSING

Deze toets kan worden toegepast als de waarnemingen in de k geordende verwante waarnemingsreeksen een dichotomie (met de kenmerken A en B) betreffen. De nulhypothese die kan worden getoetst luidt dat voor elk element de kans op de uitkomst A bij elke waarnemingsreeks constant is, met als alternatieve hypothese dat deze kans voor een overwegend aantal der elementen in de volgende 1, 2, ..., k toe- of afneemt.

Voorbeeld

12.2. Tabel 12.5 geeft de aanwezigheid van *Staphylococcus aureus* in de neusholte van 8 zuigelingen op 12 opeenvolgende tijdstippen (met tussenruimten van 2 dagen); 0 = geen, 1 = wel Staphylococci.

Tabel 12.5. Twaalf geordende verwante waarnemingsreeksen

Zuigeling nr.	Tijdstippen van waarneming												Totaal R_i	R_i^2
	1	2	3	4	5	6	7	8	9	10	11	12		
	−11	−9	−7	−5	−3	−1	+1	+3	+5	+7	+9	+11		
1	1	1	0	0	1	0	0	0	0	0	0	1	4	16
2	1	0	0	0	1	0	0	0	1	0	0	0	3	9
3	1	1	1	1	0	1	1	0	1	1	1	0	9	81
4	0	1	1	1	0	1	0	1	0	1	0	0	6	36
5	1	0	0	1	1	0	0	1	0	0	1	0	5	25
6	0	0	0	1	0	0	0	1	0	0	0	0	2	4
7	1	0	0	1	1	0	1	1	1	1	0	0	7	49
8	1	1	1	0	0	1	0	0	0	1	0	0	5	25
S_j	6	4	3	5	4	3	2	4	3	4	2	1	$41 = S$	$245 = \Sigma R_i^2$
$S_j g_j$	−66	−36	−21	−25	−12	−3	2	12	15	28	18	11	$-77 = \Sigma S_j g_j$	

De nulhypothese is dat voor elke zuigeling de kans op het voorkomen van *Staphylococcus aureus* in de tijd constant is met als alternatieve hypothese dat deze kans voor een overwegend aantal zuigelingen een dalend verloop vertoont. Men toetst met $\alpha = 0{,}05$.

12.3.2. DE TOETSINGSGROOTHEID

Men kent aan de k waarnemingsreeksen op de in 12.1 behandelde wijze waarden van g toe; daar de tijdstippen equidistant zijn worden deze waarden volgens (12.1.1) gekozen (zie tabel 12.5). Vervolgens geeft men alle waarnemingen die kenmerk A (wel staphylococci) hebben opgeleverd met een 1 en alle andere waarnemingen met een 0 aan en bepaalt per waarnemingsreeks de som S_j. Als toetsingsgrootheid kan men nu nemen

$$\underline{T} = \frac{\Sigma S_j g_j}{\sqrt{\frac{k \cdot S - \Sigma R_i^2}{k\,(k-1)} \Sigma g_j^2}} \tag{12.5}$$

waarin S = het totaal aantal enen in de tabel en R_i = het aantal enen dat behoort bij het i-de element (= op de i-de regel); zie weer tabel 12.5. Onder H_0 volgt de grootheid $\underline{T}$ bij benadering een standaardnormale verdeling.

VAN EEDEN (1955, par. 5) komt tot een identieke toets door per element (per rij) de enen als een steekproef van de omvang t_1 en de nullen als een steekproef van de omvang t_2 te beschouwen ($t_1 + t_2 = k$) en de toets van WILCOXON toe te passen. De waarde U_i van de toetsingsgrootheid U van de toets van WILCOXON kan dan worden verkregen door de som te nemen van het aantal nullen dat achter elk der enen staat; voor zuigeling 1 vindt men op deze wijze $U_1 = 2 \times 8 + 1 \times 6 + 1 \times 0 = 22$, met $t_1 = 4$ en $t_2 = 8$. Als toetsingsgrootheid voor de toets tegen verloop neemt zij dan $U' = \Sigma U_i$. Onder H_0 volgt voor elk element de grootheid $\underline{U}_i$ een bij benadering normale verdeling met

gemiddelde $\mu_i = \frac{1}{2}t_1t_2$ en variantie $\sigma_i^2 = t_1t_2(t_1 + t_2 + 1)/12 = t_1t_2(k + 1)/12$, zodat dan de grootheid $\underline{U}'$ eveneens bij benadering normaal verdeeld is met gemiddelde

$$\sigma_{U'} = \Sigma t_1t_2/2$$

en met variantie

$$\sigma^2_{U'} = \frac{k + 1}{12}\,\Sigma t_1t_2.$$

Onder H_0 is dus de grootheid

$$\underline{T} = \frac{2\underline{U}' - \Sigma t_1t_2}{\sqrt{\frac{k + 1}{3}\,\Sigma t_1t_2}} \tag{12.6}$$

bij benadering standaardnormaal verdeeld. Deze toetsingsgrootheid is identiek aan (12.5), maar neemt een tegengesteld teken aan.

De uitvoering van de toetsing verloopt op de wijze die in 12.2.3 is aangegeven.

Bij het voorbeeld in tabel 12.5 vindt men met (12.5)

$$T_0 = \frac{-77}{\sqrt{\frac{12 \times 41 - 245}{12 \times 11}\,572}} = -2{,}35.$$

Er wordt tegen een dalend verloop, dus linkszijdig, getoetst. Daar $P_L(-2{,}35)$ volgens tabel A iets kleiner is dan 0,01 wordt H_0 verworpen. De conclusie is dus dat de kans op het voorkomen van *Staphylococcus aureus* in het algemeen in de tijd een dalend verloop vertoont.

12.4. Toets tegen verloop in k aselecte steekproeven

12.4.1. TOEPASSING

De toets kan worden toegepast als men beschikt over k geordende aselecte steekproeven en de hypothese H_0 wil toetsen, dat zij afkomstig zijn uit populaties met dezelfde verdeling. De alternatieve hypothese luidt dat de gemiddelden van de populaties in de volgorde 1, 2, . . ., k een overwegend stijgend of dalend verloop vertonen.

Voorbeeld

12.3. De gegevens in tabel 12.6 betreffen het ceruloplasminegehalte (y) bij 42 donores, die zijn ingedeeld in vier leeftijdsgroepen van dezelfde breedte.

Men wil de hypothese H_0 toetsen dat de leeftijdsgroepen dezelfde verdeling van y bezitten met als alternatieve hypothese dat met toenemende leeftijd gemiddeld een stijgend verloop van de y-waarden optreedt. Men kiest $\alpha = 0{,}05$.

Tabel 12.6. Vier geordende aselecte steekproeven

Leeftijdsgroepen							
20–29		30–39		40–49		50–59	
y	Rang-nummer	y	Rang-nummer	y	Rang-nummer	y	Rang-nummer
77	1	81	4,5	81	4,5	88	12
79	2	81	4,5	90	18	89	15,5
81	4,5	82	7,5	94	23,5	100	29,5
82	7,5	88	12	95	25,5	103	32,5
85	9	93	22	97	27	112	39
88	12	94	23,5	100	29,5	163	42
88	12	98	28	105	34		
88	12	102	31	109	36		
89	15,5	103	32,5				
90	18	118	40				
90	18	120	41				
92	20,5						
92	20,5						
95	25,5						
106	35						
110	37,5						
110	37,5						
17	288	11	246,5	8	198	6	170,5
n_1	S_1	n_2	S_2	n_3	S_3	n_4	S_4

12.4.2. DE TOETSINGSGROOTHEID

Volgens de in 12.1 gegeven richtlijnen kent men de steekproeven waarden van g toe. Daar de leeftijdsgroepen dezelfde breedte hebben kiezen wij deze bij het voorbeeld equidistant, dus resp. als —3, —1, +1 en +3. Vervolgens kent men — evenals bij de toets van KRUSKAL en WALLIS — de waarnemingen, onafhankelijk van de steekproef waarin zij voorkomen, van laag naar hoog de rangnummers 1, 2, . . ., n toe ($n = n_1 + n_2 + \ldots + n_k$).

Zijn er geen of slechts weinig gelijken, dan kan men volgens DE JONGE (1960) als toetsingsgrootheid nemen

$$\underline{T} = \frac{2 \cdot \Sigma S_j g_j - (n+1)\,\Sigma n_j g_j}{\sqrt{\frac{n\,(n+1)}{3}\,[\Sigma n_j g_j^2 - (\Sigma n_j g_j)^2/n]}}. \tag{12.7}$$

De wijze waarop deze grootheid moet worden berekend blijkt uit tabel 12.7.

Bij steekproeven van dezelfde omvang m is

$$n = km \text{ en } \Sigma n_j g_j = m \,.\, \Sigma g_j = 0,$$

zodat (12.7) overgaat in

(12.8) $$\underline{T} = \frac{2 \cdot \Sigma S_j g_j}{\sqrt{\frac{mn\,(n+1)}{3}\,\Sigma g_j^2}}.$$

Treden betrekkelijk veel gelijken op, dan dient men in de formules (12.7) en (12.8) onder het wortelteken in de noemer $n\,(n+1)/3$ te vervangen door

(12.9) $$\frac{n^3 - D}{3\,(n-1)}$$

waarin D wordt berekend zoals in 10.6. is aangegeven.

Onder H_0 volgt de grootheid $\underline{T}$ bij benadering een standaardnormale verdeling. De toetsing kan dus ook hier worden uitgevoerd op de in 12.2.3 vermelde wijze.

De berekeningen die bij toepassing van de toets op het voorbeeld moeten worden uitgevoerd zijn in tabel 12.7 opgenomen. Men vindt eerst de berekening van $\Sigma S_j g_j$, $\Sigma n_j g_j$ en $\Sigma n_j g_j^2$ en vervolgens de berekening van D uit de frequentieverdeling van de groepen gelijken (die tijdens het toekennen van de rangnummers kan worden opgesteld). Er zijn 14 waarnemingen die niet tot een groep gelijken behoren en dus een bijdrage van $14 \times 1^3 = 14$ tot D leveren; er zijn 8 groepen van 2 gelijken, waarvan de bijdrage tot D gelijk is aan $8 \times 2^3 = 64$, enz.

Tabel 12.7. Berekeningen bij het voorbeeld in tabel 12.6

Steekproef	1	2	3	4	
g_j	−3	−1	+1	+3	$\Sigma g_j = 0$
g_j^2	9	1	1	9	
S_j	288	246,5	198	170,5	$\Sigma S_j = 903$
n_j	17	11	8	6	$n = \Sigma n_j = 42$
$S_j g_j$	−864	−246,5	198	511,5	$\Sigma S_j g_j = -401$
$n_j g_j$	−51	−11	8	18	$\Sigma n_j g_j = -36$
$n_j g_j^2$	153	11	8	54	$\Sigma n_j g_j^2 = 226$

Groepen gelijken						Controle
t	1	2	3	4	5	$\Sigma S_j = \frac{n\,(n+1)}{2} = 903$
f	14	8	1	1	1	
ft	14	16	3	4	5	$\Sigma ft = 42 = n$
ft^3	14	64	27	64	125	$\Sigma ft^3 = 294 = D$

Formule (12.7) levert tenslotte, inclusief de door (12.9) gegeven correctie voor gelijken, op

$$T_0 = \frac{2\,(-401) - (43)\,(-36)}{\sqrt{\frac{(42^3 - 294)\,[226 - (-36)^2/42]}{3 \times 41}}} = 2{,}18.$$

Er wordt tegen een stijgend verloop, dus rechtséénzijdig, getoetst met $\alpha = 0{,}05$. Uit tabel A blijkt dat de rechtse kans van T_0 ongeveer 0,015 is, zodat de nulhypothese wordt verworpen. De conclusie luidt dus dat de ceruloplasminegehalten bij toenemende leeftijd gemiddeld hogere waarden aannemen.

Merk op dat de correctie voor gelijken bij dit voorbeeld met slechts enkele grotere groepen gelijken praktisch niet van betekenis is. Laat men de correctie achterwege, dan komt dit erop neer dat men $D = n = 43$ in plaats van $D = 294$ neemt; de op twee decimalen afgeronde waarde T_0 verandert hierdoor niet!

12.4.3. UITSCHAKELING VAN DE INVLOED VAN VERSCHILLEN IN STEEKPROEFOMVANG

Door TERPSTRA (1955—56) is voor k aselecte steekproeven een toets tegen verloop ontwikkeld, waarbij de invloed van verschillen in de omvang der steekproeven wordt uitgeschakeld. Deze toets, die o.m. door RÜMKE en VAN EEDEN (1961) wordt beschreven, vereist bij een wat groter aantal steekproeven veel rekenwerk, daar men moet beginnen met toepassing van de toets van WILCOXON op elk steekproefpaar met $i < j$, d.i. bijv. bij $k = 4$ steekproeven op de $C_2^4 = 6$ paren (1,2), (1,3), (1,4), (2,3), (2,4) en (3,4). Bij 6 steekproeven moet men dus reeds $C_2^6 = 15$ toetsen van WILCOXON toepassen.

In het algemeen is een toetsing tegen verloop waarbij met verschillen in steekproefomvang geen rekening wordt gehouden niet gebruikelijk en ons inziens ook niet gewenst.[1] Desgewenst kan men zo'n toetsing echter — met veel minder rekenwerk dan bij de toets van TERPSTRA — uitvoeren door als toetsingsgrootheid (12.7) te nemen met g_j/n_j in plaats van g_j. Daar $\Sigma n_j \cdot g_j/n_j = \Sigma g_j = 0$ en $\Sigma n_j \cdot (g_j/n_j)^2 = \Sigma g_j^2/n_j$ wordt daarbij de toetsingsgrootheid

$$\underline{T} = \frac{2 \cdot \Sigma \underline{S}_j \frac{g_j}{n_j}}{\sqrt{\frac{n(n+1)}{3} \Sigma \frac{g_j^2}{n_j}}} \tag{12.10}$$

Evenals bij (12.7) kan men de correctie voor gelijken aanbrengen door in de noemer $n(n+1)/3$ door $(n^3 - D)/3(n-1)$ te vervangen. Als alle steekproeven dezelfde omvang m bezitten is steeds $n_j = m$ en gaat formule (12.10) in (12.8) over.

Tabel 12.8 bevat de berekening bij toepassing van de verlooptoets volgens (12.10) op de gegevens in tabel 12.6 (de correctie voor gelijken is hierbij, wegens de reeds geconstateerde geringe betekenis, achterwege gelaten).

[1] Zie de beschouwing aan het eind van 12.5.

Tabel 12.8. Verlooptoets met uitschakeling van de steekproefomvang, toegepast op de gegevens in tabel 12.6

g_j	-3	-1	$+1$	$+3$	$\Sigma g_j = 0$
n_j	17	11	8	6	$\Sigma n_j = 42 = n$
g_j/n_j	$\frac{-3}{17}$	$\frac{-1}{11}$	$\frac{1}{8}$	$\frac{3}{6} = \frac{1}{2}$	
S_j	288	246,5	198	170,5	$\Sigma S_j = 903$
$S_j g_j/n_j$	$\frac{-3 \times 288}{17}$	$\frac{-246,5}{11}$	$\frac{198}{8}$	$\frac{170,5}{2}$	$\Sigma = 36,7674$
g_j^2/n_j	$\frac{9}{17}$	$\frac{1}{11}$	$\frac{1}{8}$	$\frac{9}{6}$	$\Sigma = 2,2453$

$$T_0 = \frac{2 \times 36,7674}{\sqrt{\frac{42 \times 43}{3} 2,2453}} = 2,00$$

Opmerking. Wanneer men reeds beschikbaar waarnemingsmateriaal bewerkt, zoals bij het voorbeeld in tabel 12.6 het geval is, kan men de omvang van de 'steekproeven' niet zelf kiezen, zodat men gewoonlijk met verschillen in steekproefomvang te doen heeft. Stelt men echter zelf de proefopzet van een onderzoek vast, dan verdient het in het algemeen aanbeveling de steekproeven even groot te nemen, zodat zich geen problemen betreffende het al of niet uitschakelen van de invloed van verschillen in steekproefomvang voordoen.

12.5. Toets tegen verloop in een $k \times 2$ tabel

Wanneer de waarnemingen in k geordende aselecte steekproeven een dichotomie (met de kenmerken A en B) betreffen, verkrijgt men gegevens van de in tabel 12.9 gegeven vorm.

Tabel 12.9. k geordende steekproeven uit dichotome populaties

Kenmerk	Geordende steekproeven 1	2	...	j	...	k	Totaal
A	f_1	f_2	...	f_j	...	f_k	$r = \Sigma f_j$
B	$n_1 - f_1$	$n_2 - f_2$	...	$n_j - f_j$	...	$n_k - f_k$	$s = \Sigma (n_j - f_j)$
Totaal	n_1	n_2	...	n_j	...	n_k	$n = r + s$

Met de in 12.4 behandelde toets kan men de nulhypothese toetsen dat de steekproeven uit dezelfde dichotomie afkomstig zijn, met als alternatieve hypothese dat zij stammen uit populaties waarin de kansen van kenmerk A in de volgorde $1, 2, \ldots, k$ een stijgend of dalend verloop bezitten. De toetsingsgrootheid volgens formule (12.7) kan nu, bij gebruik van de in tabel 12.9 gegeven notatie, worden geschreven als[1]

[1] Zie COCHRAN (1954) en ARMITAGE (1955).

$$(12.11) \qquad T = \frac{\Sigma f_j g_j - r \cdot \Sigma n_j g_j / n}{\sqrt{\frac{rs}{n(n-1)} [\Sigma n_j g_j^2 - (\Sigma n_j g_j)^2/n]}}$$

Voor k steekproeven van dezelfde omvang m (waarbij dus $km = n$) kan (12.11) worden vereenvoudigd tot

$$(12.12) \qquad T = \frac{\Sigma f_j g_j}{\sqrt{\frac{rs}{k(n-1)} \Sigma g_j^2}}$$

(bij equidistante g_j kan men Σg_j^2 weer uit tabel 12.4 aflezen).

Voorbeeld

12.4. Tabel 12.10 bevat gegevens over het voorkomen van folliculosis bij vier groepen kinderen van 7 t/m 9 jaar, die beschouwd kunnen worden als aselecte steekproeven uit bevolkingsgroepen die volgens bepaalde criteria in de volgorde van 1 t/m 4 in welstand toenemen. Men wil (met $\alpha = 0{,}05$) de hypothese H_0 toetsen dat de kans op folliculosis voor de vier welstandsgroepen gelijk is, met als alternatieve hypothese dat deze kans bij toenemende welstand afneemt.

Tabel 12.10. Folliculosis bij kinderen uit vier welstandsgroepen

Folliculosis	Welstand 1	2	3	4	Totaal
Wel (f_j)	16	34	42	8	100 $= r$
Geen	20	74	102	28	224 $= s$
Totaal (n_j)	36	108	144	36	324 $= n$
g_j	−3	−1	+1	+3	Σg = 0
$f_j g_j$	−48	−34	42	24	Σfg = −16
$n_j g_j$	−108	−108	144	108	Σng = 36
$n_j g_j^2$	324	108	144	324	Σng^2 = 900

De berekeningen die nodig zijn voor de toepassing van de toets tegen verloop op deze $k \times 2$ tabel zijn beneden in tabel 12.10 uitgevoerd. De waarden van g zijn hier arbitrair equidistant gekozen. Formule (12.11) levert op

$$T_0 = \frac{-16 - 100 \times 36/324}{\sqrt{\frac{100\,(224)\,[900 - (36)^2/324]}{324 \times 323}}} = -1{,}96,$$

met $P_R = 0{,}025$. De nulhypothese wordt dus verworpen ten gunste van een dalend verloop.

VAN EEDEN (1955) bespreekt de toetsing waarbij de invloed van verschillen in steekproefomvang wordt uitgeschakeld. Haar toetsingsgrootheid is (gebruikmakend van onze notatie)

$$T = \frac{\Sigma f_j g_j / n_j}{\sqrt{\frac{rs}{n(n-1)} \Sigma g_j^2 / n_j}} \tag{12.13}$$

Als de steekproeven dezelfde omvang m bezitten gaat (12.13) over in (12.12).

Wij zullen het verschil tussen de toetsing met en zonder uitschakeling van de verschillen in steekproefomvang met enkele voorbeelden illustreren. Het eerste is ontleend aan RÜMKE en VAN EEDEN (1961) en betreft de aantallen opgenomen en overleden patiënten met tuberculeuze meningitis in de jaren 1947–1956 in tabel 12.11 (T.N.O.-nieuws 12 (1957), 500).

Tabel 12.11

Jaar	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	Totaal
Niet overleden	1	6	9	10	15	23	41	24	11	5	145
Overleden	2	12	13	12	10	18	7	4	0	2	80
Opgenomen	3	18	22	22	25	41	48	28	11	7	225

Men wil onderzoeken of voor de opgenomen patiënten met deze aandoening tussen 1947 en 1956 de overlevingskans een (stijgend of dalend) verloop heeft.

De waarden van g worden vanzelfsprekend equidistant genomen en zijn dus respectievelijk $-9, -7, -5, \ldots, +5, +7, +9$. Wij laten het aan de lezer over om de berekeningen uit te voeren; met (12.11) vindt men

$$T_0 = \frac{247 - 145 \times 111/225}{\sqrt{\frac{145 \times 80}{225 \times 224}\,[4177 - (111)^2/225]}} = 5{,}70$$

en met (12.13)

$$T_0 = \frac{11{,}4955}{\sqrt{\frac{145 \times 80}{225 \times 224}\,48{,}4382}} = 3{,}44.$$

Het grote verschil tussen beide toetsingsuitkomsten wordt verklaarbaar als wij de *percentages* overlevenden beschouwen:

Jaar	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956
% Overleving	33	33	41	45	60	56	85	86	100	71
n_j	3	18	22	22	25	41	48	28	11	7

De toets volgens (12.13) berust in wezen op het verloop van deze percentages, zonder daarbij rekening te houden met de totalen n_j waarop deze gebaseerd zijn. Hierdoor zal bijvoorbeeld de uitkomst van 1956, die het stijgende verloop procentueel sterk doorbreekt, ernstig storend werken. De toets volgens (12.11) brengt daarentegen het feit, dat er in dat jaar in totaal slechts 7 waarnemingen zijn (zodat het toeval een betrekkelijk grote rol kan spelen), mede in rekening en wordt zodoende door de afwijkende uitkomst minder beïnvloed.

Tabel 12.12 bevat een gefingeerd voorbeeld met als waargenomen percentages A in drie steekproeven resp. 25%, 23% en 60%. De eerste twee percentages berusten op vrij grote steekproeven en zij verschillen weinig. Het sterk afwijkende percentage in de derde steekproef heeft echter slechts betrekking op 5 waarnemingen. Hier levert de toets volgens (12.13) $T_0 = 1{,}76$ met een rechtse kans van ongeveer 0,04 op. De toets volgens (12.11) leidt tot $T_0 = 0{,}52$ met een rechtse kans van ongeveer 0,30. Het behoeft naar onze mening geen betoog dat men in de praktijk op grond van deze gegevens niet tot een stijgend verloop zal willen besluiten.

Tabel 12.12

	Steekproeven			
	1	2	3	
A	25	22	3	50
B	75	73	2	150
	100	95	5	200

12.6. Opgaven

12.1. Onderstaande waarnemingen betreffen de lichaamstemperatuur van 6 personen, gemeten op vrijdag te 9.00 uur in 7 opeenvolgende weken. Toets de hypothese dat de waargenomen verschillen tussen de weken toevallig zijn met als alternatieve hypothese dat een stijgend verloop optreedt. Kies $\alpha = 0{,}01$.

Pers. nr	Opeenvolgende weken						
	1	2	3	4	5	6	7
1	36,27	36,20	36,47	36,54	36,58	37,00	37,12
2	35,90	36,04	36,18	36,12	36,17	36,40	36,22
3	35,78	35,85	35,70	36,10	36,23	36,20	35,98
4	35,95	35,74	36,26	36,10	36,31	36,20	36,24
5	35,81	35,99	35,97	36,08	36,17	36,28	36,25
6	35,70	35,49	35,78	35,68	36,00	35,88	35,93

12.2. De (gefingeerde) gegevens in onderstaande tabel betreffen een onderzoek naar het voorkomen van diphterietoxine in speekselmonsters van 9 patiënten, bij wie de diagnose diphterie bacteriologisch werd bevestigd, op 9 opeenvolgende tijdstippen (+ = positief, — = negatief resultaat van de toxineproef). Onderzoek (met $\alpha = 0{,}05$) of de kans op een positieve reactie in het algemeen in de tijd afneemt.

Patiënten	Tijdstippen								
	1	2	3	4	5	6	7	8	9
1	+	+	—	+	+	—	+	+	—
2	—	+	+	—	+	—	—	+	—
3	+	+	—	—	—	—	—	—	—
4	+	+	+	+	—	+	+	—	+
5	+	—	—	+	—	+	—	—	—
6	—	—	+	+	+	—	—	+	—
7	+	+	+	—	—	—	+	—	—
8	+	—	+	—	—	—	—	—	—
9	—	—	—	—	—	+	+	—	—

12.3. Een aselecte steekproef van 40 kinderen uit een bepaalde populatie wordt naar voedingstoestand (volgens dieetopname) in vier categorieën gesplitst.

Bij deze kinderen werden de volgende waarden van carotinoïden (γ per 10 cc serum) waargenomen (fictieve gegevens):

Voedingstoestand	Waarden								
1. Onvoldoende	3,8	3,9	4,9	5,8	6,4	7,8	8,6		
2. Matig	3,9	4,8	5,9	6,3	6,5	7,7	9,8	10,5	
3. Voldoende	4,1	4,7	6,3	6,8	6,9	7,1	7,3	8,2	
	8,7	8,8	9,1	9,7	10,8	11,1	12,4	12,8	
4. Goed	4,6	5,7	6,7	8,7	9,4	11,3	12,6	12,9	15,3

Onderzoek of bij verbetering van de voedingstoestand een stijgend verloop optreedt. Kies $\alpha = 0{,}05$.

Tijdstip van ontstaan	Totaal	Binoculair zien
1e levensjaar	31	4
2e levensjaar	50	15
3e levensjaar	29	10
na 3e jaar	29	19
Totaal	139	48

12.4. Uit N.T.v.G. 99 (1955), 1948. Nevenstaande tabel bevat gegevens over het herstel van het binoculair zien na operatie bij patiëntjes met een restscheelzienhoek van minder dan 5°, geordend naar de tijd van het ontstaan van het scheelzien. Toets tegen verloop met $\alpha = 0{,}01$.

12.5. Uit N.T.v.G. 102 (1958), 759. Tabel III. Halsklierextirpatie wegens metastasen van carcinomen gezeteld in het slijmvliesgebied van hoofd en hals. Vijf-jaars-overleving:

Grootte van het primaire gezwel	Totaal aantal patiënten	Na vijf jaar in leven
Groot	88	18 = 20,4 pct
Matig	85	21 = 24,7 pct
Klein	36	12 = 33,3 pct

Kunt u zich verenigen met de conclusie '... dat degenen met kleine, primaire gezwellen een betere kans op vijf-jaars-overleving hebben dan de patiënten met grote gezwellen'?

BIJLAGE I

TABELLENVERZAMELING

A. 1. Linkse kansen van de standaardnormaal verdeelde grootheid $\underline{T}$, $T = 0{,}00\ (0{,}01)\ 2{,}99$. 351
2. Idem, $T = 3{,}0\ (0{,}1)\ 6{,}9$ 352

B. Fractielen van de standaardnormaal verdeelde grootheid $\underline{T}$ 352

C. Fractielen van STUDENT-verdelingen ($\underline{t}$) 353

D. Fractielen van χ^2-verdelingen 354

E. Fractielen van de verdeling van de correlatiecoëfficiënt $\underline{r}$ voor $\varrho = 0$ 356

F. Fractielen van $\underline{F}$-verdelingen
1. $F_{0,95}$ 357
2. $F_{0,975}$ 358
3. $F_{0,99}$ 359
4. $F_{0,995}$ 360

G. De z-transformatie van FISHER:
$\underline{z} = \frac{1}{2}\log_e(1 + \underline{r})/(1 - \underline{r}) = 1{,}1513\,[\log_{10}(1 + \underline{r}) - \log_{10}(1 - \underline{r})]$ . . . 361

H. De toets van HARTLEY
Rechter kritieke waarden van de grootheid $\underline{F}_{max} = \underline{s}^2_{max}/\underline{s}^2_{min}$, $\alpha = 0{,}01$ en $0{,}05$. 363

I. De tekentoets
Linker kritieke waarden van $\underline{x}$ bij tweezijdige toetsing, $\alpha = 0{,}01$, $0{,}02$, $0{,}05$ en $0{,}10$. 364

J. De rang-tekentoets (symmetrietoets van WILCOXON)
Rechter kritieke waarden van $\underline{V}$ bij tweezijdige toetsing, $\alpha = 0{,}01$, $0{,}02$, $0{,}05$ en $0{,}10$ 365

K. De toets van FRIEDMAN (methode der m rangschikkingen)
1. Exacte rechtse overschrijdingskansen van $\underline{K}$ voor kleine waarden van k en n 366
2. Rechter kritieke waarden van $\underline{K}$, $\alpha = 0{,}05$ 367

L. De toets van WILCOXON
Linker en rechter kritieke waarden van $\underline{W}$ bij tweezijdige toetsing, $\alpha = 0{,}01$, $0{,}02$, $0{,}05$ en $0{,}10$ 368

M. De toets van KRUSKAL & WALLIS
Exacte rechtse overschrijdingskansen van $\underline{H}$ voor $k = 3$ en kleine steekproeven 372

N. De rangcorrelatietoets van SPEARMAN
Linker en rechter kritieke waarden van $\underline{R}$ bij tweezijdige toetsing, $\alpha = 0{,}01$, $0{,}02$, $0{,}05$ en $0{,}10$ 373

O. De rangcorrelatietoets van KENDALL
Rechter kritieke waarden van $\underline{S}$ bij tweezijdige toetsing, $\alpha = 0{,}01$, $0{,}02$, $0{,}05$ en $0{,}10$. 374

P. De serietoets op de waarnemingen
1. Linker en rechter kritieke waarden van $\underline{u}$ bij tweezijdige toetsing, $m_1 \neq m_2$, $\alpha = 0{,}01$, $0{,}02$, $0{,}05$ en $0{,}10$ 375
2. Linker en rechter kritieke waarden van $\underline{u}$ bij tweezijdige toetsing en $m_1 = m_2 = \frac{1}{2}n$, $\alpha = 0{,}01$, $0{,}02$, $0{,}05$ en $0{,}10$ 377

Q. De serietoets op verschillen tussen opeenvolgende waarnemingen. Linker en rechter kritieke waarden van $\underline{v}$ bij éénzijdige toetsing, $\alpha = 0{,}01$ en $0{,}05$. 378

R. Betrouwbaarheidsgrenzen van percentages, $1 - \alpha = 0{,}95$ en $0{,}99$
1. Voor $n = 5$, 6, 7, 8, 9, 10, 12, 14, 16, 20, 25, 30, 40, 50 en 75 op basis van x = aantal elementen met kenmerk A 379
2. Voor $n = 100$, 150, 200, 300, 500 en 1000 op basis van x = aantal elementen met kenmerk A 380
3. Voor $n = 100$, 200, 500 en 1000 op basis van het percentage elementen met kenmerk A 381

S. De toets van KOLMOGOROV-SMIRNOV
Kritieke waarden van $\underline{D}$ bij tweezijdige toetsing, $\alpha = 0{,}01$, $0{,}05$, $0{,}10$, $0{,}15$ en $0{,}20$. 382

T. De normaliteitstoetsen van FISHER
Linker en rechter kritieke waarden van de grootheden $\underline{g}_1$ en $\underline{g}_2$ bij éénzijdige toetsing en $\alpha = 0{,}05$ en $0{,}01$ 383

U. De uitbijtertoets van DIXON
Rechter kritieke waarden van de grootheid $\underline{r}$ voor $\alpha = 0{,}01$, $0{,}02$ en $0{,}05$ en $0{,}10$ 384

V. De bgsin transformatie
1. Waarden van de grootheid $\underline{y} = 2$ bgsin $\sqrt{\underline{p}}$ 385
2. Waarden van y (0) bij $p = 0$ en van y (1) bij $p = 1$ voor $10 < n < 50$ 387

W. De vierkantswortel transformatie van FREEMAN en TUKEY
Waarden van $y = \sqrt{x} + \sqrt{x+1}$ en van $y^2 = (\sqrt{x} + \sqrt{x+1})^2$ . . 387

X. Pseudo-normaal verdeelde modelpopulatie 388

Y. Aselecte getallen (toevalcijfers) 389

Z. Aselecte permutaties (lotingsgetallen) 390

AA. Rechter kritieke waarden van de grootheid $\underline{w}$ voor $\alpha = 0{,}05$ en $0{,}01$. 391

BB. Rankits in aselecte steekproeven van n elementen uit een standaard-normale verdeling 393

CC. Betrouwbaarheidsgrenzen van de mediaan 394

TABEL A–1

LINKSE KANSEN VAN DE STANDAARDNORMAAL VERDEELDE GROOTHEID T

$T=\frac{x-\mu}{\sigma}$		$P_L(T)$										$T=\frac{x-\mu}{\sigma}$	
		0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09		
0,0	—	0,5000	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641	—	0,0
	+	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359	+	
0,1	—	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247	—	0,1
	+	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753	+	
0,2	—	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859	—	0,2
	+	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141	+	
0,3	—	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483	—	0,3
	+	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517	+	
0,4	—	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121	—	0,4
	+	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879	+	
0,5	—	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776	—	0,5
	+	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224	+	
0,6	—	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451	—	0,6
	+	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549	+	
0,7	—	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148	—	0,7
	+	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852	+	
0,8	—	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867	—	0,8
	+	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133	+	
0,9	—	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611	—	0,9
	+	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389	+	
1,0	—	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379	—	1,0
	+	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621	+	
1,1	—	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170	—	1,1
	+	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830	+	
1,2	—	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985	—	1,2
	+	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015	+	
1,3	—	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823	—	1,3
	+	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177	+	
1,4	—	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681	—	1,4
	+	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319	+	
1,5	—	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559	—	1,5
	+	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441	+	
1,6	—	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455	—	1,6
	+	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545	+	
1,7	—	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367	—	1,7
	+	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633	+	
1,8	—	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294	—	1,8
	+	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706	+	
1,9	—	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233	—	1,9
	+	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767	+	
2,0	—	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183	—	2,0
	+	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817	+	
2,1	—	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143	—	2,1
	+	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857	+	
2,2	—	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110	—	2,2
	+	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890	+	
2,3	—	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084	—	2,3
	+	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916	+	
2,4	—	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064	—	2,4
	+	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936	+	
2,5	—	0,00621	0,00604	0,00587	0,00570	0,00554	0,00539	0,00523	0,00508	0,00494	0,00480	—	2,5
	+	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520	+	
2,6	—	0,00466	0,00453	0,00440	0,00427	0,00415	0,00402	0,00391	0,00379	0,00368	0,00357	—	2,6
	+	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643	+	
2,7	—	0,00347	0,00336	0,00326	0,00317	0,00307	0,00298	0,00289	0,00280	0,00272	0,00264	—	2,7
	+	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736	+	
2,8	—	0,00256	0,00248	0,00240	0,00233	0,00226	0,00219	0,00212	0,00205	0,00199	0,00193	—	2,8
	+	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807	+	
2,9	—	0,00187	0,00181	0,00175	0,00169	0,00164	0,00159	0,00154	0,00149	0,00144	0,00139	—	2,9
	+	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861	+	

Uit PEARSON en HARTLEY (1954).

Tabel A–2

Linkse kansen van de standaardnormaal verdeelde grootheid T

$T = \frac{x-\mu}{\sigma}$		$P_L(T)$									
		0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
3	—	0,00135	$0,0^3968$	$0,0^3687$	$0,0^3483$	$0,0^3337$	$0,0^3233$	$0,0^3159$	$0,0^3108$	$0,0^4723$	$0,0^4481$
	+	0,99865	$0,9^3032$	$0,9^3313$	$0,9^3517$	$0,9^3663$	$0,9^3767$	$0,9^3841$	$0,9^3892$	$0,9^4277$	$0,9^4519$
4	—	$0,0^4317$	$0,0^4207$	$0,0^4133$	$0,0^5854$	$0,0^5541$	$0,0^5340$	$0,0^5211$	$0,0^5130$	$0,0^6793$	$0,0^6479$
	+	$0,9^4683$	$0,9^4793$	$0,9^4867$	$0,9^5146$	$0,9^5459$	$0,9^5660$	$0,9^5789$	$0,9^5870$	$0,9^6207$	$0,9^6521$
5	—	$0,0^5287$	$0,0^6170$	$0,0^7996$	$0,0^7579$	$0,0^7333$	$0,0^7190$	$0,0^7107$	$0,0^8599$	$0,0^8332$	$0,0^8182$
	+	$0,9^5713$	$0,9^6830$	$0,9^7004$	$0,9^7421$	$0,9^7667$	$0,9^7810$	$0,9^7893$	$0,9^8401$	$0,9^8668$	$0,9^8818$
6	—	$0,0^6987$	$0,0^6530$	$0,0^6282$	$0,0^6149$	$0,0^{10}777$	$0,0^{10}402$	$0,0^{10}206$	$0,0^{10}104$	$0,0^{11}523$	$0,0^{11}260$
	+	$0,9^6013$	$0,9^6470$	$0,9^6718$	$0,9^6851$	$0,9^{10}223$	$0,9^{10}598$	$0,9^{10}794$	$0,9^{10}896$	$0,9^{11}477$	$0,9^{11}740$

Voorbeeld: $T = -4$, $P_L(-4) = 0,0000317$; $T = +4$, $P_L(+4) = 0,9999683$.

Tabel B

Fractielen van de standaardnormaal verdeelde grootheid T

$P_L(T)$	T	$P_L(T)$	T	$P_L(T)$	T	$P_L(T)$	T	$P_L(T)$	T
0,0001	–3,719	0,045	–1,695	0,280	–0,583	0,700	0,524	0,950	1,645
0,0002	–3,540	0,050	–1,645	0,300	–0,524	0,720	0,583	0,955	1,695
0,0003	–3,432	0,055	–1,598	0,320	–0,468	0,740	0,643	0,960	1,751
0,0004	–3,353	0,060	–1,555	0,340	–0,412	0,750	0,6745	0,965	1,812
0,0005	–3,291	0,065	–1,514	0,360	–0,358	0,760	0,706	0,970	1,881
0,001	–3,090	0,070	–1,476	0,380	–0,305	0,780	0,772	0,975	1,960
0,002	–2,878	0,075	–1,440	0,400	–0,253	0,800	0,842	0,980	2,054
0,003	–2,748	0,080	–1,405	0,420	–0,202	0,820	0,915	0,985	2,170
0,004	–2,652	0,085	–1,372	0,440	–0,151	0,840	0,994	0,990	2,326
0,005	–2,576	0,090	–1,341	0,460	–0,100	0,860	1,080	0,991	2,366
0,006	–2,512	0,095	–1,311	0,480	–0,050	0,880	1,175	0,992	2,409
0,007	–2,457	0,100	–1,282	0,500	0,000	0,900	1,282	0,993	2,457
0,008	–2,409	0,120	–1,175	0,520	0,050	0,905	1,311	0,994	2,512
0,009	–2,366	0,140	–1,080	0,540	0,100	0,910	1,341	0,995	2,576
0,010	–2,326	0,160	–0,994	0,560	0,151	0,915	1,372	0,996	2,652
0,015	–2,170	0,180	–0,915	0,580	0,202	0,920	1,405	0,997	2,748
0,020	–2,054	0,200	–0,842	0,600	0,253	0,925	1,440	0,998	2,878
0,025	–1,960	0,220	–0,772	0,620	0,305	0,930	1,476	0,999	3,090
0,030	–1,881	0,240	–0,706	0,640	0,358	0,935	1,514	0,9995	3,291
0,035	–1,812	0,250	–0,6745	0,660	0,412	0,940	1,555	0,9996	3,353
0,040	–1,751	0,260	–0,643	0,680	0,468	0,945	1,598	0,9999	3,719

Uit Walker en Lev (1953).

TABEL C. FRACTIELEN VAN STUDENT-VERDELINGEN (t)

ν	$t_{0,75}$	$t_{0,90}$	$t_{0,95}$	$t_{0,975}$	$t_{0,99}$	$t_{0,995}$	$t_{0,999}$	$t_{0,9995}$	ν
1	1,000	3,078	6,314	12,706	31,821	63,657	318,31	636,62	1
2	0,816	1,886	2,920	4,303	6,965	9,925	22,326	31,598	2
3	0,765	1,638	2,353	3,182	4,541	5,841	10,213	12,924	3
4	0,741	1,533	2,132	2,776	3,747	4,604	7,173	8,610	4
5	0,727	1,476	2,015	2,571	3,365	4,032	5,893	6,869	5
6	0,718	1,440	1,943	2,447	3,143	3,707	5,208	5,959	6
7	0,711	1,415	1,895	2,365	2,998	3,499	4,785	5,408	7
8	0,706	1,397	1,860	2,306	2,896	3,355	4,501	5,041	8
9	0,703	1,383	1,833	2,262	2,821	3,250	4,297	4,781	9
10	0,700	1,372	1,812	2,228	2,764	3,169	4,144	4,587	10
11	0,697	1,363	1,796	2,201	2,718	3,106	4,075	4,437	11
12	0,695	1,356	1,782	2,179	2,681	3,055	3,930	4,318	12
13	0,694	1,350	1,771	2,160	2,650	3,012	3,852	4,221	13
14	0,692	1,345	1,761	2,145	2,624	2,977	3,787	4,140	14
15	0,691	1,341	1,753	2,131	2,602	2,947	3,733	4,073	15
16	0,690	1,337	1,746	2,120	2,583	2,921	3,686	4,015	16
17	0,689	1,333	1,740	2,110	2,567	2,898	3,646	3,965	17
18	0,688	1,330	1,734	2,101	2,552	2,878	3,610	3,922	18
19	0,688	1,328	1,729	2,093	2,539	2,861	3,579	3,883	19
20	0,687	1,325	1,725	2,086	2,528	2,845	3,552	3,850	20
21	0,686	1,323	1,721	2,080	2,518	2,831	3,527	3,819	21
22	0,686	1,321	1,717	2,074	2,508	2,819	3,505	3,792	22
23	0,685	1,319	1,714	2,069	2,500	2,807	3,485	3,767	23
24	0,684	1,318	1,711	2,064	2,492	2,797	3,467	3,745	24
25	0,684	1,316	1,708	2,060	2,485	2,787	3,450	3,725	25
26	0,684	1,315	1,706	2,056	2,479	2,779	3,435	3,707	26
27	0,684	1,314	1,703	2,052	2,473	2,771	3,421	3,690	27
28	0,683	1,313	1,701	2,048	2,467	2,763	3,408	3,674	28
29	0,683	1,311	1,699	2,045	2,462	2,756	3,396	3,659	29
30	0,683	1,310	1,697	2,042	2,457	2,750	3,385	3,646	30
35	0,682	1,306	1,690	2,030	2,438	2,724	3,341	3,592	35
40	0,681	1,303	1,684	2,021	2,423	2,704	3,307	3,551	40
50	0,679	1,300	1,677	2,009	2,403	2,678	3,262	3,495	50
60	0,679	1,296	1,671	2,000	2,390	2,660	3,232	3,460	60
90	0,677	1,291	1,662	1,987	2,368	2,632	3,184	3,402	90
120	0,677	1,289	1,658	1,980	2,358	2,617	3,160	3,374	120
160	0,676	1,290	1,654	1,975	2,350	2,607	3,139	3,353	160
200	0,676	1,286	1,652	1,972	2,345	2,601	3,131	3,340	200
∞	0,674	1,282	1,645	1,960	2,326	2,576	3,090	3,291	∞
P_R	0,25	0,10	0,05	0,025	0,01	0,005	0,001	0,0005	P_R
P_D	0,50	0,20	0,10	0,05	0,02	0,01	0,002	0,001	P_D

Gedeeltelijk uit PEARSON & HARTLEY (1954) en HALD (1952).

TABEL D. FRACTIELEN VAN χ^2-VERDELINGEN *

ν	$\chi^2_{0,005}$	$\chi^2_{0,01}$	$\chi^2_{0,025}$	$\chi^2_{0,05}$	$\chi^2_{0,10}$	$\chi^2_{0,25}$	$\chi^2_{0,50}$	$\chi^2_{0,75}$
1	$0{,}0^4393$	$0{,}0^3157$	$0{,}0^3982$	$0{,}0^2393$	0,01579	0,1015	0,4549	1,3233
2	0,01003	0,02010	0,05064	0,10259	0,2107	0,5754	1,3863	2,7726
3	0,0717	0,1148	0,2158	0,3518	0,5844	1,2125	2,3660	4,10835
4	0,2070	0,2971	0,4844	0,7107	1,0636	1,92255	3,3567	5,3853
5	0,4117	0,5543	0,8312	1,1455	1,6103	2,6746	4,3515	6,6257
6	0,6757	0,8721	1,2373	1,6354	2,2041	3,4546	5,3481	7,8408
7	0,9893	1,2390	1,6899	2,16735	2,8331	4,25485	6,3458	9,0372
8	1,3444	1,6465	2,1797	2,7326	3,4895	5,0706	7,3441	10,2188
9	1,7349	2,0879	2,7004	3,3251	4,1682	5,8988	8,3428	11,3887
10	2,1559	2,5582	3,2470	3,9403	4,8652	6,7372	9,3418	12,5489
11	2,6032	3,0535	3,8158	4,5748	5,5778	7,5841	10,3410	13,7007
12	3,0738	3,5706	4,4038	5,2260	6,3038	8,4384	11,3403	14,8454
13	3,5650	4,1069	5,0087	5,8919	7,0415	9,2991	12,3398	15,9839
14	4,0747	4,6604	5,6287	6,5706	7,7895	10,1653	13,3393	17,1170
15	4,6009	5,2294	6,2621	7,2609	8,5468	11,0365	14,3389	18,2451
16	5,1422	5,8122	6,9077	7,9616	9,3122	11,9122	15,3385	19,3688
17	5,6972	6,4078	7,5642	8,6718	10,0852	12,7919	16,3381	20,4887
18	6,2648	7,0149	8,2308	9,3905	10,8649	13,6753	17,3379	21,6049
19	6,8440	7,6327	8,9066	10,1170	11,6509	14,5620	18,3376	22,7178
20	7,4339	8,2604	9,5908	10,8508	12,4426	15,4518	19,3374	23,8277
21	8,0337	8,8972	10,2829	11,5913	13,2396	16,3444	20,3372	24,9348
22	8,6427	9,5425	10,9823	12,3380	14,0415	17,2396	21,3370	26,0393
23	9,2604	10,1957	11,6885	13,0905	14,8479	18,1373	22,3369	27,1413
24	9,8862	10,8564	12,4001	13,8484	15,6587	19,0372	23,3367	28,2412
25	10,5197	11,5240	13,1197	14,6114	16,4734	19,9393	24,3366	29,3389
26	11,1603	12,1981	13,8439	15,3791	17,2919	20,8434	25,3364	30,4345
27	11,8076	12,8786	14,5733	16,1513	18,1138	21,7494	26,3363	31,5284
28	12,4613	13,5648	15,3079	16,9279	18,9392	22,6572	27,3363	32,6205
29	13,1211	14,2565	16,0471	17,7083	19,7677	23,5666	28,3362	33,7109
30	13,7867	14,9535	16,7908	18,4926	20,5992	24,4776	29,3360	34,7998
40	20,7065	22,1643	24,4331	26,5093	29,0505	33,6603	39,3354	45,6160
50	27,9907	29,7067	32,3574	34,7642	37,6886	42,9421	49,3349	56,3336
60	35,5346	37,4848	40,4817	43,1879	46,4589	52,2938	59,3347	66,9814
70	43,2752	45,4418	48,7576	51,7393	55,3290	61,6983	69,3344	77,5766
80	51,1720	53,5400	57,1532	60,3915	64,2778	71,1445	79,3343	88,1303
90	59,1963	61,7541	65,6466	69,1260	73,2912	80,6247	89,3342	98,6499
100	67,3276	70,0648	74,2219	77,9295	82,3581	90,1332	99,3341	109,141

Uit THOMPSON (1941) en GEIGY (1960).

* Zie voor normale benaderingen van χ^2-verdelingen de formules *(9.7)* en *(9.8)*.

TABEL D (*vervolg*)

ν	$\chi^2_{0,90}$	$\chi^2_{0,95}$	$\chi^2_{0,975}$	$\chi^2_{0,99}$	$\chi^2_{0,995}$	$\chi^2_{0,999}$	$\chi^2_{0,9995}$
1	2,7055	3,8415	5,0239	6,6349	7,8794	10,828	12,116
2	4,6052	5,9915	7,3778	9,2103	10,5966	13,816	15,202
3	6,2514	7,8147	9,3484	11,3449	12,8381	16,266	17,730
4	7,7794	9,4877	11,1433	13,2767	14,8602	18,467	19,998
5	9,23635	11,0705	12,8325	15,0863	16,7496	20,515	22,105
6	10,6446	12,5916	14,4494	16,8119	18,5476	22,458	24,103
7	12,0170	14,0671	16,0128	18,4753	20,2777	24,322	26,018
8	13,3616	15,5073	17,5346	20,0902	21,9550	26,125	27,868
9	14,6837	16,9190	19,0228	21,6660	23,5893	27,877	29,666
10	15,9871	18,3070	20,4831	23,2093	25,1882	29,588	31,419
11	17,2750	19,6751	21,9200	24,7250	26,7569	31,264	33,136
12	18,5494	21,0261	23,3367	26,2170	28,2995	32,909	34,821
13	19,8119	22,3621	24,7356	27,6883	29,8194	34,528	36,478
14	21,0642	23,6848	26,1190	29,1413	31,3193	36,123	38,109
15	22,3072	24,9958	27,4884	30,5779	32,8013	37,697	39,719
16	23,5418	26,2962	28,8454	31,9999	34,2672	39,252	41,308
17	24,7690	27,5871	30,1910	33,4087	35,7185	40,790	42,879
18	25,9894	28,8693	31,5264	34,8053	37,1564	42,312	44,434
19	27,2036	30,1435	32,8523	36,1908	38,5822	43,820	45,973
20	28,4120	31,4104	34,1696	37,5662	39,9968	45,315	47,498
21	29,6151	32,6705	35,4789	38,9321	41,4010	46,797	49,010
22	30,8133	33,9244	36,7807	40,2894	42,7956	48,268	50,511
23	32,0069	35,1725	38,0757	41,6384	44,1813	49,728	52,000
24	33,1963	36,4151	39,3641	42,9798	45,5585	51,179	53,479
25	34,3816	37,6525	40,6465	44,3141	46,9278	52,620	54,947
26	35,5631	38,8852	41,9232	45,6417	48,2899	54,052	56,407
27	36,7412	40,1133	43,1944	46,9630	49,6449	55,476	57,858
28	37,9159	41,3372	44,4607	48,2782	50,9933	56,892	59,300
29	39,0875	42,5569	45,7222	49,5879	52,3356	58,302	60,734
30	40,2560	43,7729	46,9792	50,8922	53,6720	59,703	62,161
40	51,8050	55,7585	59,3417	63,6907	66,7659	73,402	76,095
50	63,1671	67,5048	71,4202	76,1539	79,4900	86,661	89,561
60	74,3970	79,0819	83,2976	88,3794	91,9517	99,607	102,695
70	85,5271	90,5312	95,0231	100,425	104,215	112,317	115,577
80	96,5782	101,879	106,629	112,329	116,321	124,839	128,261
90	107,565	113,145	118,136	124,116	128,299	137,208	140,783
100	118,498	124,342	129,561	135,807	140,169	149,448	153,165

Voor $30 < \nu < 100$ kan lineair worden geïnterpoleerd. Men verkrijgt dan een uitkomst met 4, resp. 3, betrouwbare cijfers.

TABEL E. FRACTIELEN VAN DE VERDELING VAN DE CORRELATIECOËFFICIËNT r VOOR $\varrho = 0$

$n-2$	$r_{0,95}$	$r_{0,975}$	$r_{0,99}$	$r_{0,995}$	$r_{0,9995}$	$n-2$
1	0,988	0,997	0,9995	0,9999	1,000	1
2	0,900	0,950	0,980	0,990	0,999	2
3	0,805	0,878	0,934	0,959	0,991	3
4	0,729	0,811	0,882	0,917	0,974	4
5	0,669	0,754	0,833	0,874	0,951	5
6	0,622	0,707	0,789	0,834	0,925	6
7	0,582	0,666	0,750	0,798	0,898	7
8	0,549	0,632	0,716	0,765	0,872	8
9	0,521	0,602	0,685	0,735	0,847	9
10	0,497	0,576	0,658	0,708	0,823	10
11	0,476	0,553	0,634	0,684	0,801	11
12	0,458	0,532	0,612	0,661	0,780	12
13	0,441	0,514	0,592	0,641	0,760	13
14	0,426	0,497	0,574	0,623	0,742	14
15	0,412	0,482	0,558	0,606	0,725	15
16	0,400	0,468	0,542	0,590	0,708	16
17	0,389	0,456	0,528	0,575	0,693	17
18	0,378	0,444	0,516	0,561	0,679	18
19	0,369	0,433	0,503	0,549	0,665	19
20	0,360	0,423	0,492	0,537	0,652	20
22	0,344	0,404	0,472	0,515	0,629	22
24	0,330	0,388	0,453	0,496	0,607	24
25	0,323	0,381	0,445	0,487	0,597	25
30	0,296	0,349	0,409	0,449	0,554	30
35	0,275	0,325	0,381	0,418	0,519	35
40	0,257	0,304	0,358	0,393	0,490	40
45	0,243	0,288	0,338	0,372	0,465	45
50	0,231	0,273	0,322	0,354	0,443	50
55	0,220	0,261	0,307	0,338	0,424	55
60	0,211	0,250	0,295	0,325	0,408	60
65	0,203	0,240	0,284	0,312	0,393	65
70	0,195	0,232	0,274	0,302	0,380	70
75	0,189	0,224	0,264	0,292	0,368	75
80	0,183	0,217	0,256	0,283	0,357	80
85	0,178	0,211	0,249	0,275	0,347	85
90	0,173	0,205	0,242	0,267	0,338	90
95	0,168	0,200	0,236	0,260	0,329	95
100	0,164	0,195	0,230	0,254	0,321	100
125	0,147	0,174	0,206	0,228	0,288	125
150	0,134	0,159	0,189	0,208	0,264	150
175	0,124	0,148	0,174	0,194	0,248	175
200	0,116	0,138	0,164	0,181	0,235	200
300	0,095	0,113	0,134	0,148	0,188	300
500	0,074	0,088	0,104	0,115	0,148	500
1000	0,052	0,062	0,073	0,081	0,104	1000
2000	0,037	0,044	0,056	0,058	0,074	2000
	$-r_{0,05}$	$-r_{0,025}$	$-r_{0,01}$	$-r_{0,005}$	$-r_{0,0005}$	

Uit WALKER en LEV (1953).

Tabel F–1. Fractielen van F-verdelingen: $F_{0,95}$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

De tabellen F–1 t/m F–4 zijn ontleend aan Pearson & Hartley (1954). In deze tabellenverzameling vindt men ook $F_{0,90}$ en $F_{0,75}$.

TABEL F–2. FRACTIELEN VAN F-VERDELINGEN: $F_{0,975}$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	648	800	864	900	922	937	948	957	963	969	977	985	993	997	1001	1006	1010	1014	1018
2	38,5	39,0	39,2	39,2	39,3	39,3	39,4	39,4	39,4	39,4	39,4	39,4	39,4	39,5	39,5	39,5	39,5	39,5	39,5
3	17,4	16,0	15,4	15,1	14,9	14,7	14,6	14,5	14,5	14,4	14,3	14,3	14,2	14,1	14,1	14,0	14,0	13,9	13,9
4	12,2	10,6	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,75	8,66	8,56	8,51	8,46	8,41	8,36	8,31	8,26
5	10,0	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,52	6,43	6,33	6,28	6,23	6,18	6,12	6,07	6,02
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,37	5,27	5,17	5,12	5,07	5,01	4,96	4,90	4,85
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,67	4,57	4,47	4,42	4,36	4,31	4,25	4,20	4,14
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,20	4,10	4,00	3,95	3,89	3,84	3,78	3,73	3,67
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,87	3,77	3,67	3,61	3,56	3,51	3,45	3,39	3,33
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,62	3,52	3,42	3,37	3,31	3,26	3,20	3,14	3,08
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53	3,43	3,33	3,23	3,17	3,12	3,06	3,00	2,94	2,88
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,28	3,18	3,07	3,02	2,96	2,91	2,85	2,79	2,72
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25	3,15	3,05	2,95	2,89	2,84	2,78	2,72	2,66	2,60
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15	3,05	2,95	2,84	2,79	2,73	2,67	2,61	2,55	2,49
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	2,96	2,86	2,76	2,70	2,64	2,59	2,52	2,46	2,40
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	3,05	2,99	2,89	2,79	2,68	2,63	2,57	2,51	2,45	2,38	2,32
17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,98	2,92	2,82	2,72	2,62	2,56	2,50	2,44	2,38	2,32	2,25
18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,93	2,87	2,77	2,67	2,56	2,50	2,44	2,38	2,32	2,26	2,19
19	5,92	4,51	3,90	3,56	3,33	3,17	3,05	2,96	2,88	2,82	2,72	2,62	2,51	2,45	2,39	2,33	2,27	2,20	2,13
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,68	2,57	2,46	2,41	2,35	2,29	2,22	2,16	2,09
21	5,83	4,42	3,82	3,48	3,25	3,09	2,97	2,87	2,80	2,73	2,64	2,53	2,42	2,37	2,31	2,25	2,18	2,11	2,04
22	5,79	4,38	3,78	3,44	3,22	3,05	2,93	2,84	2,76	2,70	2,60	2,50	2,39	2,33	2,27	2,21	2,14	2,08	2,00
23	5,75	4,35	3,75	3,41	3,18	3,02	2,90	2,81	2,73	2,67	2,57	2,47	2,36	2,30	2,24	2,18	2,11	2,04	1,97
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,70	2,64	2,54	2,44	2,33	2,27	2,21	2,15	2,08	2,01	1,94
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61	2,51	2,41	2,30	2,24	2,18	2,12	2,05	1,98	1,91
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,57	2,51	2,41	2,31	2,20	2,14	2,07	2,01	1,94	1,87	1,79
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39	2,29	2,18	2,07	2,01	1,94	1,88	1,80	1,72	1,64
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,33	2,27	2,17	2,06	1,94	1,88	1,82	1,74	1,67	1,58	1,48
120	5,15	3,80	3,23	2,89	2,67	2,52	2,39	2,30	2,22	2,16	2,05	1,94	1,82	1,76	1,69	1,61	1,53	1,43	1,31
∞	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,11	2,05	1,94	1,83	1,71	1,64	1,57	1,48	1,39	1,27	1,00

Tabel F–3. Fractielen van F-verdelingen: $F_{0,99}$

		ν_1																		
		1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
ν_2	1	4052	5000	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
	2	98,5	99,0	99,2	99,2	99,3	99,3	99,4	99,4	99,4	99,4	99,4	99,4	99,4	99,5	99,5	99,5	99,5	99,5	99,5
	3	34,1	30,8	29,5	28,7	28,2	27,9	27,7	27,5	27,3	27,2	27,1	26,9	26,7	26,6	26,5	26,4	26,3	26,2	26,1
	4	21,2	18,0	16,7	16,0	15,5	15,2	15,0	14,8	14,7	14,5	14,4	14,2	14,0	13,9	13,8	13,7	13,7	13,6	13,5
	5	16,3	13,3	12,1	11,4	11,0	10,7	10,5	10,3	10,2	10,1	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
	6	13,7	10,9	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
	7	12,2	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,65
	8	11,3	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
	9	10,6	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31
	10	10,0	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
	11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
	12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45	3,36
	13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25	3,17
	14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3,00
	15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
	16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84	2,75
	17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75	2,65
	18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66	2,57
	19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,49
	20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
	21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46	2,36
	22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40	2,31
	23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35	2,26
	24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
	25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,62	2,54	2,45	2,36	2,27	2,17
	30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11	2,01
	40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92	1,80
	60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
	120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53	1,38
	∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00

Tabel F–4. Fractielen van F-verdelingen: $F_{0,995}$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24426	24630	24836	24940	25044	25148	25253	25359	25465
2	198	199	199	199	199	199	199	199	199	199	199	199	199	199	199	199	199	199,5	200
3	55,6	49,8	47,5	46,2	45,4	44,8	44,4	44,1	43,9	43,7	43,4	43,1	42,8	42,6	42,5	42,3	42,1	42,0	41,8
4	31,3	26,3	24,3	23,2	22,5	22,0	21,6	21,4	21,1	21,0	20,7	20,4	20,2	20,0	19,9	19,8	19,6	19,5	19,3
5	22,8	18,3	16,5	15,6	14,9	14,5	14,2	14,0	13,8	13,6	13,4	13,1	12,9	12,8	12,7	12,5	12,4	12,3	12,1
6	18,6	14,5	12,9	12,0	11,5	11,1	10,8	10,6	10,4	10,3	10,0	9,81	9,59	9,47	9,36	9,24	9,12	9,00	8,88
7	16,2	12,4	10,9	10,1	9,52	9,16	8,89	8,68	8,51	8,38	8,18	7,97	7,75	7,65	7,53	7,42	7,31	7,19	7,08
8	14,7	11,0	9,60	8,81	8,30	7,95	7,69	7,50	7,34	7,21	7,01	6,81	6,61	6,50	6,40	6,29	6,18	6,06	5,95
9	13,6	10,1	8,72	7,96	7,47	7,13	6,88	6,69	6,54	6,42	6,23	6,03	5,83	5,73	5,62	5,52	5,41	5,30	5,19
10	12,8	9,43	8,08	7,34	6,87	6,54	6,30	6,12	5,97	5,85	5,66	5,47	5,27	5,17	5,07	4,97	4,86	4,75	4,64
11	12,2	8,91	7,60	6,88	6,42	6,10	5,86	5,68	5,54	5,42	5,24	5,05	4,86	4,76	4,65	4,55	4,44	4,34	4,23
12	11,8	8,51	7,23	6,52	6,07	5,76	5,52	5,35	5,20	5,09	4,91	4,72	4,53	4,43	4,33	4,23	4,12	4,01	3,90
13	11,4	8,19	6,93	6,23	5,79	5,48	5,25	5,08	4,94	4,82	4,64	4,46	4,27	4,17	4,07	3,97	3,87	3,76	3,65
14	11,1	7,92	6,68	6,00	5,56	5,26	5,03	4,86	4,72	4,60	4,43	4,25	4,06	3,96	3,86	3,76	3,66	3,55	3,44
15	10,8	7,70	6,48	5,80	5,37	5,07	4,85	4,67	4,54	4,42	4,25	4,07	3,88	3,79	3,69	3,58	3,48	3,37	3,26
16	10,6	7,51	6,30	5,64	5,21	4,91	4,69	4,52	4,38	4,27	4,10	3,92	3,73	3,64	3,54	3,44	3,33	3,22	3,11
17	10,4	7,35	6,16	5,50	5,07	4,78	4,56	4,39	4,25	4,14	3,97	3,79	3,61	3,51	3,41	3,31	3,21	3,10	2,98
18	10,2	7,21	6,03	5,37	4,96	4,66	4,44	4,28	4,14	4,03	3,86	3,68	3,50	3,40	3,30	3,20	3,10	2,99	2,87
19	10,1	7,09	5,92	5,27	4,85	4,56	4,34	4,18	4,04	3,93	3,76	3,59	3,40	3,31	3,21	3,11	3,00	2,89	2,78
20	9,94	6,99	5,82	5,17	4,76	4,47	4,26	4,09	3,96	3,85	3,68	3,50	3,32	3,22	3,12	3,02	2,92	2,81	2,69
21	9,83	6,89	5,73	5,09	4,68	4,39	4,18	4,01	3,88	3,77	3,60	3,43	3,24	3,15	3,05	2,95	2,84	2,73	2,61
22	9,73	6,81	5,65	5,02	4,61	4,32	4,11	3,94	3,81	3,70	3,54	3,36	3,18	3,08	2,98	2,88	2,77	2,66	2,55
23	9,63	6,73	5,58	4,95	4,54	4,26	4,05	3,88	3,75	3,64	3,47	3,30	3,12	3,02	2,92	2,82	2,71	2,60	2,48
24	9,55	6,66	5,52	4,89	4,49	4,20	3,99	3,83	3,69	3,59	3,42	3,25	3,06	2,97	2,87	2,77	2,66	2,55	2,43
25	9,48	6,60	5,46	4,84	4,43	4,15	3,94	3,78	3,64	3,54	3,37	3,20	3,01	2,92	2,82	2,72	2,61	2,50	2,38
30	9,18	6,35	5,24	4,62	4,23	3,95	3,74	3,58	3,45	3,34	3,18	3,01	2,82	2,73	2,63	2,52	2,42	2,30	2,18
40	8,83	6,07	4,98	4,37	3,99	3,71	3,51	3,35	3,22	3,12	2,95	2,78	2,60	2,50	2,40	2,30	2,18	2,06	1,93
60	8,49	5,79	4,73	4,14	3,76	3,49	3,29	3,13	3,01	2,90	2,74	2,57	2,39	2,29	2,19	2,08	1,96	1,83	1,69
120	8,18	5,54	4,50	3,92	3,55	3,28	3,09	2,93	2,81	2,71	2,54	2,37	2,19	2,09	1,98	1,87	1,75	1,61	1,43
∞	7,88	5,30	4,28	3,72	3,35	3,09	2,90	2,74	2,62	2,52	2,36	2,19	2,00	1,90	1,79	1,67	1,53	1,36	1,00

Tabel G. De z-transformatie van Fisher:

$$z = \tfrac{1}{2} \log_e \frac{1+r}{1-r} = 1{,}1513\,[\log_{10}(1+r) - \log_{10}(1-r)]$$

r	0	1	2	3	4	5	6	7	8	9
0,00	0,000000	0,001000	0,002000	0,003000	0,004000	0,005000	0,006000	0,007000	0,008000	0,009000
0,01	0,010000	0,011000	0,012001	0,013001	0,014001	0,015001	0,016001	0,017002	0,018002	0,019002
0,02	0,020003	0,021003	0,022004	0,023004	0,024005	0,025005	0,026006	0,027007	0,028007	0,029008
0,03	0,030009	0,031010	0,032011	0,033012	0,034013	0,035014	0,036016	0,037017	0,038018	0,039020
0,04	0,040021	0,041023	0,042025	0,043027	0,044028	0,045030	0,046032	0,047035	0,048037	0,049039
0,05	0,050042	0,051044	0,052047	0,053050	0,054053	0,055056	0,056059	0,057062	0,058065	0,059069
0,06	0,060072	0,061076	0,062080	0,063084	0,064088	0,065092	0,066096	0,067101	0,068105	0,069110
0,07	0,070115	0,071120	0,072125	0,073130	0,074136	0,075141	0,076147	0,077153	0,078159	0,079165
0,08	0,080171	0,081178	0,082185	0,083192	0,084198	0,085206	0,086213	0,087220	0,088228	0,089236
0,09	0,090244	0,091253	0,092261	0,093270	0,094278	0,095287	0,096297	0,097306	0,098316	0,099325
0,10	0,100335	0,101346	0,102356	0,103367	0,104377	0,105389	0,106400	0,107411	0,108423	0,109435
0,11	0,110447	0,111459	0,112472	0,113485	0,114498	0,115511	0,116525	0,117538	0,118552	0,119567
0,12	0,120581	0,121596	0,122611	0,123626	0,124642	0,125657	0,126673	0,127690	0,128706	0,129723
0,13	0,130740	0,131757	0,132775	0,133793	0,134811	0,135829	0,136848	0,137867	0,138886	0,139906
0,14	0,140926	0,141946	0,142966	0,143987	0,145008	0,146029	0,147051	0,148073	0,149095	0,150118
0,15	0,151140	0,152164	0,153187	0,154211	0,155235	0,156260	0,157284	0,158309	0,159335	0,160361
0,16	0,161387	0,162413	0,163440	0,164467	0,165495	0,166522	0,167551	0,168579	0,169608	0,170637
0,17	0,171667	0,172697	0,173727	0,174758	0,175789	0,176820	0,177852	0,178884	0,179917	0,180949
0,18	0,181983	0,183016	0,184051	0,185085	0,186120	0,187155	0,188191	0,189227	0,190263	0,191300
0,19	0,192337	0,193375	0,194413	0,195451	0,196490	0,197530	0,198569	0,199610	0,200650	0,201691
0,20	0,202733	0,203774	0,204817	0,205860	0,206903	0,207946	0,208991	0,210035	0,211080	0,212126
0,21	0,213171	0,214218	0,215265	0,216312	0,217360	0,218408	0,219457	0,220506	0,221555	0,222606
0,22	0,223656	0,224707	0,225759	0,226811	0,227864	0,228917	0,229970	0,231024	0,232079	0,233134
0,23	0,234190	0,235246	0,236302	0,237359	0,238417	0,239475	0,240534	0,241593	0,242653	0,243713
0,24	0,244774	0,245836	0,246898	0,247960	0,249023	0,250087	0,251151	0,252215	0,253281	0,254347
0,25	0,255413	0,256480	0,257547	0,258616	0,259684	0,260753	0,261823	0,262894	0,263965	0,265036
0,26	0,266108	0,267181	0,268255	0,269329	0,270403	0,271479	0,272554	0,273631	0,274708	0,275786
0,27	0,276864	0,277943	0,279022	0,280103	0,281184	0,282265	0,283347	0,284430	0,285513	0,286597
0,28	0,287682	0,288768	0,289854	0,290940	0,292028	0,293116	0,294205	0,295294	0,296384	0,297475
0,29	0,298566	0,299659	0,300751	0,301845	0,302939	0,304034	0,305130	0,306226	0,307323	0,308421
0,30	0,309520	0,310625	0,311719	0,312820	0,313921	0,315023	0,316126	0,317230	0,318334	0,319440
0,31	0,320546	0,321652	0,322760	0,323868	0,324977	0,326087	0,327197	0,328309	0,329421	0,330534
0,32	0,331647	0,332762	0,333877	0,334993	0,336110	0,337228	0,338346	0,339465	0,340586	0,341707
0,33	0,342828	0,343951	0,345074	0,346199	0,347324	0,348450	0,349577	0,350704	0,351833	0,352962
0,34	0,354093	0,355224	0,356356	0,357489	0,358623	0,359757	0,360893	0,362029	0,363167	0,364305
0,35	0,365444	0,366584	0,367725	0,368867	0,370010	0,371153	0,372298	0,373444	0,374590	0,375738
0,36	0,376886	0,378035	0,379186	0,380337	0,381489	0,382643	0,383797	0,384952	0,386108	0,387265
0,37	0,388423	0,389582	0,390743	0,391904	0,393066	0,394229	0,395393	0,396556	0,397724	0,398892
0,38	0,400060	0,401229	0,402399	0,403571	0,404743	0,405917	0,407091	0,408267	0,409444	0,410621
0,39	0,411800	0,412980	0,414161	0,415343	0,416527	0,417711	0,418896	0,420083	0,421270	0,422459
0,40	0,423649	0,424840	0,426032	0,427226	0,428420	0,429616	0,430813	0,432010	0,433210	0,434410
0,41	0,435611	0,436814	0,438018	0,439223	0,440429	0,441637	0,442845	0,444055	0,445266	0,446479
0,42	0,447692	0,448907	0,450123	0,451340	0,452559	0,453779	0,455000	0,456222	0,457446	0,458671
0,43	0,459897	0,461124	0,462353	0,463583	0,464815	0,466047	0,467281	0,468517	0,469754	0,470992
0,44	0,472231	0,473472	0,474714	0,475957	0,477202	0,478448	0,479696	0,480945	0,482195	0,483447
0,45	0,484700	0,485955	0,487211	0,488469	0,489728	0,490988	0,492250	0,493513	0,494778	0,496044
0,46	0,497311	0,498581	0,499851	0,501123	0,502397	0,503672	0,504949	0,506227	0,507507	0,508788
0,47	0,510070	0,511355	0,512641	0,513928	0,515217	0,516508	0,517800	0,519094	0,520389	0,521686
0,48	0,522985	0,524285	0,525586	0,526890	0,528195	0,529502	0,530810	0,532120	0,533432	0,534745
0,49	0,536061	0,537377	0,538696	0,540016	0,541338	0,542662	0,543987	0,545314	0,546643	0,547974
0,50	0,549306	0,550641	0,551977	0,553314	0,554654	0,555995	0,557339	0,558684	0,560031	0,561379

vervolg op blz. 362

r	0	1	2	3	4	5	6	7	8	9
0,51	0,562730	0,564082	0,565437	0,566793	0,568151	0,569511	0,570873	0,572237	0,573603	0,574970
0,52	0,576340	0,577712	0,579085	0,580461	0,581838	0,583218	0,584599	0,585983	0,587368	0,588756
0,53	0,590145	0,591537	0,592931	0,594327	0,595724	0,597124	0,598526	0,599931	0,601337	0,602745
0,54	0,604156	0,605568	0,606983	0,608400	0,609820	0,611241	0,612665	0,614091	0,615519	0,616949
0,55	0,618382	0,619816	0,621253	0,622693	0,624134	0,625579	0,627025	0,628473	0,629924	0,631378
0,56	0,632833	0,634292	0,635752	0,637215	0,638680	0,640148	0,641618	0,643091	0,644566	0,646043
0,57	0,647523	0,649006	0,650491	0,651978	0,653468	0,654961	0,656456	0,657954	0,659454	0,660957
0,58	0,662463	0,663971	0,665482	0,666996	0,668512	0,670031	0,671553	0,673077	0,674604	0,676134
0,59	0,677666	0,679202	0,680740	0,682281	0,683825	0,685371	0,686921	0,688473	0,690028	0,691586
0,60	0,693147	0,694711	0,696278	0,697848	0,699421	0,700997	0,702576	0,704158	0,705742	0,707330
0,61	0,708922	0,710516	0,712113	0,713713	0,715317	0,716924	0,718534	0,720147	0,721763	0,723382
0,62	0,725005	0,726631	0,728261	0,729893	0,731529	0,733169	0,734811	0,736458	0,738107	0,739760
0,63	0,741416	0,743076	0,744740	0,746406	0,748077	0,749751	0,751428	0,753109	0,754794	0,756482
0,64	0,758174	0,759870	0,761569	0,763272	0,764979	0,766689	0,768403	0,770122	0,771844	0,773569
0,65	0,775299	0,777033	0,778770	0,780511	0,782257	0,784006	0,785760	0,787517	0,789279	0,791044
0,66	0,792814	0,794588	0,796366	0,798148	0,799934	0,801725	0,803520	0,805320	0,807123	0,808931
0,67	0,810744	0,812560	0,814381	0,816207	0,818037	0,819872	0,821711	0,823555	0,825404	0,827257
0,68	0,829114	0,830977	0,832844	0,834716	0,836593	0,838474	0,840361	0,842252	0,844149	0,846050
0,69	0,847956	0,849867	0,851784	0,853705	0,855632	0,857564	0,859500	0,861443	0,863390	0,865343
0,70	0,867300	0,869264	0,871233	0,873208	0,875187	0,877173	0,879164	0,881160	0,883163	0,885171
0,71	0,887184	0,889204	0,891229	0,893260	0,895297	0,897340	0,899389	0,901444	0,903505	0,905572
0,72	0,907645	0,909725	0,911811	0,913903	0,916001	0,918106	0,920217	0,922335	0,924459	0,926590
0,73	0,928728	0 930872	0,933023	0,935181	0,937345	0,939517	0,941695	0,943881	0,946073	0,948273
0,74	0,950480	0,952694	0,954915	0,957144	0,959380	0,961623	0,963875	0,966133	0,968400	0,970674
0,75	0,972955	0,975245	0,977543	0,979848	0,982162	0,984483	0,986813	0,989151	0,991498	0,993852
0,76	0,996216	0,998587	1,000967	1,003356	1,005754	1,008161	1,010576	1,013000	1,015434	1,017876
0,77	1,020328	1,022789	1,025260	1,027739	1,030229	1,032728	1,035237	1,037755	1,040284	1,042822
0,78	1,045371	1,047930	1,050499	1,053078	1,055668	1,058268	1,060879	1,063501	1,066134	1,068777
0,79	1,071432	1,074098	1,076775	1,079579	1,082164	1,084876	1,087599	1,090334	1,093082	1,095841
0,80	1,098613	1,101397	1,104193	1,107002	1,109824	1,112659	1,115506	1,118367	1,121241	1,124128
0,81	1,127030	1,129944	1,132873	1,135815	1,138772	1,141743	1,144728	1,147728	1,150743	1,153773
0,82	1,156818	1,159878	1,162954	1,166045	1,169152	1,172275	1,175414	1,178570	1,181742	1,184931
0,83	1,188137	1,191360	1,194600	1,197858	1,201134	1,204428	1,207740	1,211070	1,214419	1,217787
0,84	1,221174	1,224581	1,228007	1,231452	1,234919	1,238405	1,241912	1,245441	1,248990	1,252561
0,85	1,256153	1,259768	1,263405	1,267065	1,270747	1,274454	1,278183	1,281937	1,285715	1,289518
0,86	1,293345	1,297198	1,301077	1,304982	1,308913	1,312871	1,316857	1,320870	1,324911	1,328981
0,87	1,333080	1,337209	1,341367	1,345555	1,349775	1,354026	1,358308	1,362623	1,366971	1,371353
0,88	1,375768	1,380218	1,384703	1,389224	1,393782	1,398376	1,403008	1,407678	1,412388	1,417137
0,89	1,421926	1,426757	1,431630	1,436545	1,441504	1,446507	1,451556	1,456651	1,461792	1,466982
0,90	1,472220	1,477508	1,482847	1,488239	1,493683	1,499181	1,504734	1,510344	1,516012	1,521738
0,91	1,527525	1,533373	1,539285	1,545261	1,551302	1,557411	1,563590	1,569839	1,576160	1,582556
0,92	1,589027	1,595577	1,602207	1,608913	1,615715	1,622597	1,629568	1,636631	1,643787	1,651039
0,93	1,658391	1,665844	1,673402	1,681069	1,688846	1,696738	1,704749	1,712881	1,721139	1,729528
0,94	1,738050	1,746711	1,755516	1,764469	1,773577	1,782843	1,792274	1,801877	1,811658	1,821624
0,95	1,831782	1,842139	1,852705	1,863488	1,874497	1,885742	1,897234	1,908985	1,921005	1,933309
0,96	1,945911	1,958825	1,972068	1,985656	1,999611	2,013951	2,028699	2,043880	2,059519	2,075648
0,97	2,092296	2,109501	2,127300	2,145738	2,164861	2,184725	2,205389	2,226922	2,249400	2,272913
0,98	2,297561	2,323460	2,350746	2,379577	2,410142	2,442663	2,477411	2,514717	2,554990	2,598747
0,99	2,646653	2,699585	2,758727	2,825744	2,903070	2,994482	3,106304	3,250396	3,453379	3,800203
1,00	∞									

Tabel H. De toets van Hartley

Rechter kritieke waarden van de grootheid $F_{max} = \frac{s^2_{max}}{s^2_{min}}$

$\alpha = 0{,}05$

$n-1$ \ k	2	3	4	5	6	7	8	9	10	11	12
2	39,0	87,5	142	202	266	333	403	475	550	626	704
3	15,4	27,8	39,2	50,7	62,0	72,9	83,5	93,9	104	114	124
4	9,60	15,5	20,6	25,2	29,5	33,6	37,5	41,1	44,6	48,0	51,4
5	7,15	10,8	13,7	16,3	18,7	20,8	22,9	24,7	26,5	28,2	29,9
6	5,82	8,38	10,4	12,1	13,7	15,0	16,3	17,5	18,6	19,7	20,7
7	4,99	6,94	8,44	9,70	10,8	11,8	12,7	13,5	14,3	15,1	15,8
8	4,43	6,00	7,18	8,12	9,03	9,78	10,5	11,1	11,7	12,2	12,7
9	4,03	5,34	6,31	7,11	7,80	8,41	8,95	9,45	9,91	10,3	10,7
10	3,72	4,85	5,67	6,34	6,92	7,42	7,87	8,28	8,66	9,01	9,34
12	3,28	4,16	4,79	5,30	5,72	6,09	6,42	6,72	7,00	7,25	7,48
15	2,86	3,54	4,01	4,37	4,68	4,95	5,19	5,40	5,59	5,77	5,93
20	2,46	2,95	3,29	3,54	3,76	3,94	4,10	4,24	4,37	4,49	4,59
30	2,07	2,40	2,61	2,78	2,91	3,02	3,12	3,21	3,29	3,36	3,39
60	1,67	1,85	1,96	2,04	2,11	2,17	2,22	2,26	2,30	2,33	2,36
∞	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00

$\alpha = 0{,}01$

$n-1$ \ k	2	3	4	5	6	7	8	9	10	11	12
2	199	448	729	1036	1362	1705	2063	2432	2813	3204	3605
3	47,5	85	120	151	184	216*	249*	281*	310*	337*	361*
4	23,2	37	49	59	69	79	89	97	106	113	120
5	14,9	22	28	33	38	42	46	50	54	57	60
6	11,1	15,5	19,1	22	25	27	30	32	34	36	37
7	8,89	12,1	14,5	16,5	18,4	20	22	23	24	26	27
8	7,50	9,9	11,7	13,2	14,5	15,8	16,9	17,9	18,9	19,8	21
9	6,54	8,5	9,9	11,1	12,1	13,1	13,9	14,7	15,3	16,0	16,6
10	5,85	7,4	8,6	9,6	10,4	11,1	11,8	12,4	12,9	13,4	13,9
12	4,91	6,1	6,9	7,6	8,2	8,7	9,1	9,5	9,9	10,2	10,6
15	4,07	4,9	5,5	6,0	6,4	6,7	7,1	7,3	7,5	7,8	8,0
20	3,32	3,8	4,3	4,6	4,9	5,1	5,3	5,5	5,6	5,8	5,9
30	2,63	3,0	3,3	3,4	3,6	3,7	3,8	3,9	4,0	4,1	4,2
60	1,96	2,2	2,3	2,4	2,4	2,5	2,5	2,6	2,6	2,7	2,7
∞	1,00	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0

Overgenomen uit Pearson en Hartley (1954), table 31.

s^2_{max} is the largest and s^2_{min} is the smallest in a set of k independent variances, each based on $n-1$ degrees of freedom.
Values in the column $k = 2$ and in the rows $n-1 = 2$ and ∞ are exact. Elsewhere the third digit may be in error by a few units for the 5% points and several units for the 1% points. The third digit figures marked * for $n-1 = 3$ are the most uncertain.

eenzijdig α=0,01 | eenzijdig α=0,05

TABEL I. DE TEKENTOETS

Linker kritieke waarden van x bij tweezijdige toetsing ($x_R = n - x_L$)

n	Onbetrouwbaarheidsdrempel α				n	Onbetrouwbaarheidsdrempel α			
	0,01	0,02	0,05	0,10		0,01	0,02	0,05	0,10
1	—	—	—	—	51	15	16	18	19
2	—	—	—	—	52	16	17	18	19
3	—	—	—	—	53	16	17	18	20
4	—	—	—	—	54	17	18	19	20
5	—	—	—	0	55	17	18	19	20
6	—	—	0	0	56	17	18	20	21
7	—	0	0	0	57	18	19	20	21
8	0	0	0	1	58	18	19	21	22
9	0	0	1	1	59	19	20	21	22
10	0	0	1	1	60	19	20	21	23
11	0	1	1	2	61	20	20	22	23
12	1	1	2	2	62	20	21	22	24
13	1	1	2	3	63	20	21	23	24
14	1	2	2	3	64	21	22	23	24
15	2	2	3	3	65	21	22	24	25
16	2	2	3	4	66	22	23	24	25
17	2	3	4	4	67	22	23	25	26
18	3	3	4	5	68	22	23	25	26
19	3	4	4	5	69	23	24	25	27
20	3	4	5	5	70	23	24	26	27
21	4	4	5	6	71	24	25	26	28
22	4	5	5	6	72	24	25	27	28
23	4	5	6	7	73	25	26	27	28
24	5	5	6	7	74	25	26	28	29
25	5	6	7	7	75	25	26	28	29
26	6	6	7	8	76	26	27	28	30
27	6	7	7	8	77	26	27	29	30
28	6	7	8	9	78	27	28	29	31
29	7	7	8	9	79	27	28	30	31
30	7	8	9	10	80	28	29	30	32
31	7	8	9	10	81	28	29	31	32
32	8	8	9	10	82	28	30	31	33
33	8	9	10	11	83	29	30	32	33
34	9	9	10	11	84	29	30	32	33
35	9	10	11	12	85	30	31	32	34
36	9	10	11	12	86	30	31	33	34
37	10	10	12	13	87	31	32	33	35
38	10	11	12	13	88	31	32	34	35
39	11	11	12	13	89	31	33	34	36
40	11	12	13	14	90	32	33	35	36
41	11	12	13	14	91	32	33	35	37
42	12	13	14	15	92	33	34	36	37
43	12	13	14	15	93	33	34	36	38
44	13	13	15	16	94	34	35	37	38
45	13	14	15	16	95	34	35	37	38
46	13	14	15	16	96	34	36	37	39
47	14	15	16	17	97	35	36	38	39
48	14	15	16	17	98	35	37	38	40
49	15	15	17	18	99	36	37	39	40
50	15	16	17	18	100	36	37	39	41

TABEL J. DE RANG-TEKENTOETS

Rechter kritieke waarden van $\underline{V}$ bij tweezijdige toetsing ($V_L = -V_R$)

n	Onbetrouwbaarheidsdrempel α				n	Onbetrouwbaarheidsdrempel α			
	0,01	0,02	0,05	0,10		0,01	0,02	0,05	0,10
1	—	—	—	—	51	552	498	420	352
2	—	—	—	—	52	568	512	432	364
3	—	—	—	—	53	583	527	445	373
4	—	—	—	—	54	601	543	457	385
5	—	—	—	15	55	616	558	470	394
6	—	—	21	17	56	634	572	482	406
7	—	28	24	22	57	651	587	495	417
8	36	34	30	26	58	667	603	509	427
9	43	39	35	29	59	684	618	522	438
10	49	45	39	35	60	702	634	534	448
11	56	52	46	40	61	719	649	547	461
12	64	60	52	44	62	737	665	561	471
13	73	67	57	49	63	754	682	574	482
14	81	75	63	55	64	772	698	588	494
15	90	82	70	60	65	791	713	601	505
16	98	90	78	66	66	809	731	615	517
17	107	99	85	71	67	826	746	630	528
18	117	107	91	77	68	846	764	644	540
19	126	116	98	84	69	863	781	657	553
20	136	124	106	90	70	883	797	671	565
21	147	133	115	97	71	902	814	686	576
22	157	143	123	103	72	920	830	700	588
23	168	152	130	110	73	939	849	715	601
24	178	162	138	118	74	959	865	729	613
25	189	173	147	125	75	978	882	744	626
26	201	183	155	131	76	998	900	760	638
27	212	194	164	140	77	1017	919	775	649
28	224	204	174	146	78	1037	935	789	663
29	235	215	183	155	79	1056	954	804	676
30	247	225	191	163	80	1076	972	820	688
31	260	236	202	170	81	1097	991	835	701
32	272	248	210	178	82	1117	1009	849	713
33	285	259	221	187	83	1136	1026	866	726
34	299	271	231	195	84	1158	1046	880	740
35	312	284	240	204	85	1177	1063	897	753
36	324	296	250	212	86	1199	1083	913	767
37	339	307	261	221	87	1220	1102	928	780
38	353	319	271	229	88	1240	1120	944	792
39	366	332	282	238	89	1261	1139	961	807
40	380	344	292	248	90	1283	1159	977	819
41	395	357	303	257	91	1304	1178	992	834
42	409	371	315	265	92	1326	1196	1008	846
43	424	384	326	274	93	1347	1217	1025	861
44	438	398	336	284	94	1369	1235	1041	875
45	453	411	349	293	95	1390	1256	1058	888
46	467	425	359	303	96	1412	1276	1074	902
47	484	438	372	314	97	1433	1295	1091	917
48	498	452	384	324	98	1455	1315	1109	931
49	515	467	395	333	99	1478	1334	1126	944
50	529	481	407	343	100	1500	1354	1142	958

n = 5 t/m 50 exact berekend, 51 t/m 100 vlgs. de normale benadering.
Uit 'Statistische Tabellen en Nomogrammen'.

TABEL K–1. DE TOETS VAN FRIEDMAN (METHODE DER m RANGSCHIKKINGEN)
Exacte rechtse overschrijdingskansen van K voor kleine waarden van k en n

$k = 3$, $n = 3$	
$K = 18$	0,028

$k = 3$, $n = 4$	
$K = 18$	0,125
24	0,069
26	0,042
32	0,0046

$k = 3$, $n = 5$	
$K = 18$	0,182
24	0,124
26	0,093
32	0,039
38	0,024
42	0,0085
50	$0,0^{3}77$

$k = 3$, $n = 6$	
$K = 26$	0,142
32	0,072
38	0,052
42	0,029
50	0,012
54	0,0081
56	0,0055
62	0,0017
72	$0,0^{3}13$

$k = 3$, $n = 7$	
$K = 26$	0,192
32	0,112
38	0,085
42	0,051
50	0,027
54	0,021
56	0,016
62	0,0084
72	0,0036
74	0,0027
78	0,0012
86	$0,0^{3}32$
98	$0,0^{4}21$

$k = 3$, $n = 8$	
$K = 32$	0,149
38	0,120
42	0,079
50	0,047
54	0,038
56	0,030
62	0,018
72	0,0099
74	0,0080

$k = 3$, $n = 8$ *(vervolg)*	
$K = 78$	0,0048
86	0,0024
96	0,0011
98	$0,0^{3}86$
104	$0,0^{3}26$
114	$0,0^{4}61$
128	$0,0^{3}36$

$k = 3$, $n = 9$	
$K = 32$	0,187
38	0,154
42	0,107
50	0,069
54	0,057
56	0,048
62	0,031
72	0,019
74	0,016
78	0,010
86	0,0060
96	0,0035
98	0,0029
104	0,0013
114	$0,0^{3}66$
122	$0,0^{3}35$
126	$0,0^{3}20$
128	$0,0^{4}97$
134	$0,0^{4}54$
146	$0,0^{4}11$
162	$0,0^{5}60$

$k = 3$, $n = 10$	
$K = 38$	0,187
42	0,135
50	0,092
54	0,078
56	0,066
62	0,046
72	0,030
74	0,026
78	0,018
86	0,012
96	0,0075
98	0,0063
104	0,0034
114	0,0020
122	0,0013
126	$0,0^{3}83$
128	$0,0^{3}51$
134	$0,0^{3}37$
146	$0,0^{3}18$

$k = 3$, $n = 10$ *(vervolg)*	
$K = 150$	$0,0^{3}11$
152	$0,0^{4}85$
158	$0,0^{4}44$
162	$0,0^{4}20$
168	$0,0^{4}11$
182	$0,0^{5}21$
200	$0,0^{7}99$

$k = 4$, $n = 2$	
$K = 18$	0,167
20	0,042

$k = 4$, $n = 3$	
$K = 27$	0,175
29	0,148
33	0,075
35	0,054
37	0,033
41	0,017
45	0,0017

$k = 4$, $n = 4$	
$K = 34$	0,190
36	0,158
38	0,141
40	0,105
42	0,094
44	0,077
46	0,068
48	0,054
50	0,052
52	0,036
54	0,033
56	0,019
58	0,014
62	0,012
64	0,0069
66	0,0062
68	0,0027
72	0,0016
74	$0,0^{3}94$
80	$0,0^{4}72$

$k = 4$, $n = 5$	
$K = 43$	0,162
45	0,141
49	0,123
51	0,107
53	0,093
57	0,075
59	0,067

$k = 4$, $n = 5$ *(vervolg)*	
$K = 61$	0,055
65	0,044
67	0,034
69	0,031
73	0,023
75	0,020
77	0,017
81	0,012
83	0,0087
85	0,0067
89	0,0055
91	0,0031
93	0,0023
97	0,0018
99	0,0016
101	0,0014
105	$0,0^{3}64$
107	$0,0^{3}33$
109	$0,0^{3}21$
113	$0,0^{3}14$
117	$0,0^{4}48$
125	$0,0^{5}30$

$k = 4$, $n = 6$	
$K = 52$	0,163
54	0,155
56	0,127
58	0,114
62	0,108
64	0,089
66	0,088
68	0,073
70	0,066
72	0,060
74	0,056
76	0,043
78	0,041
80	0,037
82	0,035
84	0,032
86	0,029
88	0,023
90	0,022
94	0,017
96	0,014
98	0,013
100	0,010
102	0,0096
104	0,0085
106	0,0073
108	0,0061

$k = 4$, $n = 6$ *(vervolg)*	
$K = 110$	0,0057
114	0,0040
116	0,0033
118	0,0028
120	0,0023
122	0,0020
126	0,0015
128	$0,0^{3}90$
130	$0,0^{3}87$
132	$0,0^{3}73$
134	$0,0^{3}65$
136	$0,0^{3}40$
138	$0,0^{3}36$
140	$0,0^{3}28$
144	$0,0^{3}24$
146	$0,0^{3}22$
148	$0,0^{3}12$
150	$0,0^{4}95$
152	$0,0^{4}62$
154	$0,0^{4}46$
158	$0,0^{4}24$
160	$0,0^{4}16$
162	$0,0^{4}12$
164	$0,0^{5}80$
170	$0,0^{5}24$
180	$0,0^{6}13$

$k = 5$, $n = 3$	
$K = 50$	0,163
52	0,127
54	0,117
56	0,096
58	0,080
60	0,063
62	0,056
64	0,045
66	0,038
68	0,028
70	0,026
72	0,017
74	0,015
76	0,0078
78	0,0053
80	0,0040
82	0,0028
86	$0,0^{3}90$
90	$0,0^{4}69$

Ontleend aan KENDALL (1948), Tabel 5, Appendix; $0,0^{3}11 = 0,00011$.

TABEL K–2. DE TOETS VAN FRIEDMAN

Rechter kritieke waarden van $\underline{K}$ bij $\alpha = 0{,}05$

k \ n	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	—	—	—	—	*18*	*24,5*	*32*	*24,5*	*32*	*40,5*	*32*	*40,5*	*50*	*40,5*
3	—	*18*	*26*	*32*	*42*	*50*	*50*	*56*	*62*	66	72	78	84	90
4	*20*	*37*	*52*	*65*	*76*	92	105	118	131	144	157	170	183	196
5	*38*	*64*	89	113	137	167	190	214	238	261	285	309	333	356
6	*64*	104	144	183	222	272	310	349	388	427	465	504	543	582
7	*96*	158	217	277	336	412	471	529	588	647	706	764	823	882
8	*138*	225	311	396	482	591	676	760	845	929	1013	1098	1182	1267
9	*192*	311	429	547	664	815	931	1047	1164	1280	1396	1512	1629	1745
10	*258*	416	574	731	887	1086	1241	1396	1551	1706	1862	2017	2172	2327
11	336	542	747	950	1155	1410	1612	1813	2014	2216	2417	2618	2820	3021
12	429	691	950	1210	1469	1791	2047	2302	2558	2814	3070	3326	3581	3837
13	538	865	1189	1512	1831	2233	2552	2871	3189	3508	3827	4146	4465	4784
14	664	1063	1460	1859	2253	2740	3131	3523	3914	4305	4697	5088	5479	5871
15	808	1292	1770	2254	2738	3316	3790	4264	4737	5211	5685	6159	6632	7106

Uit ENTERS (1957). De cursief gedrukte waarden berusten op exacte berekening. De overige waarden zijn volgens verschillende benaderingsmethoden berekend door de Statistische Afdeling van het Mathematisch Centrum te Amsterdam.

Opmerking betreffende tabellen met kritieke waarden

Een streepje(–) in deze tabellen betekent, dat bij de betrokken drempelwaarde α geen significante waarde van de toetsingsgrootheid aanwezig is.

Voor alle tabellen met kritieke waarden bij tweezijdige toetsing geldt, dat zij voor éénzijdige toetsing met een drempelwaarde α kunnen worden gebruikt, door de kritieke waarde in de kolom met 2α af te lezen.

TABEL L–1. DE TOETS VAN WILCOXON

Linker en rechter kritieke waarden van S bij tweezijdige toetsing

$\alpha = 0{,}10$

n_2 \ n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
4	11– 25	17– 33	24– 42	32– 52	41– 63	51– 75	62– 88	74–102	87–117	101–133	116–150	132–168	150–186	168–206	187–227	207–249	228–272	250–296	273–321	297–347	322–374	348–402
5	12– 28	19– 36	26– 46	34– 57	44– 68	54– 81	66– 94	78–109	91–125	106–141	121–159	138–177	155–197	173–218	193–239	213–262	235–285	257–310	281–335	305–362	330–390	357–418
6	13– 31	20– 40	28– 50	36– 62	46– 74	57– 87	69–101	82–116	95–133	110–150	126–168	143–187	161–207	179–229	199–251	220–274	242–298	265–323	289–349	313–377	339–405	366–434
7	14– 34	21– 44	29– 55	39– 66	49– 79	60– 93	72–108	85–124	99–141	115–158	131–177	148–197	166–218	186–239	206–262	227–286	249–311	272–337	297–363	322–391	348–420	375–450
8	15– 37	23– 47	31– 59	41– 71	51– 85	63– 99	75–115	89–131	104–148	119–167	136–186	153–207	172–228	192–250	212–274	234–298	257–323	280–350	305–377	330–406	357–435	385–465
9	16– 40	24– 51	33– 63	43– 76	54– 90	66–105	79–121	93–138	108–156	124–178	141–195	159–216	178–238	198–261	219–285	241–310	264–336	288–363	313–391	339–420	366–450	394–481
10	17– 43	26– 54	35– 67	45– 81	56– 96	69–111	82–128	97–145	112–164	128–184	146–204	164–226	184–248	204–272	226–297	248–322	272–348	296–376	321–405	348–434	375–465	403–497
11	18– 46	27– 58	37– 71	47– 86	59–101	72–117	86–134	100–153	116–172	133–192	151–213	170–235	190–258	210–283	232–308	255–334	279–361	304–389	329–419	356–449	384–480	413–512
12	19– 49	28– 62	38– 76	49– 91	62–106	75–123	89–141	104–160	120–180	138–200	156–222	175–245	196–268	217–293	239–319	262–346	286–374	312–402	338–432	365–463	393–495	423–527
13	20– 52	30– 65	40– 80	52– 95	64–112	78–129	92–148	108–167	125–187	142–209	161–231	181–254	201–279	223–304	245–331	269–358	294–386	320–415	346–446	374–477	403–509	433–542
14	21– 55	31– 69	42– 84	54–100	67–117	81–135	96–154	112–174	129–195	147–217	166–240	186–264	207–289	229–315	252–342	276–370	301–399	328–428	355–459	383–491	412–524	442–558
15	22– 58	33– 72	44– 88	56–105	69–123	84–141	99–161	116–181	133–203	152–225	171–249	191–274	213–299	235–326	259–353	284–381	309–411	336–441	363–473	392–505	422–538	452–573
16	24– 60	34– 76	46– 92	58–110	72–128	87–147	103–167	120–188	138–210	156–234	176–258	197–283	219–309	242–336	266–364	291–393	317–423	344–454	372–486	401–519	431–553	462–588
17	25– 63	35– 80	47– 97	61–114	75–133	90–153	106–174	123–196	142–218	161–242	181–267	202–293	225–319	248–347	273–375	298–405	325–435	352–467	380–500	410–533	440–568	472–603
18	26– 66	37– 83	49–101	63–119	77–139	93–159	110–180	127–203	146–226	165–251	186–276	208–302	231–329	255–357	280–386	305–417	332–448	360–480	389–513	419–547	450–582	482–618
19	27– 69	38– 87	51–105	66–124	80–144	96–165	113–187	131–210	150–234	170–259	191–285	214–311	237–339	261–368	286–398	313–428	340–460	368–493	398–526	428–561	459–597	492–633
20	28– 72	40– 90	53–109	67–129	83–149	99–171	117–193	135–217	154–242	175–267	196–294	219–321	243–349	268–378	293–409	320–440	348–472	376–506	406–540	437–575	469–611	501–649
21	29– 75	41– 94	55–113	69–134	86–155	102–177	120–200	139–224	159–249	180–275	202–302	225–330	249–359	274–389	300–420	327–452	355–485	385–518	415–553	446–589	478–626	511–664
22	30– 78	43– 97	57–117	72–138	88–160	105–183	123–207	142–232	163–257	184–284	207–311	230–340	255–369	280–400	307–431	335–463	363–497	393–531	423–567	455–603	488–640	521–679
23	31– 81	44–101	58–122	74–143	90–166	108–189	127–213	146–239	167–265	189–292	212–320	236–349	261–379	287–410	314–442	342–475	371–509	401–544	432–580	464–617	497–655	531–694
24	32– 84	45–105	60–126	76–148	93–171	111–195	130–220	150–246	171–273	194–300	217–329	242–358	267–389	293–421	321–453	349–487	379–521	409–557	441–593	473–631	507–669	541–709
25	33– 87	47–108	62–130	78–153	96–176	114–201	133–227	154–253	176–280	199–308	222–338	247–368	273–399	300–431	328–464	357–498	386–534	417–570	449–607	482–645	516–684	551–724

De cursieve kritieke waarden zijn exact [zij zijn berekend uit de kritieke waarden van W in WABEKE en VAN EEDEN (1955)]. De overige kritieke waarden zijn benaderend en ontleend aan GEIGY (1960).

Tabel L–2. De toets van Wilcoxon

Linker en rechter kritieke waarden van S bij tweezijdige toetsing

$\alpha = 0{,}05$

n_2 \ n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
4	10– 26	16– 34	23– 43	31– 53	40– 64	49– 77	60– 90	72–104	85–119	99–135	114–152	130–170	147–189	164–210	183–231	203–253	224–276	246–300	269–325	293–351	317–379	343–407
5	11– 29	17– 38	24– 48	33– 58	42– 70	52– 83	63– 97	75–112	89–127	103–144	118–162	134–181	151–201	170–221	189–243	209–266	230–290	253–314	276–340	300–367	325–395	352–423
6	12– 32	18– 42	26– 52	34– 64	44– 76	55– 89	66–104	79–119	92–136	107–153	122–172	139–191	157–211	175–233	195–255	215–279	237–303	260–328	283–355	308–382	333–411	360–440
7	13– 35	20– 45	27– 57	36– 69	46– 82	57– 96	69–111	82–127	96–144	111–162	127–181	144–201	162–222	181–244	201–267	222–291	244–316	267–342	291–369	316–397	342–426	369–456
8	14– 38	21– 49	29– 61	38– 74	49– 87	60–102	72–118	85–135	100–152	115–171	131–191	149–211	167–233	187–255	207–279	228–304	251–329	274–356	298–384	324–412	350–442	378–472
9	14– 42	22– 53	31– 65	40– 79	51– 93	62–109	75–125	89–142	104–160	119–180	136–200	154–221	173–243	192–267	213–291	235–316	258–342	281–370	306–398	332–427	359–457	387–488
10	15– 45	23– 57	32– 70	42– 84	53– 99	65–115	78–132	92–150	107–169	124–188	141–209	159–231	178–254	198–278	219–303	242–328	265–355	289–383	314–412	340–442	368–472	396–504
11	16– 48	24– 61	34– 74	44– 89	55–105	68–121	81–139	96–157	111–177	128–197	145–219	164–241	183–265	204–289	226–314	248–341	272–368	296–397	322–426	349–456	376–488	405–520
12	17– 51	26– 64	35– 79	46– 94	58–110	71–127	84–146	99–165	115–185	132–206	150–228	169–251	189–275	210–300	232–326	255–353	279–381	304–410	330–440	357–471	385–503	414–536
13	18– 54	27– 68	37– 83	48– 99	60–116	73–134	88–152	103–172	119–193	136–215	155–237	174–261	195–285	216–311	239–337	262–365	286–394	312–423	338–454	365–486	394–518	423–552
14	19– 57	28– 72	38– 88	50–104	62–122	76–140	91–159	106–108	123–201	141–223	160–246	179–271	200–296	222–322	245–349	269–377	293–407	319–437	346–468	374–500	403–533	433–567
15	20– 60	29– 76	40– 92	52–109	65–127	79–146	94–166	110–187	127–209	145–232	164–256	184–281	206–306	228–333	251–361	275–390	301–419	327–450	354–482	382–515	412–548	442–583
16	21– 63	30– 80	42– 96	54–114	67–133	82–152	97–173	113–195	131–217	150–240	169–265	190–290	211–317	234–344	258–372	282–402	308–432	335–463	362–496	391–529	421–563	451–599
17	21– 67	32– 83	43–101	56–119	70–138	84–159	100–180	117–202	135–225	154–249	174–274	195–300	217–327	240–355	264–384	289–414	315–445	342–477	370–510	399–544	429–579	461–614
18	22– 70	33– 87	45–105	58–124	72–144	87–165	103–187	121–209	139–233	159–257	179–283	200–310	223–337	246–366	271–395	296–426	322–458	350–490	378–524	408–558	438–594	470–630
19	23– 73	34– 91	46–110	60–129	74–150	90–171	107–193	124–217	143–241	163–266	184–292	205–320	228–348	252–377	277–407	303–438	330–470	358–503	387–537	416–573	447–609	479–646
20	24– 76	35– 95	48–114	62–134	77–155	93–177	110–200	128–224	147–249	167–275	188–302	211–329	234–358	258–388	283–419	310–450	337–483	365–517	395–551	425–587	456–624	489–661
21	25– 79	37– 98	50–118	64–139	79–161	95–184	113–207	131–232	151–257	172–283	193–311	216–339	240–368	264–399	290–430	317–462	344–496	373–530	403–565	434–601	465–639	498–677
22	26– 82	38–102	51–123	66–144	81–167	98–190	116–214	135–239	155–265	176–292	198–320	221–349	245–379	270–410	296–442	324–474	352–508	381–543	411–579	442–616	474–654	508–692
23	27– 85	39–106	53–127	68–149	84–172	101–196	119–221	139–246	159–273	180–301	203–329	226–359	251–389	276–421	303–453	330–487	359–521	389–556	419–593	451–630	483–669	517–708
24	27– 89	40–110	54–132	70–154	86–178	104–202	123–227	142–254	163–281	185–309	208–338	232–368	257–399	282–432	309–465	337–499	366–534	396–570	427–607	459–645	492–684	527–723
25	28– 92	42–113	56–136	72–159	89–183	107–208	126–234	146–261	167–289	189–318	213–347	237–378	262–410	289–442	316–476	344–511	374–546	404–583	436–620	468–659	502–698	536–739

De cursieve kritieke waarden zijn exact [zij zijn berekend uit de kritieke waarden van W in Wabeke en Van Eeden (1955)]. De overige kritieke waarden zijn benaderend en ontleend aan Geigy (1960).

TABEL L–3. DE TOETS VAN WILCOXON

Linker en rechter kritieke waarden van S bij tweezijdige toetsing

$\alpha = 0{,}02$

n_2 \ n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
4	—	*15– 35*	*22– 44*	*29– 55*	*38– 66*	*48– 78*	*58– 92*	*70–106*	*83–121*	*96–138*	*111–155*	*127–173*	*143–193*	*161–213*	*180–234*	*199–257*	*220–280*	*242–304*	*264–330*	*288–356*	*313–383*	*338–412*
5	*10– 30*	16– 39	*23– 49*	*31– 60*	*40– 72*	*50– 85*	*61– 99*	*73–114*	*86–130*	*100–147*	*115–165*	*131–184*	*148–204*	*166–225*	*185–247*	*205–270*	*226–294*	*248–319*	*271–345*	*295–372*	*320–400*	*346–429*
6	*11– 33*	*17– 43*	*24– 54*	*32– 66*	*42– 78*	*52– 92*	*63–107*	*75–123*	*89–139*	*103–157*	*118–176*	*135–195*	*152–216*	*171–237*	*190–260*	*210–284*	*232–308*	*254–334*	*277–361*	*302–388*	*327–417*	*354–446*
7	*11– 37*	*18– 47*	*25– 59*	34– 71	*43– 85*	*54– 99*	*66–114*	*78–131*	*92–148*	*107–166*	*122–186*	*139–206*	*157–227*	*176–249*	*195–273*	*216–297*	*238–322*	*261–348*	*284–376*	*309–404*	*335–433*	*361–464*
8	*12– 40*	*19– 51*	*27– 63*	*35– 77*	45– 91	*56–106*	*68–122*	*81–139*	*95–157*	*111–175*	*127–195*	*144–216*	*162–238*	*181–261*	*201–285*	*222– 310*	*244–336*	*267–363*	*291–391*	*316–420*	*342–450*	*370–480*
9	*13– 43*	*20– 55*	*28– 68*	*37– 82*	*47– 97*	59–112	*71–129*	*84–147*	*99–165*	*114–185*	*131–205*	*148–227*	*167–249*	*186–273*	*207–297*	*228–323*	*250–350*	*274–377*	*298–406*	*324–435*	*350–466*	*378–497*
10	*13– 47*	*21– 59*	*29– 73*	*39– 87*	*49–103*	*61–119*	74–136	*88–154*	*102–174*	*118–194*	*135–215*	*153–237*	*172–260*	*191–285*	*212–310*	*234–336*	*257–363*	*281–391*	*306–420*	*331–451*	*358–482*	*386–514*
11	*14– 50*	*22– 63*	*30– 78*	*40– 93*	*51–109*	*63–126*	*77–143*	91–162	*106–182*	*122–203*	*139–225*	*157–248*	*177–271*	*197–296*	*218–322*	*240–349*	*263–377*	*288–405*	*313–435*	*339–466*	*366–498*	394–531
12	*15– 53*	*23– 67*	*32– 82*	*42– 98*	*53–115*	*66–132*	*79–151*	*94–170*	109–191	*126–212*	*143–235*	*162–258*	*182–282*	*202–308*	*224–334*	*246–362*	*270–390*	*295–419*	*320–450*	347–481	374–514	403–547
13	*15– 57*	*24– 71*	*33– 87*	*44–103*	*56–120*	*68–139*	*82–158*	*97–178*	*113–199*	130–221	*148–244*	*167–268*	*187–293*	*208–319*	*230–346*	*253–374*	276–404	301–434	327–465	354–497	382–530	411–564
14	*16– 60*	*25– 75*	*34– 92*	*45–109*	*58–126*	*71–145*	*86–165*	*100–186*	*116–208*	*134–230*	152–254	*171–279*	*192–304*	*213–331*	235–359	259–387	283–417	308–448	335–479	362–512	391–545	420–580
15	*17– 63*	*26– 79*	*36– 96*	*47–114*	*60–132*	*73–152*	*88–172*	*103–194*	*120–216*	*138–239*	*156–264*	176–289	197–315	218–343	241–371	265–400	290–430	315–462	342–494	370–527	399–561	429–596
16	*17– 67*	*27– 83*	*37–101*	*49–119*	*62–138*	*76–158*	*91–179*	*107–201*	*124–224*	*142–248*	*161–273*	181–299	202–326	224–354	247–383	271–413	296–444	323–475	350–508	378–542	407–577	437–613
17	*18– 70*	*28– 87*	*39–106*	*51–124*	*64–144*	*78–165*	*93–187*	*110–209*	*127–233*	*146–257*	*165–283*	185–310	207–337	229–366	253–395	278–425	303–457	330–489	357–523	386–557	416–592	446–629
18	*19– 73*	*29– 91*	*40–110*	*52–130*	*66–150*	*81–171*	*96–194*	*113–217*	*131–214*	*150–266*	169–293	190–320	212–348	235–377	259–407	284–438	310–470	337–503	365–537	394–572	424–608	455–645
19	*19– 77*	*30– 95*	*41–116*	*54–135*	*68–156*	*83–178*	*99–201*	*116–225*	*134–250*	*154–275*	174–302	195–330	217–359	241–388	265–419	290–451	317–483	344–517	373–551	402–587	432–624	464–661
20	*20– 80*	*31– 99*	*43–119*	*56–140*	*70–162*	*85–185*	*102–208*	*119–233*	*138–258*	157–285	178–312	200–340	222–370	246–400	271–431	297–463	323–497	351–531	380–566	410–602	441–639	473–677
21	*21– 83*	*32–103*	*44–124*	*58–145*	*72–168*	*88–191*	*105–215*	*123–240*	*142–266*	161–294	182–322	204–351	228–380	252–411	277–443	303–476	330–510	359–544	388–580	418–617	449–655	482–693
22	*21– 87*	*33–107*	*45–129*	*59–151*	*74–174*	*90–198*	*108–222*	*126–248*	*145–275*	165–303	187–331	209–361	233–391	257–423	283–455	310–488	337–523	366–558	395–595	426–632	458–670	490–710
23	*22– 90*	*34–111*	*47–133*	*61–156*	*76–180*	*93–204*	*110–230*	*129–256*	149–283	169–312	191–341	214–371	238–402	263–434	289–467	316–501	344–536	373–572	403–609	434–647	466–686	499–726
24	*23– 93*	*35–115*	*48–138*	*63–161*	*78–186*	*95–211*	*113–237*	*132–264*	152–292	173–321	196–350	219–381	243–413	269–445	295–479	322–514	351–549	380–586	411–623	442–662	475–701	508–742
25	*23– 97*	*36–119*	*50–142*	*64–167*	*81–191*	*98–217*	*116–244*	135–272	156–300	177–330	200–360	224–391	248–424	274–457	301–491	329–528	358–562	388–599	418–638	450–677	483–717	517–758

De cursieve kritieke waarden zijn exact [zij zijn berekend uit de kritieke waarden van W in WABEKE en VAN EEDEN (1955)]. De overige kritieke waarden zijn benaderend en ontleend aan GEIGY (1960).

Tabel L–4. De toets van Wilcoxon

Linker en rechter kritieke waarden van S bij tweezijdige toetsing

$\alpha = 0{,}01$

n_2 \ n_1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
4	—	—	21– 45	28– 56	37– 67	46– 80	57– 93	68–108	81–123	94–140	109–157	125–175	141–195	159–215	177–237	197–259	218–282	239–307	262–332	285–359	310–386	335–415
5	—	15– 40	22– 50	29– 62	38– 74	48– 87	59–101	71–116	84–132	98–149	112–168	128–187	145–207	163–228	182–250	202–273	223–297	245–322	267–349	291–376	316–404	342–433
6	10– 34	16– 44	23– 55	31– 67	40– 80	50– 94	61–109	73–125	87–141	101–159	116–178	132–198	149–219	168–240	187–263	207–287	228–312	250–338	274–364	298–392	323–421	349–451
7	10– 38	16– 49	24– 60	32– 73	42– 86	52–101	64–116	76–133	90–150	104–169	120–188	136–209	154–230	172–253	192–276	212–301	234–326	256–353	280–380	305–408	330–438	357–468
8	11– 41	17– 53	25– 65	34– 78	43– 93	54–108	66–124	79–141	93–159	108–178	123–199	140–220	158–242	177–265	197–289	218–314	240–340	263–367	287–395	311–425	337–455	364–486
9	11– 45	18– 57	26– 70	35– 84	45– 99	56–115	68–132	82–149	96–168	111–188	127–209	144–231	163–253	182–277	202–302	223–328	246–354	269–382	293–411	319–440	345–471	372–503
10	12– 48	19– 61	27– 75	37– 89	47–105	58–122	71–139	84–158	99–177	115–197	131–219	149–241	167–265	187–289	208–314	229–341	252–368	275–397	300–426	326–456	352–488	380–520
11	12– 52	20– 65	28– 80	38– 95	49–111	61–128	73–147	87–166	102–186	118–207	135–229	153–252	172–276	192–301	213–327	235–354	258–382	282–411	307–441	333–472	360–504	388–537
12	13– 55	21– 69	30– 84	40–100	51–117	63–135	76–154	90–174	105–195	122–216	139–239	157–263	177–287	197–313	218–340	241–367	264–396	289–425	314–456	340–488	368–520	396–554
13	13– 59	22– 73	31– 89	41–106	53–123	65–142	79–161	93–182	109–203	125–226	143–249	162–273	181–299	202–325	224–352	247–380	271–409	295–440	321–471	348–503	376–536	404–571
14	14– 62	22– 78	32– 94	43–111	54–130	67–149	81–169	96–190	112–212	129–235	147–259	166–284	186–310	208–336	230–364	253–393	277–423	302–454	328–486	355–519	383–553	413–587
15	15– 65	23– 82	33– 99	44–117	56–136	69–156	84–176	99–198	115–221	133–244	151–269	171–294	191–321	213–348	235–377	259–406	283–437	309–468	335–501	363–534	391–569	421–604
16	15– 69	24– 86	34–104	46–122	58–142	72–162	86–184	102–206	119–229	136–254	155–279	175–305	196–332	218–360	241–389	265–419	290–450	315–483	342–516	370–550	399–585	429–621
17	16– 72	25– 90	36–108	47–128	60–148	74–169	89–191	105–214	122–238	140–263	160–288	180–315	201–343	223–372	246–402	271–432	296–464	322–497	350–530	378–565	407–601	437–638
18	16– 76	26– 94	37–113	49–133	62–154	76–176	92–198	108–222	125–247	144–272	164–298	184–326	206–354	228–384	252–414	277–445	302–478	329–511	357–545	385–581	415–617	446–654
19	17– 79	27– 98	38–118	50–139	64–160	78–183	94–206	111–230	129–255	148–281	168–308	189–336	211–365	234–395	258–426	283–458	309–491	336–525	364–560	393–596	423–633	454–671
20	18– 82	28–102	39–123	52–144	66–166	81–189	97–213	114–238	132–264	152–290	172–318	193–347	216–376	239–407	263–439	289–471	315–505	343–539	371–575	401–611	431–649	463–687
21	18– 86	29–106	40–128	53–150	68–172	83–196	99–221	117–246	136–272	155–300	176–328	198–357	220–388	244–419	269–451	295–484	322–518	350–553	378–590	408–627	439–665	471–704
22	19– 89	29–111	42–132	55–155	70–178	85–203	102–228	120–254	139–281	159–309	180–338	202–368	225–399	250–430	275–463	301–497	328–532	356–568	386–604	416–642	447–681	480–720
23	19– 93	30–115	43–137	57–160	71–185	88–209	105–235	123–262	142–290	163–318	184–348	207–378	230–410	255–442	280–476	307–510	335–545	363–582	393–619	424–657	455–697	488–737
24	20– 96	31–119	44–142	58–166	73–191	90–216	107–243	126–270	146–298	167–327	188–358	211–389	235–421	260–454	286–488	313–523	341–559	370–596	400–634	431–673	464–712	497–753
25	20–100	32–123	45–147	60–171	75–197	92–223	110–250	129–278	149–307	170–337	193–367	216–399	240–432	265–466	292–500	319–536	348–572	377–610	408–648	439–688	472–728	505–770

De cursieve kritieke waarden zijn exact [zij zijn berekend uit de kritieke waarden van W in Wabeke en Van Eeden (1955)]. De overige kritieke waarden zijn benaderend en ontleend aan Geigy (1960).

TABEL M. DE TOETS VAN KRUSKAL EN WALLIS

Exacte rechtse overschrijdingskansen van H voor $k = 3$ en kleine steekproeven

n_1 n_2 n_3	H	P
2 1 1	2,7000	0,500
2 2 1	3,6000	0,200
2 2 2	4,5714	0,067
	3,7143	0,200
3 1 1	3,2000	0,300
3 2 1	4,2857	0,100
	3,8571	0,133
3 2 2	5,3572	0,029
	4,7143	0,048
	4,5000	0,067
	4,4643	0,105
3 3 1	5,1429	0,043
	4,5714	0,100
	4,0000	0,129
3 3 2	6,2500	0,011
	5,3611	0,032
	5,1389	0,061
	4,5556	0,100
	4,2500	0,121
3 3 3	7,2000	0,004
	6,4889	0,011
	5,6889	0,029
	5,6000	0,050
	5,0667	0,086
	4,6222	0,100
4 1 1	3,5714	0,200
4 2 1	4,8214	0,057
	4,5000	0,076
	4,0179	0,114
4 2 2	6,0000	0,014
	5,3333	0,033
	5,1250	0,052
	4,4583	0,100
	4,1667	0,105
4 3 1	5,8333	0,021
	5,2083	0,050
	5,0000	0,057
	4,0556	0,093
	3,8889	0,129
4 3 2	6,4444	0,008
	6,3000	0,011
	5,4444	0,046
	5,4000	0,051
	4,5111	0,098
	4,4444	0,102
4 3 3	6,7455	0,010
	6,7091	0,013
	5,7909	0,046
	5,7273	0,050
	4,7091	0,092
	4,7000	0,101
4 4 1	6,6667	0,010
	6,1667	0,022
	4,9667	0,048
	4,8667	0,054
	4,1667	0,082
	4,0667	0,102
4 4 2	7,0364	0,006
	6,8727	0,011
	5,4545	0,046
	5,2364	0,052
	4,5545	0,098
	4,4455	0,103
4 4 3	7,1439	0,010
	7,1364	0,011
	5,5985	0,049
	5,5758	0,051
	4,5455	0,099
	4,4773	0,102
4 4 4	7,6538	0,008
	7,5385	0,011
	5,6923	0,049
	5,6538	0,054
	4,6539	0,097
	4,5001	0,104
5 1 1	3,8571	0,143
5 2 1	5,2500	0,036
	5,0000	0,048
	4,4500	0,071
	4,2000	0,095
	4,0500	0,119
5 2 2	6,5333	0,008
	6,1333	0,013
	5,1600	0,034
	5,0400	0,056
	4,3733	0,090
	4,2933	0,122
5 3 1	6,4000	0,012
	4,9600	0,048
	4,8711	0,052
	4,0178	0,095
	3,8400	0,123
5 3 2	6,9091	0,009
	6,8218	0,010
	5,2509	0,049
	5,1055	0,052
	4,6509	0,091
	4,4945	0,101
5 3 3	7,0788	0,009
	6,9818	0,011
	5,6485	0,049
	5,5152	0,051
	4,5333	0,097
	4,4121	0,109
5 4 1	6,9545	0,008
	6,8400	0,011
	4,9855	0,044
	4,8600	0,056
	3,9873	0,098
	3,9600	0,102
5 4 2	7,2045	0,009
	7,1182	0,010
	5,2727	0,049
	5,2682	0,050
	4,5409	0,098
	4,5182	0,101
5 4 3	7,4449	0,010
	7,3949	0,011
	5,6564	0,049
	5,6308	0,050
	4,5487	0,099
	4,5231	0,103
5 4 4	7,7604	0,009
	7,7440	0,011
	5,6571	0,049
	5,6176	0,050
	4,6187	0,100
	4,5527	0,102
5 5 1	7,3091	0,009
	6,8364	0,011
	5,1273	0,046
	4,9091	0,053
	4,1091	0,086
	4,0364	0,105
5 5 2	7,3385	0,010
	7,2692	0,010
	5,3385	0,047
	5,2462	0,051
	4,6231	0,097
	4,5077	0,100
5 5 3	7,5780	0,010
	7,5429	0,010
	5,7055	0,046
	5,6264	0,051
	4,5451	0,100
	4,5363	0,102
5 5 4	7,8229	0,010
	7,7914	0,010
	5,6657	0,049
	5,6429	0,050
	4,5229	0,099
	4,5200	0,101
5 5 5	8,0000	0,009
	7,9800	0,010
	5,7800	0,049
	5,6600	0,051
	4,5600	0,100
	4,5000	0,102

Naar KRUSKAL en WALLIS (1952), incl. latere correcties (J.Am.Stat.Ass. 48, 910).

Tabel N. De rangcorrelatietoets van Spearman

Linker en rechter kritieke waarden van $\underline{R}$ bij tweezijdige toetsing

1. Exact

n	Linker kritieke waarden					Rechter kritieke waarden				
	$\alpha =$ 0,001	0,01	0,02	0,05	0,10	0,10	0,05	0,02	0,01	0,001
5	—	—	0	0	2	38	40	40	—	—
6	—	0	2	4	6	64	66	68	70	—
7	0	4	6	12	16	96	100	106	108	112
8	2	10	14	22	30	138	146	154	158	166

2. Benaderende waarden

9	8	20	26	38	48	192	202	214	220	232
10	16	34	44	58	72	258	272	286	296	314
11	24	50	63	83	103	337	357	377	390	416
12	36	71	89	116	143	429	456	483	501	536
13	52	98	121	158	191	537	570	607	630	676
14	72	132	161	207	247	663	703	748	778	838
15	97	175	211	266	313	807	854	909	945	1023
16	129	227	271	335	391	969	1025	1089	1133	1231
17	167	290	341	416	480	1152	1216	1291	1342	1465
18	214	363	422	508	582	1356	1430	1516	1575	1724
19	269	447	514	613	698	1582	1667	1766	1833	2011
20	335	544	620	731	828	1832	1929	2040	2116	2325
21	412	653	738	865	973	2107	2215	2342	2427	2668
22	501	775	871	1013	1135	2407	2529	2671	2767	3041
23	604	912	1020	1178	1314	2734	2870	3028	3136	3444
24	721	1064	1184	1360	1511	3089	3240	3416	3536	3879
25	853	1232	1365	1559	1727	3473	3641	3835	3968	4347
26	1000	1418	1564	1778	1962	3888	4072	4286	4432	4850
27	1161	1621	1781	2016	2219	4333	4536	4771	4931	5391
28	1340	1842	2018	2275	2497	4811	5033	5290	5466	5968
29	1535	2083	2275	2556	2797	5323	5564	5845	6037	6585
30	1748	2344	2553	2859	3122	5868	6131	6437	6646	7242
31	1980	2627	2853	3185	3470	6450	6735	7067	7293	7940
32	2231	2931	3176	3535	3844	7068	7377	7736	7981	8681
33	2503	3259	3523	3910	4244	7724	8058	8445	8709	9465
34	2795	3610	3894	4311	4670	8420	8779	9196	9480	10295
35	3110	3985	4291	4740	5125	9155	9540	9989	10295	11170
36	3448	4386	4714	5195	5609	9931	10345	10826	11154	12092
37	3809	4814	5165	5680	6123	10749	11192	11707	12058	13063
38	4195	5268	5643	6194	6667	11611	12084	12635	13010	14083
39	4606	5751	6151	6738	7243	12517	13022	13609	14009	15154
40	5043	6263	6689	7314	7852	13468	14006	14631	15057	16277
41	5507	6804	7257	7922	8494	14466	15038	15703	16156	17453
42	5999	7376	7857	8563	9170	15512	16119	16825	17306	18683
43	6519	7980	8489	9238	9882	16606	17250	17999	18508	19969
44	7069	8616	9155	9948	10630	17750	18432	19225	19764	21311
45	7649	9285	9856	10694	11415	18945	19666	20504	21075	22711
46	8261	9988	10591	11477	12239	20191	20953	21839	22442	24169
47	8904	10727	11363	12297	13101	21491	22295	23229	23865	25688
48	9580	11501	12172	13156	14003	22845	23692	24676	25347	27268
49	10291	12312	13018	14055	14946	24254	25145	26182	26888	28909
50	11035	13161	13904	14994	15931	25719	26656	27746	28489	30615

Uit Geigy (1960).

TABEL O. DE RANGCORRELATIETOETS VAN KENDALL

Rechter kritieke waarden van S bij tweezijdige toetsing, $S_L = -S_R$

n	Onbetrouwbaarheidsdrempel α			
	0,01	0,02	0,05	0,10
4	—	—	—	6
5	—	10	10	8
6	15	13	13	11
7	19	17	15	13
8	22	20	18	16
9	26	24	20	18
10	29	27	23	21
11	33	31	27	23
12	38	36	30	26
13	44	40	34	28
14	47	43	37	33
15	53	49	41	35
16	58	52	46	38
17	64	58	50	42
18	69	63	53	45
19	75	67	57	49
20	80	72	62	52
21	86	78	66	56
22	91	83	71	61
23	99	89	75	65
24	104	94	80	68
25	110	100	86	72
26	117	107	91	77
27	125	113	95	81
28	130	118	100	86
29	138	126	106	90
30	145	131	111	95
31	151	137	117	99
32	160	144	122	104
33	166	152	128	108
34	175	157	133	113
35	181	165	139	117
36	190	172	146	122
37	198	178	152	128
38	205	185	157	133
39	213	193	163	139
40	222	200	170	144

Uit KAARSEMAKER & VAN WIJNGAARDEN (1952).

TABEL P–1. DE SERIETOETS OP DE WAARNEMINGEN

Linker en rechter kritieke waarden van u bij tweezijdige toetsing, $m_1 \neq m_2$

$\alpha = 0{,}05$

m_2 \ m_1	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5		10	11	11	—	—	—	—	—	—	—	—	—	—	—	—
6	3		12	12	13	13	13	13	—	—	—	—	—	—	—	—
7	3	3		13	14	14	14	14	15	15	15	—	—	—	—	—
8	3	3	4		14	15	15	16	16	16	16	17	17	17	17	17
9	3	4	4	5		16	16	16	17	17	18	18	18	18	18	18
10	3	4	5	5	5		17	17	18	18	18	19	19	19	20	20
11	4	4	5	5	6	6		18	19	19	19	20	20	20	21	21
12	4	4	5	6	6	7	7		19	20	20	21	21	21	22	22
13	4	5	5	6	6	7	7	8		20	21	21	22	22	23	23
14	4	5	5	6	7	7	8	8	9		22	22	23	23	23	24
15	4	5	6	6	7	7	8	8	9	9		23	23	24	24	25
16	4	5	6	6	7	8	8	9	9	10	10		24	25	25	25
17	4	5	6	7	7	8	9	9	10	10	11	11		25	26	26
18	5	5	6	7	8	8	9	9	10	10	11	11	12		26	27
19	5	6	6	7	8	8	9	10	10	11	11	12	12	13		27
20	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	

$\alpha = 0{,}01$

m_2 \ m_1	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5		11	—	—	—	—	—	—	—	—	—	—	—	—	—	—
6	2		13	13	—	—	—	—	—	—	—	—	—	—	—	—
7	2	2		14	15	15	15	—	—	—	—	—	—	—	—	—
8	2	3	3		15	16	16	17	17	17	—	—	—	—	—	—
9	2	3	3	3		17	17	18	18	18	19	19	19	—	—	—
10	3	3	3	4	4		18	19	19	19	20	20	20	21	21	21
11	3	3	4	4	5	5		19	20	20	21	21	22	22	22	22
12	3	3	4	4	5	5	6		21	21	22	22	22	23	23	23
13	3	3	4	5	5	5	6	6		22	22	23	23	24	24	24
14	3	4	4	5	5	6	6	7	7		23	24	24	25	25	25
15	3	4	4	5	6	6	7	7	7	8		24	25	25	26	26
16	3	4	5	5	6	6	7	7	8	8	9		26	26	27	27
17	3	4	5	5	6	7	7	8	8	8	9	9		27	27	28
18	4	4	5	6	6	7	7	8	8	9	9	10	10		28	29
19	4	4	5	6	6	7	8	8	9	9	10	10	10	11		29
20	4	4	5	6	7	7	8	8	9	9	10	10	11	11	12	

Uit SWED en EISENHART (1943).

TABEL P–1. DE SERIETOETS OP DE WAARNEMINGEN (*vervolg*)

Linker en rechter kritieke waarden van u bij tweezijdige toetsing, $m_1 \neq m_2$

$\alpha = 0{,}10$

m_1 / m_2	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5		10	10	11	11	11	—	—	—	—	—	—	—	—	—	—
6	3		11	12	12	12	13	13	13	13	—	—	—	—	—	—
7	3	4		13	13	13	14	14	14	14	15	15	15	15	15	15
8	3	4	4		14	14	15	15	15	16	16	16	16	16	16	17
9	4	4	5	5		15	15	16	16	17	17	17	17	18	18	18
10	4	5	5	6	6		16	17	17	17	18	18	18	19	19	19
11	4	5	5	6	6	7		17	18	18	19	19	19	20	20	20
12	4	5	6	6	7	7	8		18	19	19	20	20	21	21	21
13	4	5	6	6	7	8	8	9		20	20	21	21	22	22	22
14	5	5	6	7	7	8	8	9	9		21	21	22	22	23	23
15	5	6	6	7	8	8	9	9	10	10		22	22	23	23	24
16	5	6	6	7	8	8	9	10	10	11	11		23	24	24	25
17	5	6	7	7	8	9	9	10	10	11	11	12		24	25	25
18	5	6	7	8	8	9	10	10	11	11	12	12	13		25	26
19	5	6	7	8	8	9	10	10	11	12	12	13	13	14		27
20	5	6	7	8	9	9	10	11	11	12	12	13	13	14	14	

$\alpha = 0{,}02$

m_1 / m_2	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5		11	11	—	—	—	—	—	—	—	—	—	—	—	—	—
6	2		12	13	13	—	—	—	—	—	—	—	—	—	—	—
7	2	3		14	14	15	15	15	—	—	—	—	—	—	—	—
8	2	3	3		15	15	16	16	17	17	17	17	—	—	—	—
9	3	3	4	4		16	17	17	18	18	18	18	19	19	19	19
10	3	3	4	4	5		18	18	19	19	19	20	20	20	20	20
11	3	4	4	5	5	5		19	19	20	20	21	21	21	22	22
12	3	4	4	5	5	6	6		20	21	21	22	22	22	23	23
13	3	4	5	5	6	6	6	7		21	22	22	23	23	24	24
14	3	4	5	5	6	6	7	7	8		23	23	24	24	24	25
15	4	4	5	5	6	7	7	8	8	8		24	24	25	25	26
16	4	4	5	6	6	7	7	8	8	9	9		25	26	26	26
17	4	5	5	6	7	7	8	8	9	9	10	10		26	27	27
18	4	5	5	6	7	7	8	8	9	9	10	10	11		27	28
19	4	5	6	6	7	8	8	9	9	10	10	11	11	12		29
20	4	5	6	6	7	8	8	9	10	10	11	11	11	12	12	

TABEL P–2. DE SERIETOETS OP DE WAARNEMINGEN

Linker en rechter kritieke waarden van u bij tweezijdige toetsing, $m_1 = m_2 = \frac{1}{2}n$

n	Linker kritieke waarden				Rechter kritieke waarden			
	α=0,01	0,02	0,05	0,10	0,10	0,05	0,02	0,01
8	—	—	—	2	8	—	—	—
10	—	2	2	3	9	10	10	11
12	2	2	3	3	11	11	12	12
14	3	3	3	4	12	13	13	13
16	3	4	4	5	13	14	14	15
18	4	4	5	6	14	15	16	16
20	4	5	6	6	16	16	17	17
22	5	6	7	7	17	17	18	19
24	6	6	7	8	18	19	19	20
26	7	7	8	9	19	20	21	21
28	7	8	9	10	20	21	22	23
30	8	9	10	11	21	22	23	24
32	9	10	11	11	23	23	24	25
34	10	10	11	12	24	25	26	26
36	10	11	12	13	25	26	27	27
38	11	12	13	14	26	27	28	29
40	12	13	14	15	27	28	29	30
42	13	14	15	16	28	29	30	31
44	14	14	16	17	29	30	32	32
46	14	15	16	17	31	32	33	34
48	15	16	17	18	32	33	34	35
50	16	17	18	19	33	34	35	36
52	17	18	19	20	34	35	36	37
54	18	19	20	21	35	36	37	38
56	18	19	21	22	36	37	39	40
58	19	20	22	23	37	38	40	41
60	20	21	22	24	38	40	41	42
62	21	22	23	25	39	41	42	43
64	22	23	24	25	41	42	43	44
66	23	24	25	26	42	43	44	45
68	23	24	26	27	43	44	46	47
70	24	25	27	28	44	45	47	48
72	25	26	28	29	45	46	48	49
74	26	27	29	30	46	47	49	50
76	27	28	30	31	47	48	50	51
78	28	29	30	32	48	50	51	52
80	29	30	31	33	49	51	52	53
82	29	31	32	34	50	52	53	55
84	30	31	33	35	51	53	54	56
86	31	32	34	35	53	54	56	57
88	32	33	35	36	54	55	57	58
90	33	34	36	37	55	56	58	59

Ontleend aan SWED en EISENHART (1943).

TABEL Q. DE SERIETOETS OP VERSCHILLEN TUSSEN OPEENVOLGENDE WAARNEMINGEN

Linker en rechter kritieke waarden van v bij éénzijdige toetsing

n	Linker kritieke waarden		Rechter kritieke waarden	
	$\alpha = 0{,}01$	0,05	0,05	0,01
5	—	1	—	—
6	1	1	—	—
7	1	2	—	—
8	2	2	—	—
9	2	3	8	—
10	3	3	9	—
11	3	4	10	—
12	4	4	11	—
13	4	5	12	—
14	5	6	12	13
15	5	6	13	14
16	6	7	14	15
17	6	7	15	16
18	7	8	15	17
19	7	8	16	17
20	8	9	17	18
21	8	10	18	19
22	9	10	18	20
23	10	11	19	21
24	10	11	20	21
25	11	12	21	22
26	11	13	21	23
27	12	13	22	24
28	12	14	23	24
29	13	14	24	25
30	13	15	24	26
31	14	16	25	27
32	15	16	26	27
33	15	17	27	28
34	16	17	27	29
35	16	18	28	30
36	17	19	29	30
37	18	19	29	31
38	18	20	30	32
39	19	20	31	33
40	19	21	32	33

Uit SWED en EISENHART (1943) en EDGINGTON (1961).

TABEL R–1

95% en 99% betrouwbaarheidsgrenzen van het *percentage* elementen met kenmerk A in de populatie, indien x van de n elementen in een aselecte steekproef dit kenmerk bezitten

x	95%	99%
$n = 5$		
0	0 —52,2	0 —65,3
1	0,5—71,6	0,1—81,5
2	5,3—85,3	2,3—91,7
3	14,7—94,7	8,3—97,7
$n = 6$		
0	0 —45,9	0 —58,7
1	0,4—64,1	0,1—74,6
2	4,3—77,7	1,9—85,6
3	11,8—88,2	6,6—93,4
$n = 7$		
0	0 —41,0	0 —53,1
1	0,4—57,9	0,1—68,5
2	3,7—71,0	1,6—79,7
3	9,9—81,6	5,5—88,2
4	18,4—90,1	11,8—94,5
$n = 8$		
0	0 —36,9	0 —48,4
1	0,3—52,7	0,1—63,1
2	3,2—65,1	1,4—74,2
3	8,5—75,5	4,7—83,0
4	15,7—84,3	10,0—90,0
$n = 9$		
0	0 —33,6	0 —44,5
1	0,3—48,3	0,1—58,5
2	2,8—60,0	1,2—69,3
3	7,5—70,1	4,2—78,1
4	13,7—78,8	8,7—85,4
5	21,2—86,3	14,6—91,3
$n = 10$		
0	0 —30,9	0 —41,1
1	0,3—44,5	0,1—54,4
2	2,5—55,6	1,1—64,8
3	6,7—65,3	3,7—73,5
4	12,2—73,8	7,7—80,9
5	18,7—81,3	12,8—87,2
$n = 12$		
0	0 —26,5	0 —35,7
1	0,2—38,5	0,0—47,7
2	2,1—48,4	0,9—57,3
3	5,5—57,2	3,0—65,5
4	9,9—65,1	6,2—72,7
5	15,2—72,3	10,3—79,1
6	21,1—78,9	15,2—84,8
$n = 14$		
0	0 —23,2	0 —31,5
1	0,2—33,9	0,0—42,4
2	1,8—42,8	0,8—51,2
3	4,7—50,8	2,6—58,9
4	8,4—58,1	5,3—65,8
5	12,8—64,9	8,7—72,0
6	17,7—71,1	12,8—77,7
7	23,0—77,0	17,2—82,8
$n = 16$		
0	0 —20,6	0 —28,2
1	0,2—30,2	0,0—38,1
2	1,5—38,3	0,7—46,3
3	4,1—45,7	2,2—53,4
4	7,3—52,4	4,5—59,9
5	11,0—58,7	7,5—65,9
6	15,2—64,6	10,9—71,3
7	19,7—70,1	14,7—76,4
8	24,7—75,3	19,0—81,0
$n = 20$		
0	0 —16,8	0 —23,3
1	0,1—24,9	0,0—31,7
2	1,2—31,7	0,5—38,7
3	3,2—37,9	1,8—44,9
4	5,7—43,7	3,6—50,7
5	8,7—49,1	5,8—56,0
6	11,9—54,3	8,5—61,0
7	15,4—59,2	11,4—65,7
8	19,1—63,9	14,6—70,1
9	23,1—68,5	18,1—74,3
10	27,2—72,8	21,8—74,2
$n = 25$		
0	0 —13,7	0 —19,1
1	0,1—20,3	0,0—26,2
2	1,0—26,0	0,4—32,1
3	2,5—31,2	1,4—37,4
4	4,5—36,1	2,8—42,3
5	6,8—40,7	4,6—47,0
6	9,4—45,1	6,6—51,4
7	12,1—49,4	8,9—55,5
8	14,9—53,5	11,3—59,5
9	18,0—57,5	14,0—63,3
10	21,1—61,3	16,8—67,0
11	24,4—65,1	19,7—70,5
12	27,8—68,7	22,8—73,9
13	31,3—72,2	26,1—27,2
$n = 30$		
0	0 —11,6	0 —16,2
1	0,1—17,2	0,0—22,3
2	0,8—22,1	0,3—27,4
3	2,1—26,5	1,2—32,0
4	3,8—30,7	2,3—36,3
5	5,6—34,7	3,8—40,4
6	7,7—38,6	5,5—44,3
7	9,9—42,3	7,3—48,0
8	12,3—45,9	9,3—51,6
9	14,7—49,4	11,4—55,0
10	17,3—52,8	13,7—58,3
11	19,9—56,1	16,0—61,6
12	22,7—59,4	18,5—64,7
13	25,5—62,6	21,1—67,7
14	28,3—65,7	23,7—70,7
15	31,3—68,7	26,5—73,5
$n = 40$		
0	0 — 8,8	0 —12,4
1	0,1—13,2	0,0—17,1
2	0,6—16,9	0,3—21,2
3	1,6—20,4	0,9—24,8
4	2,8—23,7	1,7—28,3
5	4,2—26,8	2,8—31,5
6	5,7—29,8	4,0—34,6
7	7,3—32,8	5,4—37,6
8	9,1—35,7	6,8—40,5
9	10,8—38,5	8,4—43,4
10	12,7—41,2	10,0—46,1
11	14,6—43,9	11,7—48,8
12	16,6—46,5	13,4—51,4
13	18,6—49,1	15,3—54,0
14	20,6—51,7	17,1—56,5
15	22,7—54,2	19,1—59.0
16	24,9—56,7	21,1—61,4
17	27,0—59,1	23,1—63,7
18	29,3—61,5	25,2—66,1
19	31,5—63,9	27,3—68,3
20	33,8—66,2	29,5—70,5
$n = 50$		
0	0 — 7,1	0 —10,1
1	0,1—10,7	0,0—13,9
2	0,5—13,7	0,2—17,3
3	1,3—16,5	0,7—20,3
4	2,2—19,2	1,4—23,1
5	3,3—21,8	2,2—25,8
6	4,5—24,3	3,2—28,4
7	5,8—26,7	4,3—30,9
8	7,2—29,1	5,4—33,3
9	8,6—31,4	6,6—35,7
10	10,0—33,7	7,9—38,1
11	11,5—36,0	9,2—40,3
12	13,1—38,2	10,6—42,5
13	14,6—40,3	12,0—44,7
14	16,2—42,5	13,4—46,9
15	17,9—44,6	14,9—49,0
16	19,5—46,7	16,4—51,1
17	21,2—48,8	18,0—53,1
18	22,9—50,8	19,6—55,1
19	24,7—52,8	21,2—57,1
20	26,4—54,8	22,9—59,1
21	28,2—56,8	24,5—61,0
22	30,0—58,7	26,3—62,9
23	31,8—60,7	28,0—64,8
24	33,7—62,6	29,8—66,6
25	35,5—64,5	31,5—68,5
$n = 75$		
0	0 — 4,8	0 — 6,8
1	0,0— 7,2	0,0— 9,5
2	0,3— 9,3	0,1—11,8
3	0,8—11,3	0,5—13,9
4	1,5—13,1	0,9—15,9
5	2,2—14,9	1,5—17,7
6	3,0—16,6	2,1—19,6
7	3,8—18,3	2,8—21,3
8	4,7—19,9	3,5—23,1
9	5,6—21,6	4,3—24,7
10	6,6—23,2	5,1—26,4
11	7,6—24,7	6,0—28,0
12	8,5—26,3	6,9—29,6
13	9,6—27,8	7,8—31,2
14	10,6—29,3	8,7—32,7
15	11,7—30,8	9,7—34,3
16	12,7—32,3	10,6—35,8
17	13,8—33,8	11,6—37,3
18	14,9—35,3	12,6—38,8
19	16,0—36,7	13,7—40,2
20	17,1—38,1	14,7—41,7
21	18,2—39,6	15,7—43,1
22	19,4—41,0	16,8—44,5
23	20,5—42,4	17,9—46,0
24	21,7—43,8	19,0—47,4
25	22,9—45,2	20,1—48,7
26	24,0—46,5	21,2—50,1
27	25,2—47,9	22,3—51,5
28	26,4—49,3	23,5—52,8
29	27,6—50,6	24,7—54,2
30	28,9—52,0	25,8—55,5
31	30,1—53,3	27,0—56,8
32	31,3—54,6	28,2—58,1
33	32,5—55,9	29,4—59,4
34	33,8—57,3	30,6—60,7
35	35,1—58,5	31,8—62,0
36	36,3—59,9	33,0—63,3
37	37,6—61,1	34,3—64,5
38	38,9—62,4	35,5—65,7

TABEL R–2

Betrouwbaarheidsgrenzen van het *percentage* elementen met kenmerk A in de populatie, indien x van de n elementen in een aselecte steekproef dit kenmerk bezitten

1. 95% betrouwbaarheidsgrenzen

x	n = 100	150	200	300	500	1000
0	0 — 3,6	0 — 2,4	0 — 1,8	0 — 1,2	0 — 0,74	0 —0,37
1	0,025— 5,4	0,017— 3,7	0,013— 2,8	0,008— 1,8	0,005— 1,1	0,0025—0,56
2	0,24 — 7,0	0,16 — 4,7	0,12 — 3,6	0,08 — 2,4	0,05 — 1,4	0,024 —0,72
3	0,62 — 8,5	0,41 — 5,7	0,31 — 4,3	0,21 — 2,9	0,12 — 1,7	0,06 —0,87
4	1,1 — 9,9	0,73 — 6,7	0,55 — 5,0	0,36 — 3,4	0,22 — 2,0	0,11 — 1,0
5	1,6 —11,3	1,1 — 7,7	0,80 — 5,8	0,53 — 3,9	0,32 — 2,3	0,16 — 1,2
6	2,2 —12,6	1,5 — 8,6	1,1 — 6,5	0,73 — 4,3	0,43 — 2,6	0,22 — 1,3
7	2,9 —13,9	1,9 — 9,4	1,4 — 7,1	0,93 — 4,8	0,56 — 2,9	0,28 — 1,4
8	3,5 —15,2	2,3 —10,3	1,7 — 7,8	1,2 — 5,2	0,69 — 3,1	0,34 — 1,6
9	4,2 —16,4	2,8 —11,1	2,1 — 8,4	1,4 — 5,6	0,82 — 3,4	0,41 — 1,7
10	4,9 —17,6	3,2 —12,0	2,4 — 9,0	1,6 — 6,1	0,96 — 3,7	0,48 — 1,8
11	5,6 —18,8	3,7 —12,8	2,8 — 9,7	1,8 — 6,5	1,1 — 3,9	0,55 — 2,0
12	6,4 —20,0	4,2 —13,6	3,1 —10,3	2,1 — 6,9	1,2 — 4,2	0,62 — 2,1
13	7,1 —21,2	4,7 —14,4	3,5 —10,9	2,3 — 7,3	1,4 — 4,4	0,69 — 2,2
14	7,9 —22,4	5,2 —15,2	3,9 —11,5	2,6 — 7,7	1,5 — 4,7	0,77 — 2,3
15	8,7 —23,5	5,7 —16,0	4,3 —12,1	2,8 — 8,1	1,7 — 4,9	0,84 — 2,5
16	9,4 —24,7	6,2 —16,8	4,6 —12,7	3,1 — 8,5	1,8 — 5,2	0,92 — 2,6
18	11,0 —26,9	7,3 —18,3	5,4 —13,9	3,6 — 9,3	2,1 — 5,6	1,1 — 2,8
20	12,7 —29,2	8,4 —19,9	6,2 —15,0	4,1 —10,1	2,5 — 6,1	1,2 — 3,1
25	16,9 —34,7	11,1 —23,6	8,3 —17,9	5,5 —12,1	3,3 — 7,3	1,6 — 3,7
30	21,2 —40,0	13,9 —27,3	10,4 —20,7	6,9 —14,0	4,1 — 8,5	2,0 — 4,3
35	25,7 —45,2	16,9 —31,0	12,5 —23,5	8,3 —15,9	4,9 — 9,6	2,5 — 4,8
40	30,3 —50,3	19,8 —34,5	14,7 —26,2	9,7 —17,7	5,8 —10,7	2,9 — 5,4
45	35,0 —55,3	23,5 —38,7	16,9 —28,9	11,2 —19,6	6,6 —11,9	3,3 — 6,0
50	39,8 —60,2	25,9 —41,5	19,2 —31,6	12,6 —21,4	7,5 —13,0	3,7 — 6,5

2. 99% betrouwbaarheidsgrenzen

x	n = 100	150	200	300	500	1000
0	0 — 5,2	0 — 3,5	0 — 2,6	0 — 1,8	0 — 1,1	0 —0,53
1	0,005— 7,2	0,003— 4,9	0,0025— 3,7	0,0017— 2,5	0,001— 1,5	0,0005—0,74
2	0,100— 8,9	0,069— 6,0	0,052 — 4,6	0,034 — 3,1	0,021— 1,8	0,010 —0,92
3	0,34 —10,6	0,23 — 7,1	0,17 — 5,4	0,11 — 3,6	0,068— 2,2	0,034 — 1,1
4	0,68 —12,1	0,45 — 8,2	0,34 — 6,2	0,22 — 4,1	0,14 — 2,5	0,067 — 1,3
5	1,1 —13,5	0,61 — 9,1	0,46 — 6,8	0,30 — 4,6	0,18 — 2,8	0,09 — 1,4
6	1,6 —14,9	0,92 —10,0	0,69 — 7,6	0,45 — 5,1	0,27 — 3,1	0,14 — 1,5
7	2,1 —16,3	1,3 —11,0	0,94 — 8,3	0,62 — 5,6	0,37 — 3,4	0,18 — 1,7
8	2,6 —17,6	1,6 —11,9	1,2 — 9,0	0,80 — 6,0	0,48 — 3,6	0,24 — 1,8
9	3,2 —18,9	2,0 —12,8	1,5 — 9,7	1,0 — 6,5	0,59 — 3,9	0,29 — 2,0
10	3,8 —20,2	2,4 —13,6	1,8 —10,3	1,2 — 6,9	0,71 — 4,2	0,35 — 2,1
11	4,5 —21,5	2,8 —14,5	2,1 —11,0	1,4 — 7,4	0,83 — 4,5	0,41 — 2,3
12	5,1 —22,7	3,3 —15,4	2,4 —11,6	1,6 — 7,8	0,96 — 4,7	0,48 — 2,4
13	5,8 —23,9	3,7 —16,2	2,8 —12,3	1,8 — 8,3	1,1 — 5,0	0,54 — 2,5
14	6,5 —25,1	4,1 —17,1	3,1 —12,9	2,0 — 8,7	1,2 — 5,3	0,60 — 2,7
15	7,2 —26,3	4,6 —17,8	3,4 —13,6	2,3 — 9,1	1,4 — 5,8	0,67 — 2,8
16	7,9 —27,5	5,1 —18,7	3,8 —14,2	2,5 — 9,6	1,5 — 5,8	0,74 — 2,9
18	9,3 —29,8	6,0 —20,3	4,5 —15,4	3,0 —10,4	1,8 — 6.3	0,88 — 3,2
20	10,8 —32,1	7,0 —21,9	5,2 —16,6	3,4 —11,2	2,1 — 6,8	1,0 — 3,4
25	14,8 —37,7	9,6 —25,8	7,1 —19,6	4,7 —13,3	2,8 — 8,0	1,4 — 4,1
30	18,9 —43,1	12,3 —29,6	9,1 —22,5	6,0 —15,2	3,6 — 9,3	1,8 — 4,7
35	23,2 —48,3	15,0 —33,3	11,1 —25,4	7,3 —17,2	4,3 —10,4	2,2 — 5,3
40	27,6 —53,4	17,9 —36,9	13,2 —28,2	8,7 —19,1	5,1 —11,6	2,6 — 5,9
45	32,2 —58,3	20,8 —40,5	15,3 —30,9	10,1 —21,0	6,0 —12,8	3,0 — 6,5
50	36,9 —63,1	23,7 —44,0	17,5 —33,7	11,5 —22,9	6,8 —13,9	3,4 — 7,0

TABEL R–3

95% en 99% betrouwbaarheidsgrenzen van het percentage elementen met kenmerk A in de populatie, wanneer a% van de elementen in een aselecte steekproef van de omvang n dit kenmerk bezitten

a	n = 100		n = 200		n = 500		n = 1000	
	95%	99%	95%	99%	95%	99%	95%	99%
0	*	*	*	*	*	*	*	*
1	*	*	*	*	*	*	*	*
2	*	*	*	*	*	*	1,2— 3,1	1,0— 3,4
3	*	*	*	*	*	*	2,0— 4,3	1,8— 4,8
4	*	*	*	*	*	*	2,9— 5,4	2,6— 5,9
5	*	*	*	*	3,3— 7,3	2,8— 8,0	3,7— 6,5	3,4— 7,0
6	*	*	*	*	4,1— 8,5	3,6— 9,3	4,6— 7,7	4,2— 8,2
7	*	*	*	*	4,9— 9,6	4,3—10,4	5,5— 8,8	5,1— 9,3
8	*	*	*	*	5,8—10,7	5,1—11,6	6,4— 9,9	5,9—10,5
9	*	*	*	*	6,6—11,9	6,0—12,8	7,3—10,9	6,8—11,6
10	*	*	*	*	7,5—13,0	6,8—13,9	8,2—12,0	7,7—12,7
11	*	*	7,0—16,2	6,0—17,8	*8,4—14,1*	*7,7—15,1*	9,1—13,1	8,6—13,8
12	*	*	7,9—17,4	6,7—19,0	9,3—15,2	8,5—16,2	10,1—14,2	9,5—14,9
13	*	*	8,7—18,5	7,5—20,8	*10,2—16,3*	*9,4—17,3*	11,0—15,3	10,4—15,9
14	*	*	9,5—19,6	8,3—21,4	11,1—17,4	10,2—18,4	11,9—16,3	11,3—17,0
15	*	*	10,4—20,7	9,1—22,5	*12,0—18,4*	*11,1—19,5*	12,8—17,4	12,2—18,1
16	*	*	11,2—21,9	9,9—23,7	12,9—19,5	12,0—20,6	13,8—18,4	13,1—19,2
17	10,2—25,8	8,6—28,7	12,1—22,9	10,7—24,8	*13,8—20,6*	*12,9—21,7*	14,7—19,5	14,0—20,3
18	11,0—27,0	9,3—29,8	12,9—24,1	11,5—25,9	14,7—21,7	13,8—22,8	15,7—20,5	15,0—21,3
19	11,8—28,1	10,1—31,0	13,8—25,2	12,4—27,1	*15,7—22,7*	*14,7—23,9*	16,6—21,6	15,9—22,4
20	12,7—29,2	10,8—32,2	14,7—26,2	13,2—28,2	16,6—23,8	15,6—25,0	17,6—22,6	16,8—23,4
21	13,5—30,3	11,6—33,3	15,6—27,3	14,1—29,3	*17,5—24,8*	*16,5—26,0*	18,5—23,7	17,8—24,5
22	*14,3—31,4*	*12,4—34,4*	16,5—28,4	14,9—30,4	18,5—25,9	17,4—27,1	19,5—24,7	18,7—25,5
23	15,2—32,5	13,2—35,5	17,4—30,0	15,8—31,5	*19,4—26,9*	*18,3—28,2*	20,4—25,7	19,7—26,6
24	16,0—33,6	14,0—36,6	18,3—30,5	16,6—32,6	20,3—28,0	19,2—29,2	21,4—26,8	20,6—27,6
25	16,9—34,7	14,8—37,7	19,2—31,6	17,5—33,7	*21,3—29,0*	*20,1—30,3*	22,3—27,8	21,5—28,7
26	17,7—35,7	15,6—38,8	*20,1—32,7*	*18,3—34,7*	22,2—30,1	21,1—31,3	23,3—28,8	22,5—29,7
27	18,6—36,8	16,4—39,9	*21,1—33,8*	*19,2—35,8*	*23,2—31,1*	*22,0—32,4*	24,3—29,9	23,4—30,8
28	19,5—37,9	17,2—40,9	*21,9—34,8*	*20,1—36,9*	24,1—32,2	22,9—33,5	25,2—30,9	24,4—31,8
29	20,4—38,9	18,1—42,0	*22,8—35,9*	*21,0—37,9*	*25,1—33,2*	*23,9—34,5*	26,2—31,9	25,4—32,8
30	21,2—40,0	18,9—43,1	23,8—36,9	21,9—39,0	26,0—34,2	24,8—35,5	27,2—32,9	26,3—33,9
31	22,1—41,0	19,8—44,1	*24,7—38,0*	*22,8—40,1*	*27,0—35,3*	*25,8—36,6*	28,1—34,0	27,3—34,9
32	23,0—42,1	20,6—45,2	*25,6—39,0*	*23,7—41,1*	27,9—36,3	26,7—37,6	29,1—35,0	28,2—35,9
33	23,9—43,1	21,5—46,2	*26,6—40,0*	*24,7—42,2*	*28,9—37,3*	*27,7—38,7*	30,1—36,0	29,2—36,9
34	24,8—44,2	22,3—47,3	*27,5—41,0*	*25,6—43,2*	29,9—38,3	28,6—39,7	31,1—37,0	30,2—38,0
35	25,7—45,2	23,2—48,3	28,4—42,1	26,5—44,2	*30,8—39,4*	*29,6—40,7*	32,1—38,1	31,1—39,0
36	26,6—46,2	24,1—49,3	*29,4—43,1*	*27,4—45,2*	31,8—40,4	30,5—41,7	33,0—39,1	32,1—40,0
37	27,6—47,2	25,0—50,3	*30,4—44,1*	*28,4—46,2*	*32,7—41,4*	*31,5—42,8*	34,0—40,1	33,1—41,0
38	28,5—48,3	25,8—51,3	*31,3—45,1*	*29,3—47,2*	33,7—42,4	32,5—43,8	35,0—41,1	34,1—42,1
39	29,4—49,3	26,7—52,4	*32,3—46,2*	*30,3—48,2*	*34,7—43,4*	*33,4—44,8*	36,0—42,1	35,0—43,1
40	30,3—50,3	27,6—53,3	33,2—47,2	31,2—49,3	35,7—44,5	34,4—45,8	36,9—43,1	36,0—44,1
41	31,3—51,3	28,5—54,3	*34,2—48,2*	*32,1—50,3*	*36,7—45,5*	*35,3—46,8*	37,9—44,1	37,0—45,1
42	32,2—52,3	29,4—55,3	*35,2—49,2*	*33,1—51,3*	37,6—46,5	36,3—47,8	38,9—45,1	38,0—46,1
43	33,1—53,3	30,3—56,3	*36,1—50,2*	*34,0—52,3*	*38,6—47,5*	*37,3—48,9*	39,9—46,1	39,0—47,1
44	34,1—54,3	31,3—57,3	*37,1—51,2*	*35,0—53,3*	39,6—48,5	38,3—49,9	40,9—47,1	39,9—48,1
45	35,0—55,3	32,2—58,3	38,1—52,2	35,9—54,4	*40,6—49,5*	*39,3—50,9*	41,9—48,1	40,9—49,1
46	36,0—56,3	33,1—59,3	*39,1—53,2*	*36,9—55,4*	41,6—50,5	40,2—51,9	42,9—49,1	41,9—50,1
47	37,0—57,2	34,1—60,2	*40,0—54,2*	*37,8—56,4*	*42,6—51,5*	*41,2—52,9*	43,9—50,1	42,9—51,1
48	37,9—58,2	35,0—61,2	*40,9—55,1*	*38,8—57,4*	43,6—52,5	42,2—53,9	44,9—51,1	43,9—52,1
49	38,9—59,2	35,9—62,2	*41,9—56,1*	*39,7—58,3*	*44,5—53,5*	*43,2—54,8*	45,9—52,1	44,9—53,1
50	39,8—60,2	36,9—63,1	42,9—57,1	40,7—59,3	45,5—54,5	44,2—55,8	46,9—53,1	45,9—54,1

*) zie tabel R–2

Ontleend aan GEIGY (1960), met uitzondering van de cursief gedrukte getallen die door de Afd. Statistiek van het N.I.P.G. zijn berekend.

TABEL S. DE TOETS VAN KOLMOGOROV-SMIRNOV
Kritieke waarden van D bij tweezijdige toetsing

n	Onbetrouwbaarheidsdrempel α				
	0,20	0,15	0,10	0,05	0,01
1	0,900	0,925	0,950	0,975	0,995
2	0,684	0,726	0,776	0,842	0,929
3	0,565	0,597	0,642	0,708	0,828
4	0,494	0,525	0,564	0,624	0,733
5	0,446	0,474	0,510	0,565	0,669
6	0,410	0,436	0,470	0,521	0,618
7	0,381	0,405	0,438	0,486	0,577
8	0,358	0,381	0,411	0,457	0,543
9	0,339	0,360	0,388	0,432	0,514
10	0,322	0,342	0,368	0,410	0,490
11	0,307	0,326	0,352	0,391	0,468
12	0,259	0,313	0,338	0,375	0,450
13	0,284	0,302	0,325	0,361	0,433
14	0,274	0,292	0,314	0,349	0,418
15	0,266	0,283	0,304	0,338	0,404
16	0,258	0,274	0,295	0,328	0,392
17	0,250	0,266	0,286	0,318	0,381
18	0,244	0,259	0,278	0,309	0,371
19	0,237	0,252	0,272	0,301	0,363
20	0,231	0,246	0,264	0,294	0,356
25	0,21	0,22	0,24	0,27	0,32
30	0,19	0,20	0,22	0,24	0,29
35	0,18	0,19	0,21	0,23	0,27
> 35	$\frac{1,07}{\sqrt{n}}$	$\frac{1,14}{\sqrt{n}}$	$\frac{1,22}{\sqrt{n}}$	$\frac{1,36}{\sqrt{n}}$	$\frac{1,63}{\sqrt{n}}$

Naar MASSEY (1951).

TABEL T. DE NORMALITEITSTOETSEN VAN FISHER

Linker en rechter kritieke waarden van de grootheden g_1 en g_2 bij éénzijdige toetsing

n	g_1		g_2			
	Rechter kritieke waarden*		Linker kritieke waarden		Rechter kritieke waarden	
	0,05	0,01	0,01	0,05	0,05	0,01
50	0,550	0,812	—	—	—	—
75	0,454	0,664	—	—	—	—
100	0,395	0,576	—0,80	—0,62	0,87	1,53
125	0,354	0,514	—0,74	—0,57	0,78	1,34
150	0,324	0,469	—0,69	—0,53	0,71	1,22
175	0,301	0,434	—0,66	—0,50	0,66	1,11
200	0,282	0,406	—0,62	—0,47	0,62	1,04
250	0,253	0,362	—0,57	—0,44	0,55	0,91
300	0,231	0,331	—0,53	—0,40	0,50	0,82
350	0,214	0,306	—0,49	—0,37	0,47	0,75
400	0,201	0,286	—0,46	—0,35	0,43	0,69
450	0,189	0,270	—0,44	—0,33	0,40	0,65
500	0,180	0,256	—0,42	—0,32	0,38	0,62
550	0,171	0,244	—0,41	—0,30	0,37	0,58
600	0,163	0,234	—0,39	—0,29	0,35	0,55
650	0,157	0,225	—0,38	—0,28	0,33	0,53
700	0,151	0,215	—0,37	—0,27	0,32	0,51
750	0,146	0,208	—0,35	—0,26	0,31	0,49
800	0,142	0,202	—0,34	—0,25	0,30	0,47
850	0,138	0,196	—0,33	—0,25	0,29	0,46
900	0,134	0,190	—0,33	—0,24	0,29	0,44
950	0,130	0,185	—0,32	—0,23	0,28	0,43
1000	0,127	0,180	—0,31	—0,23	0,27	0,42

Naar BENNETT & FRANKLIN (1954).

*De linker kritieke waarden van g_1 zijn numeriek gelijk aan de rechter kritieke waarden, maar zij bezitten een tegengesteld teken.

TABEL U. DE UITBIJTERTOETS VAN DIXON

Rechter kritieke waarden van de grootheid r

Toetsings-grootheid	n	Onbetrouwbaarheidsdrempel α				
		0,10	0,05	0,025	0,01	0,005
r_{10}	3	0,886	0,941	0,97	0,988	0,994
	4	0,679	0,765	0,83	0,889	0,926
	5	0,557	0,642	0,71	0,780	0,821
	6	0,482	0,560	0,63	0,698	0,740
	7	0,434	0,507	0,57	0,637	0,680
	8	0,399	0,468	0,53	0,590	0,634
	9	0,370	0,437	0,50	0,555	0,598
	10	0,349	0,412	0,47	0,527	0,568
r_{22}	11	0,578	0,637	0,69	0,745	0,781
	12	0,543	0,600	0,66	0,704	0,740
	13	0,515	0,570	0,62	0,670	0,705
	14	0,492	0,546	0,60	0,641	0,674
	15	0,472	0,525	0,57	0,616	0,647
	16	0,454	0,507	0,55	0,595	0,624
	17	0,438	0,490	0,54	0,577	0,605
	18	0,424	0,475	0,52	0,561	0,589
	19	0,412	0,462	0,51	0,547	0,575
	20	0,401	0,450	0,50	0,535	0,562
	21	0,391	0,440	0,49	0,524	0,551
	22	0,382	0,430	0,48	0,514	0,541
	23	0,374	0,421	0,47	0,505	0,532
	24	0,367	0,413	0,46	0,497	0,524
	25	0,360	0,406	0,45	0,489	0,516
	26	0,354	0,399	0,445	0,482	0,508
	27	0,348	0,393	0,44	0,475	0,501
	28	0,342	0,387	0,43	0,469	0,495
	29	0,337	0,381	0,43	0,463	0,489
	30	0,332	0,376	0,42	0,457	0,483

Uit DIXON (1951). De kritieke waarden voor $\alpha = 0{,}025$ worden door DIXON niet gegeven; zij zijn door grafische interpolatie verkregen en zodoende slechts in twee decimalen opgegeven.

Tabel V. De bgsin transformatie

1. Waarden van de grootheid $y = 2\text{bgsin}\sqrt{p}$

p	y	p	y	p	y	p	y	p	y
0,001	0,0633	0,051	0,4556	0,101	0,6468	0,151	0,7982	0,201	0,9298
0,002	0,0895	0,052	0,4601	0,102	0,6501	0,152	0,8010	0,202	0,9323
0,003	0,1096	0,053	0,4646	0,103	0,6534	0,153	0,8038	0,203	0,9348
0,004	0,1266	0,054	0,4690	0,104	0,6567	0,154	0,8065	0,204	0,9373
0,005	0,1415	0,055	0,4734	0,105	0,6600	0,155	0,8093	0,205	0,9397
0,006	0,1551	0,056	0,4778	0,106	0,6632	0,156	0,8121	0,206	0,9422
0,007	0,1675	0,057	0,4822	0,107	0,6665	0,157	0,8148	0,207	0,9447
0,008	0,1791	0,058	0,4864	0,108	0,6697	0,158	0,8176	0,208	0,9471
0,009	0,1900	0,059	0,4907	0,109	0,6729	0,159	0,8203	0,209	0,9496
0,010	0,2003	0,060	0,4949	0,110	0,6761	0,160	0,8230	0,210	0,9521
0,011	0,2101	0,061	0,4991	0,111	0,6793	0,161	0,8258	0,211	0,9545
0,012	0,2195	0,062	0,5033	0,112	0,6825	0,162	0,8285	0,212	0,9570
0,013	0,2285	0,063	0,5074	0,113	0,6857	0,163	0,8312	0,213	0,9594
0,014	0,2372	0,064	0,5115	0,114	0,6888	0,164	0,8339	0,214	0,9619
0,015	0,2456	0,065	0,5156	0,115	0,6920	0,165	0,8366	0,215	0,9643
0,016	0,2537	0,066	0,5196	0,116	0,6951	0,166	0,8393	0,216	0,9667
0,017	0,2615	0,067	0,5236	0,117	0,6982	0,167	0,8420	0,217	0,9692
0,018	0,2691	0,068	0,5276	0,118	0,7013	0,168	0,8446	0,218	0,9716
0,019	0,2766	0,069	0,5316	0,119	0,7044	0,169	0,8473	0,219	0,9740
0,020	0,2838	0,070	0,5355	0,120	0,7075	0,170	0,8500	0,220	0,9764
0,021	0,2909	0,071	0,5394	0,121	0,7106	0,171	0,8526	0,221	0,9788
0,022	0,2978	0,072	0,5433	0,122	0,7136	0,172	0,8553	0,222	0,9812
0,023	0,3045	0,073	0,5472	0,123	0,7167	0,173	0,8579	0,223	0,9836
0,024	0,3111	0,074	0,5510	0,124	0,7197	0,174	0,8606	0,224	0,9860
0,025	0,3176	0,075	0,5548	0,125	0,7227	0,175	0,8632	0,225	0,9884
0,026	0,3239	0,076	0,5586	0,126	0,7258	0,176	0,8658	0,226	0,9908
0,027	0,3301	0,077	0,5624	0,127	0,7288	0,177	0,8685	0,227	0,9932
0,028	0,3362	0,078	0,5661	0,128	0,7318	0,178	0,8711	0,228	0,9956
0,029	0,3423	0,079	0,5698	0,129	0,7347	0,179	0,8737	0,229	0,9980
0,030	0,3482	0,080	0,5735	0,130	0,7377	0,180	0,8763	0,230	1,0004
0,031	0,3540	0,081	0,5772	0,131	0,7407	0,181	0,8789	0,231	1,0027
0,032	0,3597	0,082	0,5808	0,132	0,7437	0,182	0,8815	0,232	1,0051
0,033	0,3653	0,083	0,5845	0,133	0,7466	0,183	0,8841	0,233	1,0075
0,034	0,3709	0,084	0,5881	0,134	0,7495	0,184	0,8867	0,234	1,0098
0,035	0,3764	0,085	0,5917	0,135	0,7525	0,185	0,8892	0,235	1,0122
0,036	0,3818	0,086	0,5953	0,136	0,7554	0,186	0,8918	0,236	1,0146
0,037	0,3871	0,087	0,5988	0,137	0,7583	0,187	0,8944	0,237	1,0169
0,038	0,3924	0,088	0,6024	0,138	0,7612	0,188	0,8969	0,238	1,0193
0,039	0,3976	0,089	0,6059	0,139	0,7641	0,189	0,8995	0,239	1,0216
0,040	0,4027	0,090	0,6094	0,140	0,7670	0,190	0,9021	0,240	1,0239
0,041	0,4078	0,091	0,6129	0,141	0,7699	0,191	0,9046	0,241	1,0263
0,042	0,4128	0,092	0,6163	0,142	0,7727	0,192	0,9071	0,242	1,0286
0,043	0,4178	0,093	0,6198	0,143	0,7756	0,193	0,9097	0,243	1,0310
0,044	0,4227	0,094	0,6232	0,144	0,7785	0,194	0,9122	0,244	1,0333
0,045	0,4275	0,095	0,6266	0,145	0,7813	0,195	0,9147	0,245	1,0356
0,046	0,4323	0,096	0,6300	0,146	0,7841	0,196	0,9173	0,246	1,0379
0,047	0,4371	0,097	0,6334	0,147	0,7870	0,197	0,9198	0,247	1,0403
0,048	0,4418	0,098	0,6368	0,148	0,7898	0,198	0,9223	0,248	1,0426
0,049	0,4464	0,099	0,6402	0,149	0,7926	0,199	0,9248	0,249	1,0449
0,050	0,4510	0,100	0,6435	0,150	0,7954	0,200	0,9273	0,250	1,0472

TABEL V. DE BGSIN TRANSFORMATIE (*vervolg*)

p	y	p	y	p	y	p	y	p	y
0,251	1,0495	0,301	1,1615	0,351	1,2682	0,401	1,3715	0,451	1,4726
0,252	1,0518	0,302	1,1636	0,352	1,2703	0,402	1,3735	0,452	1,4746
0,253	1,0541	0,303	1,1658	0,353	1,2724	0,403	1,3756	0,453	1,4767
0,254	1,0564	0,304	1,1680	0,354	1,2745	0,404	1,3776	0,454	1,4787
0,255	1,0587	0,305	1,1702	0,355	1,2766	0,405	1,3796	0,455	1,4807
0,256	1,0610	0,306	1,1723	0,356	1,2787	0,406	1,3817	0,456	1,4827
0,257	1,0633	0,307	1,1745	0,357	1,2807	0,407	1,3837	0,457	1,4847
0,258	1,0656	0,308	1,1767	0,358	1,2828	0,408	1,3857	0,458	1,4867
0,259	1,0679	0,309	1,1788	0,359	1,2849	0,409	1,3878	0,459	1,4887
0,260	1,0701	0,310	1,1810	0,360	1,2870	0,410	1,3898	0,460	1,4907
0,261	1,0724	0,311	1,1832	0,361	1,2891	0,411	1,3918	0,461	1,4927
0,262	1,0747	0,312	1,1853	0,362	1,2912	0,412	1,3939	0,462	1,4947
0,263	1,0770	0,313	1,1875	0,363	1,2932	0,413	1,3959	0,463	1,4967
0,264	1,0792	0,314	1,1896	0,364	1,2953	0,414	1,3979	0,464	1,4987
0,265	1,0815	0,315	1,1918	0,365	1,2974	0,415	1,4000	0,465	1,5007
0,266	1,0838	0,316	1,1939	0,366	1,2995	0,416	1,4020	0,466	1,5027
0,267	1,0860	0,317	1,1961	0,367	1,3016	0,417	1,4040	0,467	1,5048
0,268	1,0883	0,318	1,1982	0,368	1,3036	0,418	1,4061	0,468	1,5068
0,269	1,0905	0,319	1,2004	0,369	1,3057	0,419	1,4081	0,469	1,5088
0,270	1,0928	0,320	1,2025	0,370	1,3078	0,420	1,4101	0,470	1,5108
0,271	1,0951	0,321	1,2047	0,371	1,3098	0,421	1,4121	0,471	1,5128
0,272	1,0973	0,322	1,2068	0,372	1,3119	0,422	1,4142	0,472	1,5148
0,273	1,0995	0,323	1,2090	0,373	1,3140	0,423	1,4162	0,473	1,5168
0,274	1,1018	0,324	1,2111	0,374	1,3161	0,424	1,4182	0,474	1,5188
0,275	1,1040	0,325	1,2132	0,375	1,3181	0,425	1,4202	0,475	1,5208
0,276	1,1063	0,326	1,2154	0,376	1,3202	0,426	1,4222	0,476	1,5228
0,277	1,1085	0,327	1,2175	0,377	1,3222	0,427	1,4243	0,477	1,5248
0,278	1,1107	0,328	1,2196	0,378	1,3243	0,428	1,4263	0,478	1,5268
0,279	1,1130	0,329	1,2217	0,379	1,3264	0,429	1,4283	0,479	1,5288
0,280	1,1152	0,330	1,2239	0,380	1,3284	0,430	1,4303	0,480	1,5308
0,281	1,1174	0,331	1,2260	0,381	1,3305	0,431	1,4324	0,481	1,5328
0,282	1,1196	0,332	1,2281	0,382	1,3325	0,432	1,4344	0,482	1,5348
0,283	1,1219	0,333	1,2303	0,383	1,3346	0,433	1,4364	0,483	1,5368
0,284	1,1241	0,334	1,2324	0,384	1,3367	0,434	1,4384	0,484	1,5388
0,285	1,1263	0,335	1,2345	0,385	1,3387	0,435	1,4404	0,485	1,5408
0,286	1,1285	0,336	1,2366	0,386	1,3408	0,436	1,4424	0,486	1,5428
0,287	1,1307	0,337	1,2387	0,387	1,3428	0,437	1,4445	0,487	1,5448
0,288	1,1329	0,338	1,2408	0,388	1,3449	0,438	1,4465	0,488	1,5468
0,289	1,1351	0,339	1,2430	0,389	1,3469	0,439	1,4485	0,489	1,5488
0,290	1,1373	0,340	1,2451	0,390	1,3490	0,440	1,4505	0,490	1,5508
0,291	1,1396	0,341	1,2472	0,391	1,3510	0,441	1,4525	0,491	1,5528
0,292	1,1418	0,342	1,2493	0,392	1,3531	0,442	1,4545	0,492	1,5548
0,293	1,1440	0,343	1,2514	0,393	1,3551	0,443	1,4565	0,493	1,5568
0,294	1,1461	0,344	1,2535	0,394	1,3572	0,444	1,4586	0,494	1,5588
0,295	1,1483	0,345	1,2556	0,395	1,3592	0,445	1,4606	0,495	1,5608
0,296	1,1505	0,346	1,2577	0,396	1,3613	0,446	1,4626	0,496	1,5628
0,297	1,1527	0,347	1,2598	0,397	1,3633	0,447	1,4646	0,497	1,5648
0,298	1,1549	0,348	1,2619	0,398	1,3654	0,448	1,4666	0,498	1,5668
0,299	1,1571	0,349	1,2640	0,399	1,3674	0,449	1,4686	0,499	1,5688
0,300	1,1593	0,350	1,2661	0,400	1,3694	0,450	1,4706	0,500	1,5708

Uit HALD (1952).

Voor $0{,}501 < p < 0{,}999$ kan men nemen: $y = 3{,}1416 - 2 \text{ bgsin } \sqrt{1-p}$.
Voorbeeld: $p = 0{,}881$, $y = 3{,}1416 - 2 \text{ bsgin } \sqrt{0{,}119} = 3{,}1416 - 0{,}7044 = 2{,}4372$.

2. Waarden van y (0) bij $p = 0$ en van y (1) bij $p = 1$ voor $10 < n < 50$

n	y (0)	y (1)	n	y (0)	y (1)	n	y (0)	y (1)	n	y (0)	y (1)
10	0,3176	2,8240	20	0,2241	2,9175	30	0,1828	2,9588	40	0,1583	2,9833
11	0,3027	2,8389	21	0,2187	2,9229	31	0,1798	2,9617	41	0,1563	2,9853
12	0,2897	2,8519	22	0,2136	2,9280	32	0,1770	2,9646	42	0,1545	2,9871
13	0,2782	2,8633	23	0,2089	2,9327	33	0,1743	2,9673	43	0,1526	2,9889
14	0,2681	2,8735	24	0,2045	2,9371	34	0,1717	2,9699	44	0,1509	2,9907
15	0,2589	2,8827	25	0,2003	2,9413	35	0,1692	2,9724	45	0,1492	2,9924
16	0,2507	2,8909	26	0,1964	2,9452	36	0,1669	2,9747	46	0,1476	2,9940
17	0,2431	2,8985	27	0,1927	2,9488	37	0,1646	2,9770	47	0,1460	2,9956
18	0,2363	2,9053	28	0,1893	2,9523	38	0,1624	2,9792	48	0,1445	2,9971
19	0,2299	2,9117	29	0,1860	2,9556	39	0,1603	2,9813	49	0,1430	2,9986
									50	0,1415	3,0001

Uit EISENHART (1947).

TABEL W. DE VIERKANTSWORTEL TRANSFORMATIE VAN FREEMAN EN TUKEY

Waarden van $y = \sqrt{x} + \sqrt{x+1}$ en van $y^2 = (\sqrt{x} + \sqrt{x+1})^2$

x	y	y^2	x	y	y^2
0	1,00	1,0000	25	10,10	102,0100
1	2,41	5,8081	26	10,30	106,0900
2	3,15	9,9225	27	10,49	110,0401
3	3,73	13,9129	28	10,68	114,0624
4	4,24	17,9776	29	10,86	117,9396
5	4,69	21,9961	30	11,04	121,8816
6	5,10	26,0100	31	11,22	125,8884
7	5,47	29,9209	32	11,40	129,9600
8	5,83	33,9889	33	11,58	134,0964
9	6,16	37,9456	34	11,75	138,0625
10	6,48	41,9904	35	11,92	142,0864
11	6,78	45,9684	36	12,08	145,9264
12	7,07	49,9849	37	12,25	150,0625
13	7,35	54,0225	38	12,41	154,0081
14	7,61	57,9121	39	12,57	158,0049
15	7,87	61,9369	40	12,73	162,0529
16	8,12	65,9344	41	12,88	165,8944
17	8,37	70,0569	42	13,04	170,0416
18	8,60	73,9600	43	13,19	173,9761
19	8,83	77,9689	44	13,34	177,9556
20	9,05	81,9025	45	13,49	181,9801
21	9,27	85,9329	46	13,64	186,0496
22	9,49	90,0601	47	13,78	189,8884
23	9,69	93,8961	48	13,93	194,0449
24	9,90	98,0100	49	14,07	197,9649
			50	14,21	201,9241

TABEL X. MODELPOPULATIE

(Pseudo-normaal verdeeld, $\mu = 20$, $\sigma = 3{,}470$)

Nr.	Waarde	Nr.	Waarde	Nr.	Waarde	Nr.	Waarde	Nr.	Waarde
1	22	41	22	81	20	121	23	161	19
2	21	42	18	82	21	122	21	162	22
3	23	43	22	83	23	123	18	163	20
4	16	44	21	84	19	124	21	164	16
5	20	45	23	85	18	125	16	165	21
6	18	46	21	86	23	126	23	166	18
7	19	47	16	87	21	127	20	167	21
8	15	48	22	88	25	128	19	168	15
9	22	49	19	89	20	129	24	169	21
10	17	50	24	90	21	130	19	170	18
11	20	51	20	91	17	131	22	171	25
12	24	52	18	92	22	132	14	172	20
13	18	53	10	93	18	133	20	173	17
14	13	54	22	94	24	134	18	174	27
15	21	55	16	95	20	135	15	175	23
16	15	56	13	96	17	136	22	176	19
17	18	57	20	97	21	137	16	177	16
18	25	58	17	98	13	138	21	178	21
19	17	59	22	99	18	139	23	179	26
20	23	60	26	100	16	140	17	180	20
21	20	61	15	101	16	141	27	181	22
22	24	62	17	102	20	142	18	182	15
23	19	63	24	103	14	143	16	183	19
24	11	64	18	104	22	144	20	184	29
25	23	65	16	105	17	145	18	185	24
26	20	66	19	106	25	146	12	186	20
27	15	67	17	107	19	147	22	187	14
28	18	68	27	108	30	148	19	188	24
29	22	69	21	109	20	149	25	189	18
30	26	70	24	110	15	150	21	190	28
31	24	71	22	111	26	151	17	191	17
32	19	72	25	112	19	152	21	192	25
33	23	73	17	113	16	153	17	193	18
34	20	74	19	114	22	154	22	194	17
35	22	75	14	115	24	155	14	195	20
36	26	76	23	116	23	156	19	196	19
37	17	77	22	117	20	157	23	197	23
38	19	78	20	118	25	158	21	198	21
39	19	79	18	119	19	159	19	199	24
40	23	80	21	120	18	160	21	200	20
								201	19

De verdeling is ontleend aan DIXON en MASSEY (1957).

TABEL Y. ASELECTE GETALLEN (TOEVALCIJFERS)

	1	2	3	4	5
1	0347437386	3696473661	4698637162	3326168045	6011141095
2	9774246762	4281145720	4253323732	2707360751	2451798973
3	1676622766	5650267107	3290797853	1355385859	8897541410
4	1256859926	9696682731	0503729315	5712101421	8826498176
5	5559563564	3854824622	3162430990	0618443253	2383013030
6	1622779439	4954435482	1737932378	8735209643	8426349164
7	8442175331	5724550688	7704744767	2176335025	8392120676
8	6301637859	1695556719	9810507175	1286735807	4439523879
9	3321123429	7864560782	5242074438	1551001342	9966027954
10	5760863244	0947279654	4917460962	9052847727	0802734328
11	1818079246	4417165809	7983861962	0676500310	5523640505
12	2662389775	8416074499	8311463224	2014858845	1093728871
13	2342406474	8297777781	0745321408	3298940772	9385791075
14	5236281995	5092261197	0056763138	8022025353	8660420453
15	3785943512	8339500830	4234079688	5442068798	3585294839
16	7029171213	4033203826	1389510374	1776371304	0774211930
17	5662183735	9683508775	9712259347	7033240354	9777464480
18	9949572277	8842954572	1664361600	0443186679	9477242190
19	1608150472	3327143409	4559346849	1272073445	9927729514
20	3116933243	5027898719	2015370049	5285666044	3868881180
21	6834301370	5574307740	4422788426	0433460952	6807970657
22	7457256576	5929976860	7191386754	1358182476	1554559552
23	2742378653	4855906572	9657693610	9646924245	9760490491
24	0039682961	6637322030	7784570329	1045650426	1104966724
25	2994989424	6849691082	5375919330	3425205727	4048735192
26	1690826659	8362641112	6719007174	6047212968	0202370331
27	1127947506	0609197466	0294373402	7670903086	3845943038
28	3524101620	3332512638	7978450491	1692535616	0275509598
29	3823168638	4238970150	8775668141	4001749162	4851840832
30	3196259147	9644334913	3486825391	0052434885	2755268962
31	6667406714	6405719586	1105650968	7683203790	5716001166
32	1490844511	7573880590	5227411486	2298122208	0752749580
33	6805511800	3396027519	0760629355	5933824390	4937384459
34	2046787390	9751401402	0402333108	3954164936	4795931330
35	6419589779	1506159320	0190107506	4078788962	0267741733
36	0526937060	2235851513	9203515977	5956780683	5291057074
37	0797108823	0998429964	6171629915	0651291693	5805770951
38	6871868585	5487664754	7332081112	4495926316	2956242948
39	2699616553	5837788070	4210506742	3217558574	9444671694
40	1465526875	8759362241	2678630655	1308270150	1529393943
41	1753775871	7141615072	1241949626	4495273699	0296743083
42	9026592119	2352233312	9693021839	0702183607	2599327023
43	4123525599	3104496996	1047484588	1341438920	9717144917
44	6020508169	3199736868	3581330376	2430124860	1899107234
45	9125380590	9458284136	4537590309	9035572912	8262546560
46	3450577437	9880330091	0977931982	7494800404	4507316649
47	8522043943	7381539479	3362468628	0831544631	5394133847
48	0979137748	7382972221	0503272483	7289440560	3580399488
49	8875801814	2295754249	3932822249	0248077037	1604616787
50	9096237000	3900030690	5585783836	9437306932	9089007633

Ontleend aan FISHER en YATES (1953).

Tabel Z. Aselecte permutaties (lotingsgetallen)

1 — 1 t/m 10

Serie 1	3	6	10	2	8	Serie 2	8	10	3	4	7	Serie 3	3	1	5	7	4
	9	1	4	5	7		6	5	1	9	2		8	10	9	2	6
Serie 4	9	3	4	10	6	Serie 5	7	3	5	10	9	Serie 6	9	10	8	1	3
	2	7	1	5	8		2	8	4	1	6		2	7	4	6	5

2 — 1 t/m 20

	06	16	09	07	19		17	01	08	13	07		08	06	19	13	03
Serie	03	18	12	11	01	Serie	10	04	05	09	14	Serie	05	11	02	01	18
1	20	08	13	17	14	2	02	19	15	03	06	3	20	04	15	07	09
	05	15	10	02	04		12	18	11	20	16		12	14	16	17	10
	20	09	10	14	11		06	04	16	15	08		05	03	09	16	06
Serie	05	19	03	06	04	Serie	19	13	12	17	03	Serie	02	17	08	15	13
4	17	18	13	02	15	5	10	18	14	02	09	6	11	10	01	04	14
	07	08	01	12	16		07	20	01	05	11		19	07	20	12	18

3 — 1 t/m 50

	42	16	32	40	17		37	21	38	30	19		22	38	27	02	35
	36	03	23	41	10		02	43	01	04	20		20	37	30	26	45
	34	25	07	47	44		24	45	50	11	27		43	47	12	17	36
	24	29	30	04	22		47	32	16	36	25		44	15	25	09	31
Serie	13	26	11	37	50	Serie	07	17	41	44	23	Serie	41	06	40	49	10
1	31	02	33	06	12	2	31	34	12	42	29	3	39	08	19	28	05
	45	05	14	27	21		18	08	49	03	40		01	18	21	34	24
	01	15	20	19	38		33	15	39	28	05		23	29	11	42	07
	35	43	39	49	28		35	14	09	26	13		50	48	32	16	04
	46	18	48	08	09		48	46	22	06	10		14	03	33	46	13

4 — 1 t/m 100

	81	74	67	95	70		03	16	15	40	71		84	05	56	07	91
	61	37	42	62	93		77	07	86	46	61		26	98	74	12	45
	52	07	16	29	39		92	67	78	72	66		80	32	73	54	69
	43	08	77	25	72		43	75	91	02	04		79	86	46	57	99
	65	32	27	40	63		69	95	29	53	39		09	100	48	36	06
	21	58	11	23	80		83	13	99	94	52		04	76	59	44	89
	31	90	55	88	13		18	37	42	44	55		90	96	68	65	67
	73	98	20	05	68		20	97	21	22	58		24	52	29	70	11
	33	15	35	26	79		89	80	79	10	41		27	50	60	97	37
Serie	75	66	99	09	06	Serie	26	60	45	01	11	Serie	66	40	49	85	33
1	56	51	54	50	53	2	35	19	62	09	34	3	95	63	01	42	30
	96	34	18	22	89		25	48	74	38	87		19	23	82	14	55
	04	71	14	76	78		12	14	17	65	76		16	41	71	77	75
	49	86	03	83	45		93	30	85	59	96		35	94	28	21	51
	57	97	84	82	87		88	08	64	81	28		64	93	18	20	25
	10	30	01	100	44		84	27	73	05	32		13	58	72	08	47
	36	24	91	19	64		23	54	51	98	63		61	88	62	22	31
	46	69	85	94	59		57	68	24	90	33		34	53	10	03	17
	92	38	12	41	17		31	70	36	47	06		87	78	81	15	43
	47	48	60	28	02		82	100	49	56	50		92	39	02	83	38

Uit Statistische Tabellen en Nomogrammen.

TABEL AA. RECHTER KRITIEKE WAARDEN VAN DE GROOTHEID $\underline{w}$

$\alpha = 0{,}05$

ν \ k	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	18,0	27,0	32,8	37,1	40,4	43,1	45,4	47,4	49,1	50,6	52,0	53,2	54,3	55,4	56,3	57,2	58,0	58,8	59,6
2	6,09	8,28	9,80	10,89	11,73	12,43	13,03	13,54	13,99	14,39	14,75	15,08	15,38	15,65	15,91	16,14	16,36	16,57	16,77
3	4,50	5,91	6,82	7,50	8,04	8,48	8,85	9,18	9,46	9,72	9,95	10,15	10,35	10,52	10,69	10,84	10,98	11,11	11,24
4	3,93	5,04	5,76	6,29	6,71	7,05	7,35	7,60	7,83	8,03	8,21	8,37	8,52	8,66	8,79	8,91	9,03	9,13	9,23
5	3,64	4,60	5,22	5,67	6,03	6,33	6,58	6,80	6,99	7,17	7,32	7,47	7,60	7,72	7,83	7,93	8,03	8,12	8,21
6	3,46	4,34	4,90	5,31	5,63	5,89	6,12	6,32	6,49	6,65	6,79	6,92	7,03	7,14	7,24	7,34	7,43	7,51	7,59
7	3,34	4,16	4,68	5,06	5,35	5,61	5,82	6,00	6,16	6,30	6,43	6,55	6,66	6,76	6,85	6,94	7,02	7,09	7,17
8	3,26	4,04	4,53	4,89	5,17	5,40	5,60	5,77	5,92	6,05	6,18	6,29	6,39	6,48	6,57	6,65	6,73	6,80	6,87
9	3,20	3,95	4,42	4,76	5,02	5,24	5,43	5,60	5,74	5,87	5,98	6,09	6,19	6,28	6,36	6,44	6,51	6,58	6,64
10	3,15	3,88	4,33	4,65	4,91	5,12	5,30	5,46	5,60	5,72	5,83	5,93	6,03	6,11	6,20	6,27	6,34	6,40	6,47
11	3,11	3,82	4,26	4,58	4,82	5,03	5,20	5,35	5,49	5,61	5,71	5,81	5,90	5,98	6,06	6,14	6,20	6,26	6,33
12	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,40	5,51	5,62	5,71	5,80	5,88	5,95	6,03	6,09	6,15	6,21
13	3,06	3,73	4,15	4,46	4,69	4,88	5,05	5,19	5,32	5,43	5,53	5,63	5,71	5,79	5,86	5,93	6,00	6,05	6,11
14	3,03	3,70	4,11	4,41	4,64	4,83	4,99	5,13	5,25	5,36	5,46	5,55	5,64	5,72	5,79	5,85	5,92	5,97	6,03
15	3,01	3,67	4,08	4,37	4,60	4,78	4,94	5,08	5,20	5,31	5,40	5,49	5,58	5,65	5,72	5,79	5,85	5,90	5,96
16	3,00	3,65	4,05	4,34	4,56	4,74	4,90	5,03	5,15	5,26	5,35	5,44	5,52	5,59	5,66	5,72	5,79	5,84	5,90
17	2,98	3,63	4,02	4,31	4,52	4,71	4,86	4,99	5,11	5,21	5,31	5,39	5,47	5,55	5,61	5,68	5,74	5,79	5,84
18	2,97	3,61	4,00	4,28	4,49	4,67	4,82	4,96	5,07	5,17	5,27	5,35	5,43	5,50	5,57	5,63	5,69	5,74	5,79
19	2,96	3,59	3,98	4,26	4,47	4,65	4,79	4,92	5,04	5,14	5,23	5,32	5,39	5,46	5,53	5,59	5,65	5,70	5,75
20	2,95	3,58	3,96	4,24	4,45	4,62	4,77	4,90	5,01	5,11	5,20	5,28	5,36	5,43	5,49	5,55	5,61	5,66	5,71
24	2,92	3,53	3,90	4,17	4,37	4,54	4,68	4,81	4,92	5,01	5,10	5,18	5,25	5,32	5,38	5,44	5,50	5,54	5,59
30	2,89	3,49	3,84	4,11	4,30	4,46	4,60	4,72	4,83	4,92	5,00	5,08	5,15	5,21	5,27	5,33	5,38	5,43	5,48
40	2,86	3,44	3,79	4,04	4,23	4,39	4,52	4,63	4,74	4,82	4,91	4,98	5,05	5,11	5,16	5,22	5,27	5,31	5,36
60	2,83	3,40	3,74	3,98	4,16	4,31	4,44	4,55	4,65	4,73	4,81	4,88	4,94	5,00	5,06	5,11	5,16	5,20	5,24
120	2,80	3,36	3,69	3,92	4,10	4,24	4,36	4,48	4,56	4,64	4,72	4,78	4,84	4,90	4,95	5,00	5,04	5,09	5,13
∞	2,77	3,31	3,63	3,86	4,03	4,17	4,29	4,39	4,47	4,55	4,62	4,68	4,74	4,80	4,85	4,89	4,93	4,97	5,01

TABEL AA (*vervolg*)
$\alpha = 0{,}01$

ν \ k	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	90,0	134	164	186	202	216	227	237	246	253	260	266	272	277	282	286	290	294	298
2	14,0	18,9	22,3	24,7	26,6	28,2	29,5	30,7	31,7	32,6	33,4	34,1	34,8	35,4	36,0	36,5	37,0	37,5	37,9
3	8,26	10,56	12,17	13,34	14,25	15,00	15,65	16,20	16,69	17,13	17,53	17,89	18,23	18,54	18,83	19,09	19,33	19,56	19,79
4	6,51	8,12	9,17	9,96	10,58	11,10	11,55	11,93	12,26	12,56	12,84	13,09	13,32	13,53	13,73	13,92	14,09	14,25	14,40
5	5,70	6,97	7,80	8,42	8,91	9,32	9,67	9,97	10,24	10,48	10,70	10,89	11,08	11,24	11,40	11,55	11,68	11,81	11,93
6	5,24	6,33	7,03	7,56	7,97	8,32	8,61	8,87	9,10	9,30	9,49	9,65	9,81	9,95	10,08	10,21	10,32	10,43	10,54
7	4,95	5,92	6,54	7,01	7,37	7,68	7,94	8,17	8,37	8,55	8,71	8,86	9,00	9,12	9,24	9,35	9,46	9,55	9,65
8	4,74	5,63	6,20	6,63	6,96	7,24	7,47	7,68	7,87	8,03	8,18	8,31	8,44	8,55	8,66	8,76	8,85	8,94	9,03
9	4,60	5,43	5,96	6,35	6,66	6,91	7,13	7,32	7,49	7,65	7,78	7,91	8,03	8,13	8,23	8,32	8,41	8,49	8,57
10	4,48	5,27	5,77	6,14	6,43	6,67	6,87	7,05	7,21	7,36	7,48	7,60	7,71	7,81	7,91	7,99	8,07	8,15	8,22
11	4,39	5,14	5,62	5,97	6,25	6,48	6,67	6,84	6,99	7,13	7,25	7,36	7,46	7,56	7,65	7,73	7,81	7,88	7,95
12	4,32	5,04	5,50	5,84	6,10	6,32	6,51	6,67	6,81	6,94	7,06	7,17	7,26	7,36	7,44	7,52	7,59	7,66	7,73
13	4,26	4,96	5,40	5,73	5,98	6,19	6,37	6,53	6,67	6,79	6,90	7,01	7,10	7,19	7,27	7,34	7,42	7,48	7,55
14	4,21	4,89	5,32	5,63	5,88	6,08	6,26	6,41	6,54	6,66	6,77	6,87	6,96	7,05	7,12	7,20	7,27	7,33	7,39
15	4,17	4,83	5,25	5,56	5,80	5,99	6,16	6,31	6,44	6,55	6,66	6,76	6,84	6,93	7,00	7,07	7,14	7,20	7,26
16	4,13	4,78	5,19	5,49	5,72	5,92	6,08	6,22	6,35	6,46	6,56	6,66	6,74	6,82	6,90	6,97	7,03	7,09	7,15
17	4,10	4,74	5,14	5,43	5,66	5,85	6,01	6,15	6,27	6,38	6,48	6,57	6,66	6,73	6,80	6,87	6,94	7,00	7,05
18	4,07	4,70	5,09	5,38	5,60	5,79	5,94	6,08	6,20	6,31	6,41	6,50	6,58	6,65	6,72	6,79	6,85	6,91	6,96
19	4,05	4,67	5,05	5,33	5,55	5,73	5,89	6,02	6,14	6,25	6,34	6,43	6,51	6,58	6,65	6,72	6,78	6,84	6,89
20	4,02	4,64	5,02	5,29	5,51	5,69	5,84	5,97	6,09	6,19	6,29	6,37	6,45	6,52	6,59	6,65	6,71	6,76	6,82
24	3,96	4,54	4,91	5,17	5,37	5,54	5,69	5,81	5,92	6,02	6,11	6,19	6,26	6,33	6,39	6,45	6,51	6,56	6,61
30	3,89	4,45	4,80	5,05	5,24	5,40	5,54	5,65	5,67	5,85	5,93	6,01	6,08	6,14	6,20	6,26	6,31	6,36	6,41
40	3,82	4,37	4,70	4,93	5,11	5,27	5,39	5,50	5,60	5,69	5,77	5,84	5,90	5,96	6,02	6,07	6,12	6,17	6,21
60	3,76	4,28	4,60	4,82	4,99	5,13	5,25	5,36	5,45	5,53	5,60	5,67	5,73	5,79	5,84	5,89	5,93	5,98	6,02
120	3,70	4,20	4,50	4,71	4,87	5,01	5,12	5,21	5,30	5,38	5,44	5,51	5,56	5,61	5,66	5,71	5,75	5,79	5,83
∞	3,64	4,12	4,40	4,60	4,76	4,88	4,99	5,08	5,16	5,23	5,29	5,35	5,40	5,45	5,49	5,54	5,57	5,61	5,65

Uit MAY (1952).

TABEL BB

Rankits in aselecte steekproeven van n elementen uit een standaardnormaal verdeelde populatie

Rangnr.	n = 6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1,27	1,35	1,42	1,49	1,54	1,59	1,63	1,67	1,70	1,74	1,76	1,79	1,82	1,84	1,87
2	0,64	0,76	0,85	0,93	1,00	1,06	1,12	1,16	1,21	1,25	1,28	1,32	1,35	1,38	1,41
3	0,20	0,35	0,47	0,57	0,66	0,73	0,79	0,85	0,90	0,95	0,99	1,03	1,07	1,10	1,13
4		0,00	0,15	0,27	0,38	0,46	0,54	0,60	0,66	0,71	0,76	0,81	0,85	0,89	0,92
5				0,00	0,12	0,22	0,31	0,39	0,46	0,52	0,57	0,62	0,67	0,71	0,75
6						0,00	0,10	0,19	0,27	0,34	0,39	0,45	0,50	0,55	0,59
7								0,00	0,09	0,17	0,23	0,30	0,35	0,40	0,45
8										0,00	0,08	0,15	0,21	0,26	0,31
9												0,00	0,07	0,13	0,19
10														0,00	0,06

Rangnr.	n = 21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	1,89	1,91	1,93	1,95	1,97	1,98	2,00	2,01	2,03	2,04	2,06	2,07	2,08	2,09	2,11
2	1,43	1,46	1,48	1,50	1,52	1,54	1,56	1,58	1,60	1,62	1,63	1,65	1,66	1,68	1,69
3	1,16	1,19	1,21	1,24	1,26	1,29	1,31	1,33	1,35	1,36	1,38	1,40	1,42	1,43	1,45
4	0,95	0,98	1,01	1,04	1,07	1,09	1,11	1,14	1,16	1,18	1,20	1,22	1,23	1,25	1,27
5	0,78	0,82	0,85	0,88	0,91	0,93	0,96	0,98	1,00	1,03	1,05	1,07	1,09	1,11	1,12
6	0,63	0,67	0,70	0,73	0,76	0,79	0,82	0,85	0,87	0,89	0,92	0,94	0,96	0,98	1,00
7	0,49	0,53	0,57	0,60	0,64	0,67	0,70	0,73	0,75	0,78	0,80	0,82	0,85	0,87	0,89
8	0,36	0,41	0,45	0,48	0,52	0,55	0,58	0,61	0,64	0,67	0,69	0,72	0,74	0,76	0,79
9	0,24	0,29	0,33	0,37	0,41	0,44	0,48	0,51	0,54	0,57	0,60	0,62	0,65	0,67	0,69
10	0,12	0,17	0,22	0,26	0,30	0,34	0,38	0,41	0,44	0,47	0,50	0,53	0,56	0,58	0,60
11	0,00	0,06	0,11	0,16	0,20	0,24	0,28	0,32	0,35	0,38	0,41	0,44	0,47	0,50	0,52
12			0,00	0,05	0,10	0,14	0,19	0,22	0,26	0,29	0,33	0,36	0,39	0,41	0,44
13					0,00	0,05	0,09	0,13	0,17	0,21	0,24	0,28	0,31	0,34	0,36
14							0,00	0,04	0,09	0,12	0,16	0,20	0,23	0,26	0,29
15									0,00	0,04	0,08	0,12	0,15	0,18	0,22
16											0,00	0,04	0,08	0,11	0,14
17													0,00	0,04	0,07
18															0,00

Rangnr.	n = 36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
1	2,12	2,13	2,14	2,15	2,16	2,17	2,18	2,19	2,20	2,21	2,22	2,22	2,23	2,24	2,25
2	1,70	1,72	1,73	1,74	1,75	1,76	1,78	1,79	1,80	1,81	1,82	1,83	1,84	1,85	1,85
3	1,46	1,48	1,49	1,50	1,52	1,53	1,54	1,55	1,57	1,58	1,59	1,60	1,61	1,62	1,63
4	1,28	1,30	1,32	1,33	1,34	1,36	1,37	1,38	1,40	1,41	1,42	1,43	1,44	1,45	1,46
5	1,14	1,16	1,17	1,19	1,20	1,22	1,23	1,25	1,26	1,27	1,28	1,30	1,31	1,32	1,33
6	1,02	1,03	1,05	1,07	1,08	1,10	1,11	1,13	1,14	1,16	1,17	1,18	1,19	1,21	1,22
7	0,91	0,92	0,94	0,96	0,98	0,99	1,01	1,02	1,04	1,05	1,07	1,08	1,09	1,11	1,12
8	0,81	0,83	0,85	0,86	0,88	0,90	0,91	0,93	0,95	0,96	0,98	0,99	1,00	1,02	1,03
9	0,72	0,73	0,75	0,77	0,79	0,81	0,83	0,84	0,86	0,88	0,89	0,91	0,92	0,94	0,95
10	0,63	0,65	0,67	0,69	0,71	0,73	0,75	0,76	0,78	0,80	0,81	0,83	0,84	0,86	0,87
11	0,54	0,57	0,59	0,61	0,63	0,65	0,67	0,69	0,71	0,72	0,74	0,76	0,77	0,79	0,80
12	0,47	0,49	0,51	0,54	0,56	0,58	0,60	0,62	0,64	0,65	0,67	0,69	0,70	0,72	0,74
13	0,39	0,42	0,44	0,46	0,49	0,51	0,53	0,55	0,57	0,59	0,60	0,62	0,64	0,66	0,67
14	0,32	0,34	0,37	0,39	0,42	0,44	0,46	0,48	0,50	0,52	0,54	0,56	0,58	0,59	0,61
15	0,24	0,27	0,30	0,33	0,35	0,37	0,40	0,42	0,44	0,46	0,48	0,50	0,52	0,53	0,55
16	0,17	0,20	0,23	0,26	0,28	0,31	0,33	0,36	0,38	0,40	0,42	0,44	0,46	0,48	0,49
17	0,10	0,14	0,16	0,19	0,22	0,25	0,27	0,29	0,32	0,34	0,36	0,38	0,40	0,42	0,44
18	0,03	0,07	0,10	0,13	0,16	0,18	0,21	0,23	0,26	0,28	0,30	0,32	0,34	0,36	0,38
19		0,00	0,03	0,06	0,09	0,12	0,15	0,17	0,20	0,22	0,25	0,27	0,29	0,31	0,33
20				0,00	0,03	0,06	0,09	0,12	0,14	0,17	0,19	0,21	0,24	0,26	0,28
21						0,00	0,03	0,06	0,09	0,11	0,14	0,16	0,18	0,21	0,23
22								0,00	0,03	0,06	0,08	0,11	0,13	0,15	0,18
23										0,00	0,03	0,05	0,08	0,10	0,13
24												0,00	0,03	0,05	0,08
25														0,00	0,03

Uit FISHER en YATES (1953).

TABEL CC. BETROUWBAARHEIDSGRENZEN VAN DE MEDIAAN

n	(>) 95%			(>) 99%		
	k	$n-k+1$	$1-\alpha$ (exact)	k	$n-k+1$	$1-\alpha$ (exact)
6	1	6	0,9844	—	—	—
7	1	7	0,9922	—	—	—
8	1	8	0,9961	1	8	0,9961
9	2	8	0,9805	1	9	0,9980
10	2	9	0,9893	1	10	0,9990
11	2	10	0,9941	1	11	0,9995
12	3	10	0,9807	2	11	0,9968
13	3	11	0,9888	2	12	0,9983
14	3	12	0,9935	2	13	0,9991
15	4	12	0,9824	3	13	0,9963
16	4	13	0,9894	3	14	0,9979
17	5	13	0,9755	3	15	0,9988
18	5	14	0,9846	4	15	0,9962
19	5	15	0,9904	4	16	0,9978
20	6	15	0,9793	4	17	0,9987
21	6	16	0,9867	5	17	0,9964
22	6	17	0,9915	5	18	0,9978
23	7	17	0,9827	5	19	0,9987
24	7	18	0,9887	6	19	0,9967
25	8	18	0,9784	6	20	0,9980
26	8	19	0,9855	7	20	0,9953
27	8	20	0,9904	7	21	0,9970
28	9	20	0,9822	7	22	0,9981
29	9	21	0,9879	8	22	0,9959
30	10	21	0,9786	8	23	0,9974
31	10	22	0,9853	8	24	0,9987
32	10	23	0,9900	9	24	0,9965
33	11	23	0,9825	9	25	0,9977
34	11	24	0,9878	10	25	0,9955
35	12	24	0,9795	10	26	0,9970
40	14	27	0,9808	12	29	0,9968
45	16	30	0,9882	14	32	0,9967
50	18	33	0,9836	16	35	0,9967
55	20	36	0,9850	18	38	0,9968
60	22	39	0,9863	20	41	0,9969
65	25	41	0,9768	22	44	0,9970
70	27	44	0,9793	24	47	0,9972
75	29	47	0,9815	26	50	0,9974

Uit Sankhya: Indian Journal of Statistics, vol. 4, part 4, 556–557 (1940).

BIJLAGE II

ENKELE BEGRIPPEN UIT DE ELEMENTAIRE ALGEBRA

1. Het gebruik van symbolen

Vijf waarnemingen aan een grootheid x verricht kunnen door het toekennen van indices worden onderscheiden: x_1, x_2, x_3, x_4 en x_5. Zijn n waarnemingen verricht, dan schrijft men $x_1, x_2, \ldots, x_n$ of x_i $(i = 1, 2, \ldots, n)$. De som van deze waarnemingen is: $x_1 + x_2 + \ldots + x_n$. Men kan dit compacter schrijven door gebruik te maken van het *somteken* Σ:

$$x_1 + x_2 + \ldots + x_n = \sum_{i=1}^{n} x_i.$$

De index onder het somteken geeft aan, dat de sommering begint bij x_1; die er boven deelt mede, dat moet worden gesommeerd tot en met x_n. Als men bijv. slechts de eerste vier waarden van x wenst te sommeren, schrijft men: $\sum_{i=1}^{4} x_i$.

Deze notatie kan worden vereenvoudigd door af te spreken dat de indices (gedeeltelijk) worden weggelaten, als men over alle n waarden sommeert, zodat men kan schrijven:

$$\sum_{i=1}^{n} x_i = \sum_{i} x_i = \Sigma x.$$

Voorbeelden

1. $\sum_{i=4}^{7} x_i = x_4 + x_5 + x_6 + x_7$

2. $\sum_{i=3}^{5} x_i y_i = x_3 y_3 + x_4 y_4 + x_5 y_5$

3. $\sum_{i=5}^{8} a x_i = a x_5 + a x_6 + a x_7 + a x_8 = a(x_5 + x_6 + x_7 + x_8) = a \sum_{i=5}^{8} x_i$

4. $\sum_{i=2}^{4}(x_i - 2) = (x_2 - 2) + (x_3 - 2) + (x_4 - 2) = x_2 + x_3 + x_4 - 3(2)$

$$= \sum_{i=2}^{4} x_i - 6$$

5. $\sum_{i=2}^{4} x_i - 2 = x_2 + x_3 + x_4 - 2$

6. $\Sigma fx = \sum_{i=1}^{n} f_i x_i = f_1 x_1 + f_2 x_2 + \ldots\ldots + f_n x_n$

7. $\sum_{i=1}^{3}(x_i + y_i) = (x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3)$

$$= (x_1 + x_2 + x_3) + (y_1 + y_2 + y_3) = \sum_{i=1}^{3} x_i + \sum_{i=1}^{3} y_i$$

8. $\sum_{i=1}^{n}(x_i - a)^2 = \Sigma(x_i^2 - 2ax_i + a^2) = \Sigma x_i^2 - 2a \,.\, \Sigma x_i + na^2$

Als een grootheid de waarden $x_i (i = 1, 2, \ldots, k)$ heeft aangenomen, noemt men het aantal malen dat x_i voorkomt, de *frequentie* van deze waarde: symbool f_i. Zijn er in totaal n waarnemingen verricht dan geldt dus:

$$f_1 + f_2 + \ldots + f_k = \sum_i f_i = \Sigma f = n.$$

Voor de som van de n waarden kan men schrijven $\sum_{i=1}^{k} f_i x_i$, of alleen Σfx.

Voorbeeld

x_i	2	3	4	5	6	7	8	
f_i	1	—	3	7	5	3	2	$\Sigma f = n = 21$
$f_i x_i$	2	—	12	35	30	21	16	$\Sigma fx = 116$

Tenslotte vermelden wij enkele tekens, die veel gebruikt worden:

$=$: gelijk aan	$\leqslant$: ten hoogste gelijk aan	$\neq$: ongelijk aan
$<$: kleiner dan	$\geqslant$: ten minste gelijk aan	$\to$: nadert tot
$>$: groter dan	$\approx$: bij benadering gelijk aan	∞ : oneindig

OPGAVEN

1. Geef onderstaande sommen een compacter vorm:

a. $x_8 + x_9 + x_{10} + x_{11} + x_{12}$ *b.* $4x_2 + 4x_3 + 4x_4 + 4x_5$

c. $ax_3y_3 + ax_4y_4 + ax_5y_5$ *d.* $(x_1 - a) + (x_2 - a) + (x_3 - a) + (x_4 - a)$

2. Men beschikt over de volgende reeks waarnemingen:

Proefpersoon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x	2	4	6	3	2	3	4	4	7	1	6	5	4	5	3

Bepaal de numerieke waarde van:

$a.\ \Sigma x_i$ $b.\ \sum_{i=1}^{5}(x_i-2)$ $c.\ \frac{\Sigma x}{15}$ $d.\ \sum_{i=1}^{5} x_i - 2$ $e.\ \sum_{i=1}^{5}(x_i-2)^2$

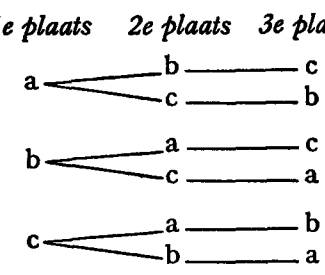
$f.\ \sum_{i=3}^{7} x_i/5$ $g.\ (\Sigma x_i)^2$ $h.\ \Sigma x_i^2$ $i.\ \sum_{i=1}^{8} 2(x_i-1)^2$

2. Permutaties en combinaties

PERMUTATIES

Zoals uit nevenstaand schema blijkt, kunnen de letters a, b en c op $3 \times 2 \times 1 = 6$ verschillende manieren worden gerangschikt. Men noemt deze rangschikkingen permutaties en het is duidelijk, dat het aantal permutaties van n objecten gelijk is aan

(1) $\quad n(n-1)\,(n-2)\,\ldots\,(3)\,(2)\,(1).$



Men kan (1) korter schrijven door gebruik te maken van de 'afkorting' $n!$ (*n faculteit*). Dus: $8! = 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 40320$. Per definitie stelt men: $0! = 1$.

Vervolgens beschouwen wij het aantal permutaties dat uit n objecten kan worden gevormd als per permutatie slechts $x\ (< n)$ van de objecten worden gebruikt. Men kan dan de eerste plaats op n manieren bezetten, de tweede op $n - 1$ manieren, enz. Als men de plaats x bereikt zijn $x-1$ objecten gebruikt, zodat er nog $n - (x-1) = n-x+1$ voor keuze in aanmerking komen. Als wij het gevraagde aantal permutaties P_x^n noemen geldt dus

(2) $\quad P_x^n = n(n-1)\,(n-2)\ldots(n-x+1) = \frac{n!}{(n-x)!},$

daar $n! = n(n-1)\,(n-2)\,\ldots\,(n-x+1)\,(n-x)!$ Als wij $x = n$ stellen gaat (2) over in (1) : $P_n^n = \frac{n!}{0!} = n!$

OPGAVEN

1. Schrijf systematisch alle permutaties van 3 letters uit de letters a, b, c en d op.
2. Verifieer, dat:

$a.\ 6! = 720$ $b.\ 20! = 20 \times 19 \times 18!$ $c.\ \frac{(x+3)!}{x!} = (x+3)(x+2)(x+1)$

$d.\ P_2^5 = 20$ $e.\ P_5^5 = 5!$ $f.\ \frac{10!}{7!} = 720$ $g.\ \frac{10!6!4!}{3!12!} = \frac{240}{11}$

3. Welke waarde heeft a in de volgende vergelijkingen?

$a.\ (n+1)! = a(n-1)!$ $b.\ 72(a!) = 9!$ $c.\ \frac{a!}{120} = 1$

$d.\ a(a+1)\,(6!) = 8!$ $e.\ 14! + a(14!) = 15!$ $f.\ 6a = 7! - a$

COMBINATIES

Met behulp van vergelijking (2) kunnen wij nagaan op hoeveel verschillende manieren x objecten uit de n objecten kunnen worden gekozen. Hierbij gaat het dus om *combinaties* van x objecten waarbij hun rangschikking niet ter zake doet. Twee combinaties zijn verschillend, indien zij niet uit dezelfde objecten bestaan. Dus: *abc* en *abd* zijn verschillende combinaties van drie letters; *abc* en *bca* zijn verschillende permutaties van dezelfde combinatie.

Voor het aanduiden van het aantal combinaties van x objecten uit n objecten zijn verschillende symbolen in gebruik, zoals

$$\binom{n}{x},\ {}_nC_x,\ C(n, x) \text{ en } C_x^n.$$

Wij gebruiken steeds de laatstgenoemde notatie.

Het is duidelijk, dat $P_x^n = x! \,.\, C_x^n$, daar elke combinatie van x objecten op $x!$ manieren gepermuteerd kan worden. Hieruit volgt

$$(3) \qquad C_x^n = \frac{P_x^n}{x!} = \frac{\overbrace{(n-x)!}^{n}}{x!} = \frac{n!}{x!(n-x)!} = C_{n-x}^n.$$

Het aantal combinaties van 4 letters, dat men uit 7 letters kan vormen, bedraagt bijv.: $C_4^7 = \frac{7!}{4!3!} = \frac{7 \times 6 \times 5}{3 \times 2 \times 1} = 35.$

Met (3) kunnen vraagstukjes worden opgelost van de volgende aard: Uit een groep van 8 mannen en 5 vrouwen wordt een commissie van 5 personen gevormd. Hierin dienen twee vrouwen zitting te hebben. Op hoeveel manieren kan deze commissie worden samengesteld?

Er dienen twee handelingen verricht te worden: 1. Het kiezen van 2 uit de 5 vrouwen, 2. Het kiezen van de andere 3 leden uit de 8 mannen. Het aantal manier, waarop men de 2 vrouwen kan kiezen is volgens (3): $C_2^5 = \frac{5!}{2!3!} = \frac{5 \times 4}{2 \times 1} = 10$. De drie mannen kunnen op $C_3^8 = \frac{8!}{3!5!} = \frac{8 \times 7 \times 6}{3 \times 2 \times 1} = 56$ manieren gekozen worden. Elke combinatie van 2 vrouwen vormt met elke combinatie van 3 mannen een verschillende combinatie van 5 personen, zodat de gevraagde samenstelling op $10 \times 56 = 560$ manieren tot stand kan komen.

Beschikt men dus over N_1 objecten van soort A en N_2 objecten van soort B, dan bedraagt het aantal manieren waarop men een groep (combinatie) van n objecten, bestaande uit x objecten A en $n - x$ objecten B, kan samenstellen

$$(4) \qquad C_x^{N_1} \,.\, C_{n-x}^{N_2}.$$

OPGAVEN

1. Verifieer, dat:
   a. $C_4^9 = 126$ *b.* $C_3^{25} = 2300$ *c.* $C_{20}^{20} = 1$ *d.* $C_4^{10} \cdot C_6^{10} = (210)^2$
2. Betreft het voorgaande voorbeeld inzake de commissiesamenstelling: *a.* Op hoeveel manieren kan de commissie worden samengesteld, als daarin ten minste 3 mannen zitting moeten hebben? *b.* Als omtrent de samenstelling niets wordt voorgeschreven?
3. Een groep kinderen bestaat uit 8 jongens en 6 meisjes. Er kunnen 5 kinderen naar een vacantie-kolonie worden gezonden. Op hoeveel verschillende manieren kunnen deze worden gekozen, als er ten hoogste 3 jongens kunnen worden uitgezonden?
4. Bij een voetbalpool dient men van 15 wedstrijden aan te geven of de thuisclub wint, verliest of gelijkspeelt. Hoeveel oplossingen moet men indienen om er zeker van te zijn, dat de juiste uitkomst der 15 gespeelde wedstrijden zich daarbij bevindt?

3. Exponenten en wortels

Het product of gedurig product van gelijke factoren noemt men een macht. Het getal dat aangeeft uit hoeveel gelijke factoren een macht bestaat, heet de *exponent*; de factor noemt men het *grondtal* der macht. Van de macht a^n is dus a het grondtal en n de exponent. Per definitie stelt men: $a^0 = 1$ en $a^{-n} = \frac{1}{a^n}$. Men noemt $\frac{1}{a} = a^{-1}$ de *reciproke* van a.

Een getal b, dat tot de macht n verheven een getal a oplevert, noemt men een wortel uit a. Daar $2^2 = 4$ is 2 de tweedemachtswortel uit 4. Daar $3^3 = 27$ is 3 de derdemachtswortel uit 27, enz. In het algemeen: als $b^n = a$, is b de n^{de} machtswortel uit a. De tweedemachtswortel uit een getal noemt men in de regel kortweg de wortel uit dat getal. Men schrijft $\sqrt[n]{a}$ ook wel als $a^{\frac{1}{n}}$ en $\sqrt[n]{a^m}$ als $a^{\frac{m}{n}}$.

Wij brengen verder de volgende eigenschappen van exponenten in herinnering:

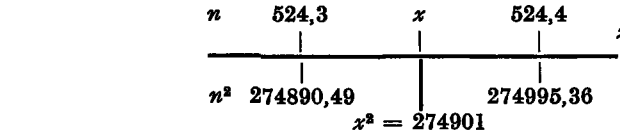
1. $(a^m)(a^n) = a^{m+n}$ . . . . . bijv.: $10^2 \times 10^5 = 10^7$.
2. $(a^m)^n = a^{mn}$ . . . . . bijv.: $(10^2)^5 = 10^{10}$.
3. $(ab)^n = a^n \cdot b^n$. . . . . bijv.: $(10 \times 4)^2 = 10^2 \cdot 4^2$.
4. $\frac{a^m}{a^n} = a^{m-n}$. . . . . bijv.: $\frac{3^5}{3^2} = 3^3$.

Bij het uitvoeren van statistische berekeningen heeft men vaak de kwadraten, (tweedemachts)wortels en reciproken van getallen nodig. Voor de getallen van 1 t/m 12.500 kan men deze opzoeken in de tafel van BARLOW (1954). Als men de wortel wil trekken uit een getal, dat niet in de tafel voorkomt, zoals bijv. het getal 274901, kan deze tafel toch worden gebruikt. Men zoekt dan het getal op, waarvan 274901

het kwadraat is. In de kwadratenkolom van de tafel staat:

$$n = 524\,3 \ldots\ldots n^2 = 27\ 48\ 90\ 49,$$
$$n = 524\,4 \ldots\ldots n^2 = 27\ 49\ 95\ 36.$$

Afgerond op één decimaal is de gewenste wortel dus 524,3. Zo nodig kan de volgende decimaal door interpolatie worden verkregen:



$$x = 524{,}3 + \frac{274901 - 274890{,}49}{274995{,}36 - 274890{,}49}\ 0{,}1$$
$$= 524{,}3 + (0{,}102)\ (0{,}1) = 524{,}31.$$

Als men over een rekenmachine beschikt kan het worteltrekken eenvoudig plaatsvinden door herhaald te delen. Bijv.: $\sqrt{78{,}393316} = ?$ Het is duidelijk, dat de gevraagde wortel tussen 8 en 9 ligt. Kies nu 8,5:

$$\frac{78{,}393316}{8{,}5} = 9{,}2 \qquad \frac{8{,}5 + 9{,}2}{2} = 8{,}85 \qquad \frac{78{,}393316}{8{,}85} = 8{,}858$$

$$\frac{8{,}85 + 8{,}858}{2} = 8{,}854 \qquad \frac{78{,}393316}{8{,}854} = 8{,}854.$$

Dus $\sqrt{78{,}393316} = 8{,}854$.

OPGAVEN

1. Schrijf elk der volgende vormen als een product van de getallen 2, 3, 5 en 7:
   a. $20^3 \cdot 45^3$ *b.* $(24 \cdot 36)^4$ *c.* $18^2 \cdot 15^3 \cdot 42^2$ *d.* $(16 \cdot 9 \cdot 18)^3$
2. Geef gelijkwaardige uitdrukkingen, waarin alle exponenten positief zijn, voor:
   a. $\dfrac{a^3b^{-5}c^2d}{a^{-2}b^2c^5d^{-2}}$ *b.* $\dfrac{pq^{-4}r^{-2}s}{s^2t^{-5}r}$ *c.* $\dfrac{(ab)^{-4}c^3d^{-5}e^2}{a^2b^{-3}c^{-2}(de)^2}$ *d.* $\dfrac{x^{-1}y^0w^{-3}z^{-4}}{y^2x^{-2}w^{-5}z^7}$
3. Schrijf elk der volgende uitdrukkingen als een wortel:
   a. $x^{\frac{3}{4}}$ *b.* $(c+d)^{\frac{1}{2}}$ *c.* $y^{\frac{c}{d}}$ *d.* $b^{\frac{a+c}{d}}$ *e.* $3^{\frac{2}{2}}$
4. Trek de vierkantswortel uit onderstaande getallen:
   a. 0,00064 *b.* 4000 *c.* 284865,20 *d.* $(2^3)^4\ (3^2)^2\ (5)^{-4}$
   e. 268,6321 *f.* 0,002106 *g.* 1399 *h.* 0,07213

4. Logarithmen

Definitie: De logarithme van een getal a is de exponent, die aangeeft tot welke macht men een aangenomen grondtal g moet verheffen, om het getal a te verkrijgen. Dus:

$$g^{\log_g a} = a. \tag{5}$$

Alle getallen met hun logarithmen, die tot een zelfde grondtal of basis behoren, vormen tezamen een logarithmenstelsel.

Logarithmen met grondtal 10 worden veel gebruikt. In dat geval geldt dus

(6) $$10^{\log_{10} a} = a.$$

Men laat bij deze logarithmen de index 10 meestal achterwege, zodat log a steeds betekent: $\log_{10} a$ (men schrijft ook vaak ${}^{10}\log a$). Op grond van (6) kunnen de logarithmen van getallen, die een macht van 10 zijn direct worden opgeschreven, zodat bijv.:

$$\begin{array}{lll} \log 1000 & = \log 10^3 & = 3 \\ \log 100 & = \log 10^2 & = 2 \\ \log 10 & = \log 10^1 & = 1 \\ \log 1 & = \log 10^0 & = 0 \\ \log 0{,}1 & = \log 10^{-1} & = -1 \\ \log 0{,}01 & = \log 10^{-2} & = -2 \\ \log 0{,}001 & = \log 10^{-3} & = -3 \end{array}$$

Negatieve getallen en nul bezitten geen logarithme.

In het algemeen bestaan logarithmen uit een aantal gehelen (de *wijzer*) en een onmeetbaar getal (de *mantisse*). De wijzer kan direct worden bepaald (zie boven): a. Voor een getal groter dan 1 is de wijzer één minder dan het aantal cijfers, dat voor het decimaalteken is geplaatst, b. Voor een getal kleiner dan 1 is de wijzer negatief en numeriek één groter dan het aantal nullen, dat achter het decimaalteken staat. De wijzer van 863 is dus 2, de wijzer van 0,00863 is —3. De mantisse kan uit een tafel worden afgelezen. Verifieer met een logarithmentafel, dat in 5 decimalen nauwkeurig:

$\log 863 = 2{,}93601$, $\log 0{,}00863 = 0{,}93601 - 3$,
$\log 3000 = 3{,}47712$, $\log 3 = 0{,}47712$, $\log 0{,}3 = 0{,}47712 - 1$.

Negatieve logarithmen worden, evenals de positieve, zo geschreven dat de decimalen de mantisse voorstellen; voor —3,51268 schrijft men dus 0,48732 — 4.

Wij vermelden de volgende eigenschappen van logarithmen, die uit de definitie volgen (deze gelden voor elk logarithmenstelsel, maar de voorbeelden betreffen het stelsel met grondtal 10):

1. $\log (a)\,(b) = \log a + \log b$ ..... $\log (5)\,(10) = \log 5 + \log 10 = 0{,}69897 + 1 = 1{,}69897 = \log 50.$
2. $\log \frac{a}{b} = \log a - \log b$ ..... $\log \frac{10}{5} = \log 10 - \log 5 = 1 - 0{,}69897 = 0{,}30103 = \log 2.$
3. $\log a^b = b \log a$ ..... $\log 10^2 = 2 \log 10 = 2 \times 1 = 2 = \log 100.$
4. $\log \frac{1}{a^b} = \log a^{-b} = -b \log a$ ... $\log 10^{-2} = -2 \log 10 = -2 = \log 0{,}01.$
5. $\log \sqrt[b]{a} = \log a^{\frac{1}{b}} = \frac{\log a}{b}$ ..... $\log \sqrt[2]{100} = \frac{\log 100}{2} = 1 = \log 10.$

Als voorbeeld geven wij enkele berekeningen, die met behulp van logarithmen zijn uitgevoerd:

$(6{,}803)\ (0{,}04362)\ (0{,}512) = ?$ $\qquad \sqrt[3]{7922} = ?$ $\qquad (2{,}91)^4 = ?$

$\log 6{,}803 = 0{,}83270$	$\frac{\log 7922}{3} = \frac{3{,}89883}{3}$	$4 \log 2{,}91 = 4(0{,}46389)$
$\log 0{,}04362 = 0{,}63969 - 2$	$= 1{,}29961$	$= 1{,}85556$
$\log 0{,}512 = 0{,}70927 - 1$		
$\log \text{product} = 2{,}18166 - 3$	wortel $= 19{,}935$	vierde macht $= 71{,}71$
$= 0{,}18166 - 1$		
product $= 0{,}15194$		

Logarithmen met grondtal e

Deze zg. *natuurlijke* logarithmen hebben als grondtal $e = 2{,}718282...$, d.i. de limiet waartoe $(1 + \frac{1}{n})^n$ nadert, als n nadert tot oneindig. Zij worden in de wiskunde veel gebruikt en aangeduid met *ln*, dus $ln\ a = \log_e a$ of ${}^e\log a$. Natuurlijke logarithmen kunnen eveneens in een tafel worden opgezocht of uit logarithmen met grondtal 10 worden afgeleid volgens:

(7) $ln\ a = \log_e a = (\log_e 10)\ (\log_{10} a) = 2{,}3026 \log_{10} a = 2{,}3026 \log a.$

OPGAVEN

1. Verifieer met een logarithmentafel, dat in 5 decimalen nauwkeurig:
   a. $\log 9{,}94 = 0{,}99739$ $\qquad$ *c.* $\log 8510 = 3{,}92993$
   b. $\log 0{,}531 = 0{,}72509 - 1$ $\qquad$ *d.* $\log 0{,}002645 = 0{,}42243 - 3$
2. Bepaal de logarithme van:
   a. 12,37 $\qquad$ *c.* 0,001152 $\qquad$ *e.* 20000 $\qquad$ *g.* 0,0016
   b. 4,03 $\qquad$ *d.* 0,2354 $\qquad$ *f.* 5375 $\qquad$ *h.* 0,00001357
3. Bepaal het getal, waarvan de logarithme is:
   a. 1,18469 $\qquad$ *c.* $0{,}82138 - 1$ $\qquad$ *e.* 0,14208 $\qquad$ *g.* $8{,}74273 - 10$
   b. 3,62552 $\qquad$ *d.* $6{,}74772 - 10$ $\qquad$ *f.* 2,99996 $\qquad$ *h.* 0,00432
4. Voer onderstaande berekeningen uit met behulp van logarithmen:
   a. $\dfrac{37{,}62 \times 4{,}921}{682{,}3}$ $\qquad$ *c.* $\dfrac{52{,}34}{(0{,}631)^2}$
   b. $(68{,}19)^{\frac{1}{6}}$ $\qquad$ *d.* $(3{,}32)\ (61{,}54)$ $\qquad$ *e.* $\left[\dfrac{9{,}723 \times 8{,}46}{0{,}041}\right]^{\frac{1}{2}}$

5. Benaderende getallen

HET AANTAL BETROUWBARE CIJFERS

Bij een discreet verdeelde grootheid worden de waarnemingsuitkomsten door tellen verkregen, zodat deze grootheid slechts gehele waarden kan aannemen. Elke uitkomst is in principe exact vast te stellen en het verschil tussen twee uitkomsten bedraagt dus steeds een of meer eenheden.

Bij een continu verdeelde grootheid worden de uitkomsten door

meten of wegen verkregen. Zij kunnen een continuum van waarden doorlopen en theoretisch kan het verschil tussen twee uitkomsten willekeurig klein worden. In de praktijk verkrijgt men echter steeds *benaderende* uitkomsten, waarvan de nauwkeurigheid afhankelijk is van de ter beschikking staande apparatuur en van de eisen, die de onderzoeker stelt. Meet men bijv. de lichaamslengte van kinderen, dan kan dit nog wel in mm nauwkeurig geschieden, maar dit zal bijzondere zorg en dus extra tijd vereisen, terwijl deze nauwkeurigheid zelden noodzakelijk zal zijn. Als men de uitkomsten van zo'n meting nu in cm nauwkeurig geeft, behoeven twee kinderen met dezelfde uitkomst dus niet even lang te zijn: meestal betekent de uitkomst $x = 132$ cm immers $131{,}5 < x < 132{,}5$, zodat bijv. zowel 131,6 cm als 132,4 cm als 132 cm genoteerd worden. Men zegt dan dat de uitkomst 132 cm *drie betrouwbare cijfers* bevat en het is duidelijk, dat dit aantal niet verandert als men een andere maateenheid gebruikt en bijv. opgeeft 1,32 m of 0,00132 km. In het laatstgenoemde geval dienen de nullen uitsluitend om de plaats van het decimaalteken vast te leggen, zodat zij niet meetellen in het aantal betrouwbare cijfers. Bij getallen als 5,09 of 60,05 tellen de nullen vanzelfsprekend daarin wel mee, zodat zij 3, resp. 4 betrouwbare cijfers bevatten. Het getal 8,5 bezit twee, maar 8,50 maakt aanspraak op drie betrouwbare cijfers.

AFRONDEN

Bij het afronden houdt men zich meestal aan de volgende regels:

a. Als het eerste cijfer dat vervalt kleiner is dan 5, blijft het daaraan voorafgaande cijfer onveranderd. Dus:
8,4644, afgerond op 4 betrouwbare cijfers, wordt 8,464.

b. Als het eerste cijfer dat vervalt groter is dan 5 of als dit een 5 is, gevolgd door cijfers die niet alle nul zijn, wordt het voorafgaande cijfer met 1 vermeerderd. Dus:
24,467, afgerond op 4 betrouwbare cijfers, wordt 24,47
0,61503, afgerond op 2 betrouwbare cijfers, wordt 0,62.

c. Als het cijfer dat vervalt een 5 is, of een 5 is en gevolgd wordt door uitsluitend nullen, volgt men in de regel de volgende (arbitraire) procedure:

1. Breng geen verandering in het voorgaande cijfer aan, als het even is,
2. Verhoog het met 1, als het oneven is (een nul wordt als even gerekend).

Dus:
26,05, afgerond op 3 betrouwbare cijfers, wordt 26,0
0,875, afgerond op 2 betrouwbare cijfers, wordt 0,88.

HET REKENEN MET BENADERENDE CIJFERS

Zonder naar volledigheid te streven geven wij enkele aanwijzingen voor het rekenen met benaderende getallen.

a. Optellen en aftrekken

Hierbij heeft men te maken met getallen, die alle in dezelfde eenheid zijn uitgedrukt. Het criterium voor de betrouwbaarheid is in dit geval dus het aantal betrouwbare decimalen en als regel geldt:

De som of het verschil van benaderende getallen bezit niet meer betrouwbare decimalen dan het kleinste aantal daarvan, dat voorkomt.

Bijv.:

276,784 + 6,51 = 283,29 en *niet* 283,294,

67,452 + 56,1 — 23,43 = 100,1 en *niet* 100,122.

Men rondt pas af, als het optellen en/of aftrekken heeft plaatsgevonden.

b. Vermenigvuldigen en delen

Hierbij dient men te letten op het aantal betrouwbare cijfers. Als regel geldt:

Het product of quotiënt van twee benaderende getallen heeft niet meer betrouwbare cijfers, dan het kleinste aantal daarvan, dat voorkomt. Het product of quotiënt van een benaderend getal en een exact getal bevat evenveel betrouwbare cijfers als het benaderend getal.

Bijv.: Beschouw het product van 8,347 en 7,93. Als beide getallen exact zijn, geldt: $8{,}347 \times 7{,}93 = 66{,}19171$. Als beide getallen benaderend zijn bevat hun product drie betrouwbare cijfers en is dus: 66,2. Als het eerste getal exact en het tweede benaderend is, verkrijgt men dezelfde uitkomst. In het omgekeerde geval bevat het product vier betrouwbare cijfers en wordt het 66,19.

Beschouw vervolgens het volgende quotiënt van benaderende getallen: $\frac{83{,}642.}{72}$ De uitkomst is 1,2, daar het getal 72 slechts twee betrouwbare cijfers telt.

c. Kwadrateren en worteltrekken

Het kwadraat van een benaderend getal telt evenveel betrouwbare cijfers als het getal zelf. Als 1,3 een benaderend getal is, geldt dus:

$$(1{,}3)^2 = 1{,}7 \text{ en niet } 1{,}69.$$

Ook de wortel uit een benaderend getal heeft evenveel betrouwbare cijfers als dat getal. Als 30,74 een benaderend getal is, geldt dus:

$$\sqrt{30{,}74} = 5{,}544.$$

OPGAVEN

1. Veronderstel dat alle onderstaande getallen benaderend zijn en geef de uitkomsten met het juiste aantal betrouwbare cijfers:

a. $56{,}3456 + 33{,}341 - 11{,}92$
b. $6{,}375 + 12{,}5 + 3{,}29$
c. $(4{,}13)^2$
d. $\sqrt{62{,}79}$
e. $(3{,}1416)\ (24{,}1)$
f. $(5{,}27834)\ (12{,}34)$
g. $\dfrac{9{,}2463}{25}$

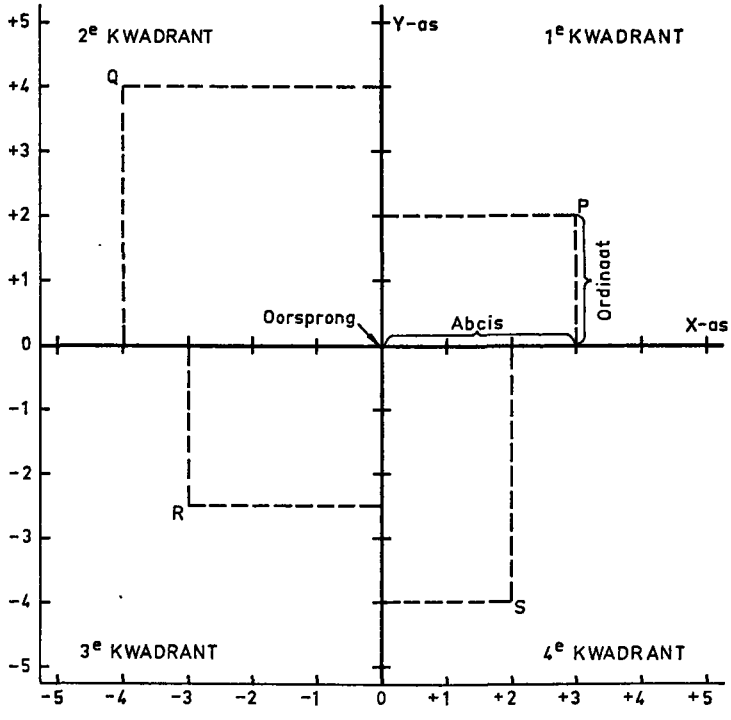
2. In onderstaande berekeningen is 25 een exact getal. Alle overige getallen zijn benaderend. Voer de berekeningen uit.

a. $(25)^2$ *b.* $(36{,}44)\ (25)$ *c.* $\dfrac{321{,}5}{25}$ *d.* $\sqrt[3]{25}$ *e.* $\dfrac{(25)\ (13{,}57)}{3{,}65}$

6. Coördinaatassen

HET AANGEVEN VAN PUNTEN IN HET ASSENSTELSEL

De meeste grafieken zijn gebaseerd op een assenstelsel, bestaande uit een horizontaal getrokken X-as en een loodrecht daarop staande Y-as. Het snijpunt der beide assen noemt men de *oorsprong* en voor beide assen is dit snijpunt het nulpunt. Het assenstelsel verdeelt het beeldvlak in vier *kwadranten* (zie figuur A). In het eerste kwadrant (rechts boven) zijn x en y beide positief, in het tweede (links boven) is x negatief en y positief, enz. De plaats van een punt t.o.v. het assenstelsel kan op ondubbelzinnige wijze met twee coördinaten worden aangeduid.



Figuur A. Coördinaatassen.

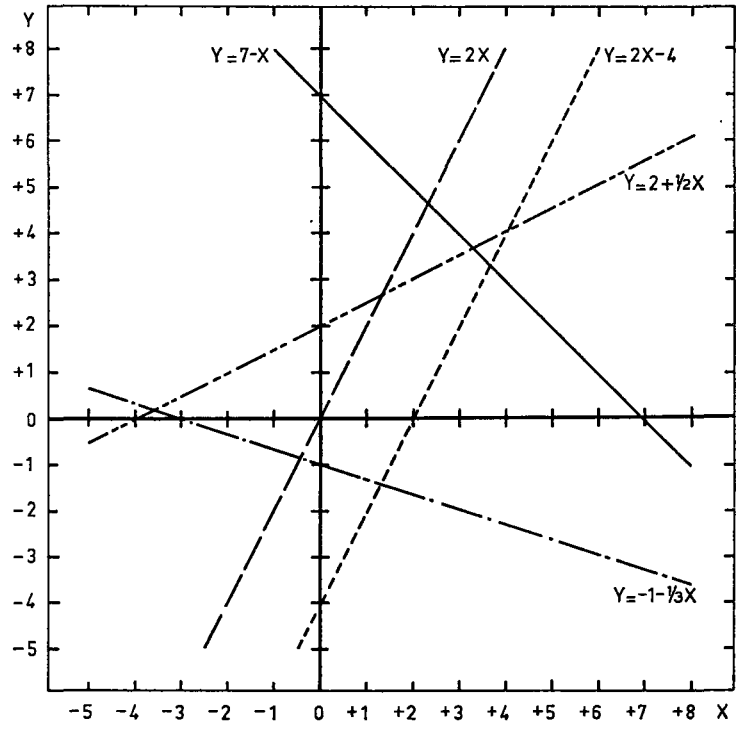
Eerst vermeldt men de *x-coördinaat* of *abscis*, d.i. de afstand van het punt tot de Y-as, gemeten evenwijdig aan de X-as. Vervolgens geeft men de *y-coördinaat* of *ordinaat*, d.i. de afstand van het punt tot de X-as gemeten evenwijdig aan de Y-as. Het punt Q in figuur A wordt zodoende aangegeven met (—4, +4), daar voor dit punt $x = -4$ en $y = 4$. Bij de grafieken, die men in de statistiek ontmoet, wordt meestal het eerste kwadrant (+, +) gebruikt.

DE RECHTE LIJN IN HET ASSENSTELSEL

Men noemt y een functie van x, als door een of andere vergelijking aan bepaalde waarden van x bepaalde waarden van y worden gekoppeld. De meest bekende functie, de *lineaire*, bezit de algemene vorm:

$$y = A + Bx,$$

waarin A en B constanten zijn. Een voorbeeld van een lineaire functie is dus: $y = 4 + 2x$. Beschouwt men x als de grootheid, waarvan de waarden worden gekozen (de primaire of onafhankelijk variabele, ofwel het argument), dan is voor elke waarde van x de waarde van y (de afhankelijk variabele of functie) vastgelegd.



Figuur B. Grafische voorstelling van enkele lineaire functies.

De grafische voorstelling van de lineaire functie is de rechte lijn. In figuur B zijn een vijftal lineaire functies in hetzelfde assenstelsel getekend. Wij zien dat de constante A in de functievergelijking het punt aangeeft, waar de rechte de Y-as snijdt; de rechte loopt door de oorsprong, als $A = 0$, dus voor de functie $y = Bx$. De constante B is bepalend voor de helling van de lijn en wordt daarom de *hellings-* of *richtingscoëfficiënt* genoemd.

Er zijn nog vele andere functies, waarvan wij slechts

$$y = AB^x$$

noemen. Door het nemen van logarithmen verkrijgt men

$$\log y = \log A + x \log B,$$

waarvan de grafische voorstelling een rechte is.

OPGAVEN

1. Teken op één assenstelsel de rechten met de volgende functievergelijkingen:
   a. $y = -3 + 2x$ *b.* $y = -3 - 2x$ *c.* $y = 3 - 2x$ *d.* $y = 3 + 2x$
2. Stel grafisch voor:
   a. $y = \frac{1}{2}x$ *b.* $y = 6 - 3x$ *c.* $y = -1 - 3x$ *d.* $y = \frac{1}{2} + \frac{1}{2}x$
   e. $\log y = \log 3 + (\log 2)x$ *f.* $\log y = \log 10 + (\log 5)x$
3. Geef de vergelijking van de rechte, die door de volgende punten loopt:
   a. (0, 0) en (−6, −2) *b.* (0, 0) en (5, −3) *c.* (−2, 3) en (0, 7)
   d. (− 1, −3) en (7, − 2) *e.* (2, −9) en (−7, −3)

BIJLAGE III

LITERATUUROPGAVE

ADAMS, J. K. (1955), Basic statistical concepts, McGraw-Hill Book Company, New York.

ALCHIAN, A. A. (1947), Analysis procedures, Army Air Forces Aviation Psychology Program, Report 18, Washington.

ANSCOMBE, F. J. (1948), The transformation of Poisson, binomial and negative binomial data, Biometrika 35, 246–254.

ARMITAGE, P. (1954), Sequential tests in prophylactic and therapeutic trials, Quart. J. Med. 23, 255–274.

ARMITAGE, P. (1955), Tests for linear trends in proportions and frequencies, Biometrics 11, 375–386.

ARMITAGE, P. (1960), Sequential medical trials, Blackwell Scientific Publications, Oxford.

ARMITAGE, P. (1961*a*), Sequential analysis in medicine, Statistica Neerlandica 15, 73–82.

ARMITAGE, P. (1961*b*), Theoretical aspects of sequential designs, Statistica Neerlandica 15, 83–90.

ARMITAGE, P. en I. ALLEN (1950), Methods of estimating the LD_{50} in quantal response data, J. of Hygiene 48, 298–322.

ASPIN, A. A. (1948), An examination and further development of a formula arising in the problem of comparing two mean values, Biometrika 35, 88–96.

ASPIN, A. A. (1949), Tables for use in comparisons whose accuracy involves two variances, seperately estimated, Biometrika 36, 290–296.

BAAS BECKING, L. G. M. en E. F. DRION (1936), On the origin of frequency distributions in biology, Acta Biotheoretica, Ser. A. 1, 133–150.

BAILEY, N. T. J. (1959), Statistical methods in biology, The English Universities Press Ltd, London.

BAKKER, O. (1955), Statistiek, deel I en II, J. Muusses, Purmerend; 8e druk.

BARLOW's TABLES (1954), (Squares, cubes, square roots, cube roots and reciprocals of all integers up to 12.500), Spon, London.

BARTLETT, M. S. (1935), The effect of non-normality on the *t* distribution, Proc. Cambr. Phil. Soc. 31, 223–231.

BARTLETT, M. S. (1936), Square root transformation in analysis of variance, Suppl. J. R. Stat. Soc. 3, 168–178.

BARTLETT, M. S. (1937), Some examples of statistical methods of research in agriculture and applied biology, Suppl. J. R. Stat. Soc. 4, 137–183.

BARTLETT, M. S. (1947), The use of transformations, Biometrics 3, 39–52.

BARTLETT, M. S., F. J. ANSCOMBE, W. G. COCHRAN en J. BERKSON (1954), Discussion of 'The analysis of variance with various binomial transformations by Sir Ronald Fisher', Biometrics 10, 140–151*.

BELLINSON, H. R., J. VON NEUMANN, R. H. KENT en B. I. HART (1941), The mean square successive difference, Ann. Math. Stat. 12, 153–162.

BENARD, A. en C. VAN EEDEN (1956), Handleiding voor de symmetrietoets van Wilcoxon, Rapport S 208 (M 76), Math. Centrum, Amsterdam.

BENJAMIN, B. (1959), Elements of vital statistics, Allen and Unwin, London.

BERKSON, J. (1942), Tests of significance considered as evidence, J. Am. Stat. Ass. 37, 325–335.

BERKSON, J. (1946), Limitations of the application of fourfold table analysis to hospital data, Biometrics Bull. 2, 47–53.

BERKSON, J. (1950), Are there two regressions, J. Am. Stat. Ass. 45, 250.

* Dit artikel bevat een afzonderlijk commentaar van elk der genoemde personen.

BENNETT, C. A. en N. L. FRANKLIN (1954), Statistical analysis in chemistry and the chemical industry, John Wiley and Sons, New York.

BEVERIDGE, W. I. B. (1951), The art of scientific investigation, Heinemann Ltd, London.

BEZEM, J. J. (1958), Toepassing van de transformatie $y = 2bg\sin\sqrt{p}$ bij de bepaling van ED_{50}, Verslag v. h. symposion over ED_{50} bepalingen, Utrecht.

BLISS, C. I. (1952), The statistics of bioassay, Academic Press, New York.

BLISS, C. I. (1956), Confidence limits for measuring the precision of bioassays, Biometrics 12, 491–526.

BLOEMENA, A. R. en R. KORSWAGEN (1960), De nauwkeurigheid van enkele methoden voor het tellen van leucocyten, Rapport S 270, Math. Centrum, Amsterdam.

BOER, J. DE (1953), Sequential test with three possible decisions for testing an unknown probability, Appl. Sc. Research B, 249–259.

BOK, S. T. (1948), De gedachtengang van de statistica, Stenfert Kroese, Leiden, 2e druk.

BOX, G. E. P. (1953), Non-normality and tests on variances, Biometrika 40, 318–335.

BOX, G. E. P. (1954), Some theorems on quadratic forms applied in the study of analysis of variance problems: I Effect of inequality of variance in one-way classification, Ann. Math. Stat. 25, 290–302.

BROSS, I. D. J. (1950), Estimates of the LD_{50}: a critique, Biometrics 6, 413–423.

BROSS, I. D. J. (1952), Sequential medical plans, Biometrics 8, 188.

BROSS, I. D. J. (1953), Design for decision, The McMillan Company, New York.

BROWNLEE, K. A. (1949), Industrial experimentation, Chemical Publishing Co., New York, 3rd ed.

BROWNLEE, K. A., J. L. HODGES en M. ROSENBLATT (1953), The up-and-down method with small samples, J. Am. Stat. Ass. 48, 262–272.

BURN, J. H. (1952), Biological standardization, Oxford University Press, London, 2nd ed.

CLARINGBOLD, P. J., J. D. BIGGERS en C. W. EMMENS (1953), The angular transformation in quantal analysis, Biometrics 9, 467–484.

CLOPPER, C. J. en E. S. PEARSON (1934), The use of confidence or fiducial limits illustrated in the case of the binomial, Biometrika 26, 404–413.

COCHRAN, W. G. (1941), The distribution of the largest of a set of estimated variances as a fraction of their total, Ann. of Eugenics 11, 47–52.

COCHRAN, W. G. (1943), Analysis of variance for percentages based on unequal numbers, J. Am. Stat. Ass. 38, 287–301.

COCHRAN, W. G. (1947), Some consequences when the assumptions for analysis of variance are not satisfied, Biometrics 3, 22.

COCHRAN, W. G. (1950), The comparison of percentages in matched samples, Biometrika 37, 256–266.

COCHRAN, W. G. (1952), The χ^2-test of goodness of fit, Ann. Math. Stat. 23, 315–345.

COCHRAN, W. G. (1953), Sampling techniques, John Wiley and Sons, New York.

COCHRAN, W. G. (1954), Some methods for strengthening the common χ^2-tests, Biometrics 10, 417–451.

COCHRAN, W. G. en G. M. COX (1950), Experimental designs, John Wiley and Sons, New York.

COCHRAN, W. G. en M. DAVIS (1961), Sequential experiments for estimating the median lethal dose, lezing voor het Colloquium 'Sur le plan d'expériences', Parijs.

CONTROLLED CLINICAL TRIALS (1960), A conference organized by C.I.O.M.S.,

Vienna 23–27 March 1959, Blackwell Scientific Publications, Oxford.
CORNFIELD J. en W. HAENSZEL (1960), Some aspects of retrospective studies, J. of Chronic Diseases 5, 523.
CROXTON, F. E. (1953), Elementary statistics with applications in medicine, Prentice-Hall Inc., New York.

DANTZIG, D. VAN (1950), Enkele historische betrekkingen tussen mathematische en verzamelende statistiek, Statistica 4, 233–248.
DANTZIG, D. VAN (1951), De natuur als tegenspeler, Statistica 5, 153–159.
DANTZIG, D. VAN (1953), De verantwoordelijkheden van de statisticus, Statistica 7, 199–208.
DAVID, F. N. (1938), Tables of the correlation coefficient, Cambridge University Press, Cambridge.
DAVID, F. N. (1949), Probability theory for statistical methods, Cambridge University Press, New York.
DAVIES, O. L. (1949), Statistical methods in research and production, Oliver & Boyd, Edinburgh.
DAVIES, O. L., H. C. HAMAKER en F. J. VAN DUN (1955), Design and analysis of industrial experiments, Statistica Neerlandica 9, 189–207.
DEMING, W. E. (1950), Some theory of sampling, John Wiley and Sons, New York.
DIXON, W. J. (1950), Analysis of extreme values, Ann. Math. Stat. 21, 488–506.
DIXON, W. J. (1951), Ratio's involving extreme values, Ann. Math. Stat. 22, 68–78.
DIXON, W. J. (1953), Processing data for outliers, Biometrics 9, 74–88.
DIXON, W. J. en F. J. MASSEY (1957), Introduction to statistical analysis, McGraw-Hill Book Company, New York, 2nd ed.
DIXON, W. J. en A. M. MOOD (1946), The statistical sign test, J. Am. Stat. Ass. 41, 557–566.
DIXON, W. J. en A. M. MOOD (1948), A method for obtaining and analyzing sensitivity data, J. Am. Stat. Ass. 43, 109–126.
DOLL, R. en A. B. HILL (1952), A study of the aetiology of carcinoma of the lung, Brit. Med. J. ii, 1271.
DOLL, R. en A. B. HILL (1956), Lung cancer and other causes of death in relation to smoking, Brit. Med. J. ii, 1071.
DOORNBOS, R. (1956), Significance of the smallest of a set of estimated normal variances, Statistica Neerlandica 10, 117–126.
DOORNBOS, R. (1959), Statistische methoden voor het aanwijzen van uitbijters, Statistica Neerlandica 13, 453–462.
DRION, E. F. (1936), On the interpretation of frequency curves in biology, Dissertatie Leiden, Amsterdam.
DRION, E. F. (1950), Waarschijnlijkheidsrekening en biologie, Vacantiecursus waarschijnlijkheidsrekening, Math. Centrum, Amsterdam.
DRION, E. F. (1953), Enige statistische begrippen, N.T.v.G. 97, 22/8/53.
DRION, E. F. (1959), Wiskundige statistiek, Technisch Repertorium I, 216–265, Elsevier, Amsterdam.
DUNCAN, D. B. (1955), Multiple range and multiple F tests, Biometrics 11, 1–42.

EDGINGTON, E. S. (1961), Probability table for number of runs of signs of first differences in ordered series, J. Am. Stat. Ass. 56, 156–159.
EDWARDS, A. L. (1960), Experimental design in psychological research, Rinehart and Comp., New York, rev. ed.
EEDEN, C. VAN (1953), Methoden voor het vergelijken, toetsen en schatten van onbekende kansen, Rapport S 115 (M 45), Math. Centrum, Amsterdam.

EEDEN, C. VAN (1955), Een toets tegen verloop voor een aantal kansen, Statistica Neerlandica 9, 131–159.

EEDEN, C. VAN (1956), Verdelingsvrije toetsen voor twee steekproeven en de methode van de 2 × 2 tabel, Statistica Neerlandica 10, 157–162.

EEDEN, C. VAN en J. HEMELRIJK (1955), A test for the equality of probabilities against a class of specified alternative hypotheses including trend, Proc. Kon. Ned. Akad. v. Wet. A 58, Indag. Math. 17, 191–198 en 301–308.

EISENHART, C. (1939), The interpretation of certain regression methods and their use in biological and industrial research, Ann. Math. Stat. 10, 162–186.

EISENHART, C. c.s. (1947), Selected techniques of statistical analysis, Columbia University, Statistical Research Group, McGraw-Hill Book Company, New York.

EKKER, W. (1961), Prospectief en retrospectief onderzoek, lezing voor de Medisch-Biologische Sectie van de Vereniging voor Statistiek.

ELTEREN, Ph. VAN (1951), Methode van de m rangschikkingen, Cursus Parametervrije methoden, Rapport S 59, Math. Centrum, Amsterdam.

ELTEREN, Ph. VAN (1961), Two restricted sequential sign tests, Statistica Neerlandica 15, 91–96.

ELTEREN, Ph. VAN en CHR. L. RÜMKE (1956), De sequente binomiale toets, Math. Centrum, Amsterdam.

EMMENS, C. W. (1948), Principles of biological assay, Chapmann & Hall, London.

ENTERS, J. H. (1951), Het gebruik van toevalscijfers, Statistica 5, 81–95.

ENTERS, J. H. (1957), De methode van de m rangschikkingen, Mededelingen Medisch-Biologische Sectie V.v.S 3.

EPSTEIN, B. en C. W. CHURCHMAN (1944), On the statistics of sensitivity data, Ann. Math. Stat. 15, 90–96.

EXALTO, L. J. H. (1950), Gebruik en toepassing van waarschijnlijkheidspapier, Statistica 4, 121–128.

FABER, M., E. ANDREASEN en H. UHRBRAND (1957), Further studies on irradiation-induced leukaemia in Denmark, Trans. 6th Congress Eur. Soc. Haematology, Copenhagen.

FERRO, H. (1954), Grafiekenpapier, Statistica 8, 123–154.

FIELLER, E. C. (1940), The biological standardization of insulin, Suppl. J. R. Stat. Soc. 7, 1–64.

FINNEY, D. J. (1948), The Fisher-Yates test of significance in 2 × 2 contingency tables, Biometrika 35, 145–156.

FINNEY, D. J. (1949), The choice of a response metameter in bio-assay, Biometrics 5, 261–272.

FINNEY, D. J. (1952*a*), Probit analysis, Cambridge University Press, London, 2nd ed.

FINNEY, D. J. (1952*b*), Statistical method in biological assay, Ch. Griffin & Comp., London.

FISHER, R. A. (1915), Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population, Biometrika 10, 507.

FISHER, R. A. (1921), On the 'probable error' of a coefficient of correlation deduced from a small sample, Metron 1, 1–32.

FISHER, R. A. (1947), The design of experiments, Oliver and Boyd, Edinburgh, 4th ed.

FISHER, R. A. (1948), Statistical methods for research workers, Oliver and Boyd, Edinburgh, 10th ed.

FISHER, R. A. (1954), The analysis of variance with various binomial transformations, Biometrics 10, 130–139.

FISHER, R. A. en F. YATES (1953), Statistical tables, Oliver and Boyd, Edinburgh, 4th ed.

FORTUIN, G. J. (1959), Wachttijden van patiënten, Statistica Neerlandica 13, 339–346.

FREEMAN, M. F. en J. W. TUKEY (1950), Transformations related to the angular and the square root, Ann. Math. Stat. 21, 607–611.

FREUDENTHAL, H. (1957), Waarschijnlijkheid en statistiek, De Erven F. Bohn N.V., Haarlem.

FRIEDMAN, M. (1937), The use of ranks to avoid the assumption of normality, J. Am. Stat. Ass. 32, 675–701.

FRY, T. C. (1938), The χ^2 test of significance, J. Am. Stat. Ass. 33, 513.

GADDUM, J. H. (1945), Lognormal distributions, Nature 156, 463–466.

GAYEN, A. K. (1949), The distribution of Student's t in random samples of any size from non-normal universes, Biometrika 36, 353–369.

GAYEN, A. K. (1950), Significance of difference between the means of two non-normal samples, Biometrika 37, 399–408.

GEARY, R. C. (1936), The distribution of 'Student's' ratio for non-normal samples, Suppl. J.R. Stat. Soc. 3, 178–184.

GEARY, R. C. (1947), Testing for normality, Biometrika 34, 209–242.

GEARY, R. C. en E. S. PEARSON (1938), Tests of normality, Biometrika Office, University College, London.

GEIGY (1960), Wissenschaftliche Tabellen, J.R. Geigy A.G., Basel, 6. Auflage.

GEZONDHEIDSORGANISATIE T.N.O. (1958), Waardering van medische gegevens uit ziekenhuizen, Van Gorcum, Assen.

GLASSER, G. J. (1962), A distribution-free test of independance with a sample of paired observations, J. Am. Stat. Ass. 57, 116–133.

GOOSSEN, J. A. H. (1958), Controle op klinisch-chemische onderzoekingen met behulp van controlekaarten, Mededelingen Medisch-Biologische Sectie V.v.S 5.

GOSLINGS, J. (1961), De medisch-ethische zijde van de sequente analyse, Statistica Neerlandica 15, 55–72.

GREENBERG, B. G. (1951), Why randomize? Biometrics 7.

GRIMM, G. (1958), Einige Probleme der Bakterienzählung, Voordracht op de Biometrische Dag te Brussel.

GRONOW, D. G. C. (1953), Non-normality in two sample t-tests, Biometrika 40, 222.

GROOT, M. J. W. DE (1958), Elementaire kwantitatieve analyse van ziekteverzuimcijfers, Mens en Onderneming 12, 219–233. Inleiding gehouden in Sectie VII, tijdens de verzuimbijeenkomst in het N.I.P.G., op 28 maart 1958.

GROOT, M. J. W. DE en P. VAN LEEUWEN (1960), Consequenties van standaardisatie naar leeftijd voor de ziekteverzuimcijfers der bedrijven, Mens en Onderneming 14, 220–237.

GRUBBS, F. E. (1950), Sample criteria for testing outlying observations, Ann. Math. Stat. 21, 27–58.

HALD, A. (1952), Statistical tables and formulas, John Wiley and Sons, New York.

HALDANE, J. B. S. (1939), The mean and variance of χ^2 when used as a test of homogeneity, when expectations are small, Biometrika 31, 341–355.

HAMAKER, H. C. (1948), Toevalscijfers, Statistica 2, 97–106.

HAMAKER, H. C. (1955*a*), Naar efficiënte experimenten, Statistica Neerlandica 9, 7–25.

HAMAKER, H. C. (1955*b*), Proefopzetten in de industrie, Statistica Neerlandica 9, 209–232.

HAMAKER, H. C. (1958), Statistiek en experiment, Statistica Neerlandica 12, 119–130.

HAMMOND, E. C. en D. HORN (1958), Smoking and death rates, report on 44 months of follow-up of 187.783 men, J. Am. Med. Ass. 155, 1316.

HART, B. I. (1942), Significance levels for the ratio of the mean square successive difference to the variance, Ann. Math. Stat. 13, 445–447.

HARTLEY, H. O. (1950), The maximum F-ratio as a short-cut test for heterogeneity of variance, Biometrika 37, 308–312.

HARVARD UNIVERSITY (1955), Tables of the cumulative binomial probability distribution, Harvard University Press, Cambridge, Mass.

HEMELRIJK, J. (1950*a*), Symmetrietoetsen, Dissertatie, Den Haag.

HEMELRIJK, J. (1950*b*), Rangcorrelatie en de schattingsproef van Varangot, Statistica 4, 216–225.

HEMELRIJK, J. (1951), Parametrische en parametervrije methoden en hun toepassingen, Statistica 5, 171–184.

HEMELRIJK, J. (1952), Het combineren van onafhankelijke toetsen, Memorandum S 73 (M 17*a*) en aanvulling: S 102 (M 17*b*), Math. Centrum, Amsterdam.

HEMELRIJK, J. (1952/53), Cursus elementaire mathematische statistiek, Math. Centrum, Amsterdam.

HEMELRIJK, J. (1957), Distributionfree tests against trend and maximum likelihood estimates of ordered parameters, Rapport SP 56, Math. Centrum, Amsterdam; Proc. 30th Session I.S.I., Stockholm.

HEMELRIJK, J. (1958*a*), Het begrip nauwkeurigheid, Mededelingen Medisch-Biologische Sectie V.v.S. 4.

HEMELRIJK, J. (1958*b*), Statistische proefopzetten; bewijs en detectie, Statistica Neerlandica 12, 111–118.

HEMELRIJK, J. (1961), Experimental comparison of Student's and Wilcoxon's two sample tests, Quantitative Methods in Pharmacology, North-Holland Publ. Comp., Amsterdam.

HEMELRIJK, J. en PH. VAN ELTEREN (1953), Cursus toegepaste statistiek, Math. Centrum, Amsterdam.

HEMELRIJK, J. en C. G. VAN HERK (1958), Een geval van statistische detectie, Statistica Neerlandica 12, 191–200.

HEMELRIJK, J. en H. R. VAN DER VAART (1950), Het gebruik van één- en tweezijdige overschrijdingskansen voor het toetsen van hypothesen, Statistica 4, 54–66.

HEY, B. B. (1938), A new method of experimental sampling illustrated on certain non-normal populations, Biometrika 30, 68–80.

HILL, B. A. (1951), The clinical trial, Br. Med. Bull. 7, 278.

HILL, B. A. (1961), Principles of medical statistics, The Lancet Ltd, London, 5th ed.

HOEFLAKE, G. J. (1959), De statistiek van de Militair Geneeskundige Dienst, Statistica Neerlandica 13, 187–196.

HOEL, P. G. (1951), Introduction to mathematical statistics, John Wiley and Sons, New York.

HOLTZ, A. H. (1960), Klinisch-chemische enquêtes in Nederland, Verslagen en mededelingen betreffende de Volksgezondheid 1, 211–236.

HOTELLING, H. (1951), The impact of R. A. Fisher on statistics, J. Am. Stat. Ass. 46, 35–46.

HSU, P. L. (1938), Contributions to the theory of Student's t-test as applied to the problem of two samples, Stat. Research Memoirs 2, 1–24.

HUFF, D. (1954), How to lie with statistics, Norton and Company, New York,

(Ned. vertaling: Gebruik en misbruik van de statistiek, Prisma 572, Het Spectrum, Utrecht, 1960).

INGLE, D. J. (1958), Principles of research in biology and medicine, Lippincott, Philadelphia.

IRWIN, J. O. (1949), A note on the subdivision of chi-square into components, Biometrika 36, 130–134.

ITERSON Jr, G. VAN (1950), Het nut van waarschijnlijkheidspapier voor de variatiestatistiek, Statistica 4, 129–157.

JERNE, N. K. en E. C. WOOD (1949), The validity and meaning of the results of biological assays, Biometrics 5, 273–299.

JESSOP, W. N. (1952), One line or two, Applied Stat. 1, 131–137.

JOHNSON, P. O. (1950), Statistical methods in research, Prentice-Hall Inc., New York.

JONGE, H. DE (1959), Klassieke methoden gebaseerd op de gemiddelden van één of twee steekproeven, Mededelingen Medisch-Biologische Sectie V.v.S. 8.

JONGE, H. DE (1960), Klassieke en verdelingsvrije toetsen voor k waarnemingsreeksen, Rapport Afd. Stat. N.I.P.G.

JONGE, H. DE (1961), The influence of non-normality on the probability of a type I error of Student's two sample test, Quantitative Methods in Pharmacology, North-Holland Publ. Comp., Amsterdam.

JONGE, H. DE (1962*a*), Toetsen voor k steekproeven, Rapport Afd. Stat. N.I.P.G.

JONGE, H. DE (1962*b*), Toetsen voor de hypothese H_0: $\mu_1 = \mu_2 = \ldots = \mu_k$, Rapport Afd. Stat. N.I.P.G.

JONGH, D. K. DE (1954), De betekenis van het placebo, N.T.v.G. 98, 1943–1950.

JONGH, D. K. DE (1958), Design for decision in het klinische experiment, Statistica Neerlandica 12, 213–230.

KAARSEMAKER, L. en A. VAN WIJNGAARDEN (1952), Tables for use in rank correlation, Math. Centrum, Amsterdam.

KAPTEYN, J. C. (1903), Skew frequency curves in biology and statistics, P. Noordhoff, Groningen.

KAPTEYN, J. C. en M. J. VAN UVEN (1916), Skew frequency curves in biology and statistics, Hoitsema Brothers, Groningen, 2nd paper.

KÄRBER, G. (1931), Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche, Arch. f. Exp. Path. u. Pharm. 162, 480–487.

KASTENBAUM, M. A. (1960), A note on the additive partitioning of chi-square in contingency tables, Biometrics 16, 416–422.

KEMPTHORNE, O. (1952), The design and analysis of experiments, John Wiley and Sons, New York.

KENDALL, M. G. (1938), A new measure of rank correlation, Biometrika 30, 81–93.

KENDALL, M. G. (1948), Rank correlation methods, Ch. Griffin and Comp., London.

KENDALL, M. G. en B. BABINGTON SMITH (1946), Tables of random sampling numbers, Cambridge University Press.

KEULS, M. (1952), The use of the 'Studentized range' in connection with an analysis of variance, Euphytica 1, 112–122.

KINCAID, W. M. (1962), The combination of tests based on discrete distributions, J. Am. Stat. Ass. 57, 10–19.

KLERK-GROBBEN, G. en H. J. PRINS (1954), Toets voor de gelijkheid van twee kleine kansen met behulp van even grote steekproeven en het onderscheidingsvermogen van deze toets, Statistica 8, 7–20.

KOLMOGOROV, A. (1941), Confidence limits for an unknown distribution function, Ann. Math. Stat. 12, 461–463.

KRUSKAL, W. H. (1952), A non-parametric test for the several sample problem, Ann. Math. Stat. 23, 525–539.

KRUSKAL, W. H. en W. A. WALLIS (1952), Use of ranks in one-criterion variance analysis, J. Am. Stat. Ass. 47, 584–618.

KUIPERS, H. A. (1957), Overzichtsrapport: Over de verdeling van het aantal 'runs' in reeksen van alternatieven, Rapport S 219 (Ov 7) Math. Centrum, Amsterdam.

LAAN, E. VAN DER en G. W. IGNATIUS (1946/47), Lijnvereffening met behulp van orthogonale polynomia, Statistica 1, 219–230.

LATSCHA, R. (1953), Tests of significance in a 2 × 2 contingency table: Extension of Finney's table, Biometrika 40, 74–86.

LEHMAN, S. Y. (1961), Exact and approximate distribution for the Wilcoxon statistic with ties, J. Am. Stat. Ass. 56, 293–298.

LINDLEY, D. V. (1947), Regression lines and the linear functional relationship, Suppl. J. R. Stat. Soc. 9, 218.

LISON, L. (1958), Statistique appliquée à la biologie expérimentale, Gauthier-Villars, Parijs.

LITCHFIELD, J. T. en F. WILCOXON (1955), The rank correlation method, Analytical Chemistry 27, 229–300.

LITCHFIELD, J. T. en F. WILCOXON (1949), A simplified method of evaluating dose-effects experiments, J. Pharmacol. 95, 99–113.

LORD, E. (1947), The use of the range in place of the standard deviation in the *t*-test, Biometrika 34, 41.

MACKINNON, W. J. (1959), Compact table of twelve probability levels of the symmetric binomial cumulative distribution for sample sizes to 1000, J. Am. Stat. Ass. 54, 164.

MADOW, W. G. (1940), Note on tests of departure from normality, J. Am. Stat. Ass. 35, 515–517.

MAINLAND, D. (1952), Elementary medical statistics, W. B. Saunders Company, Philadelphia-London.

MAINLAND, D. (1954), Statistics in medical research, Methods in Medical Research Vol. 6, Section III, The Year Book Publishers, Chicago.

MAINLAND, D. (1960), The use and misuse of statistics in medical publications, Clinical Pharmacology and Therapeutics 1, 411–422.

MAINLAND, D., L. HERRERA en M. I. SUTCLIFFE (1956), Statistical tables for use with binomial estimates, Dept. of Medical Statistics, New York University School of Medicine.

MAINLAND, D., L. HERRERA, M. I. SUTCLIFFE en E. STREET (1959), Notes from a laboratory of medical statistics: from the department of medical statistics, New York University College of Medicine.

MANN, H. B. (1945), Non parametric tests against trend, Econometrica 13, 245–259.

MANN, H. B. en D. R. WHITNEY (1947), On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Stat. 18, 50–60.

MASSEY Jr., F. J. (1951), The Kolmogorov-Smirnov test for goodness of fit, J. Am. Stat. Ass. 46, 68–78.

MATHEMATISCH CENTRUM (1952), Memoranda van de statistische afdeling, Math. Centrum, Amsterdam.

MATHEMATISCH CENTRUM (1955), Een parametervrije toets tegen verloop voor groepen waarnemingen, Rapport S 168 (M 61), Math. Centrum, Amsterdam.

MATHER, K. (1951), R. A. Fisher's 'Statistical methods for research workers', J. Am. Stat. Ass. 46, 51–54.

MAXWELL, A. E. (1958), Experimental design in psychology and the medical sciences, Methuen and Company, London.
MAXWELL, A. E. (1961), Analysing qualitative data, Methuen and Company, London.
MAY, J. M. (1952), Extended and corrected tables of the upper percentage points of the 'Studentized' range, Biometrika 39, 192–193.
MCDONALD, A. D. (1958), Maternal health and congenital defect: a prospective investigation, New Engl. J. Med. 258, 767.
MERRINGTON, M. en C. M. THOMPSON (1946), Tables for testing the homogeneity of a set of estimated variances, Biometrika 33, 296–304.
MOLINA, E. C. (1942), Poisson's exponential binomial limit, Van Nostrand, New York.
MOOD, A. M. (1940), The distribution theory of runs, Ann. Math. Stat. 11, 367–392.
MOOD, A. M. (1950), Introduction to the theory of statistics, McGraw-Hill Book Company, New York.
MOORE, G. H. en W. A. WALLIS (1943), Time series significance tests based on signs of differences, J. Am. Stat. Ass. 38, 153–164.
MORONEY, M. J. (1953), Facts from figures, Penguin Books Ltd., London, 2nd ed.
MOSTELLER, F. en R. R. BUSH, Selected quantitative techniques, ch. 8, Handbook of Social Psychology, Vol. 1, Harvard University Press.
NAIR, K. R. (1948), The distribution of the extreme deviate from the sample mean and its Studentized form, Biometrika 35, 118–144.
NASS, C. A. G. (1946/47), Het quantitatieve verband tussen dosering en letaliteit bij proefdieren, Statistica 1, 257–266.
NATIONAL BUREAU OF STANDARDS (1950), Tables of the binomial probability distribution, Washington.
NATRELLA, M. G. (1960), The relation between confidence intervals and tests of significance, The American Statistician 14, 20.
NEUMANN, J. VON (1941), Distribution of the ratio of the mean square successive difference to the variance, Ann. Math. Stat. 12, 367–395.
NEURDENBURG, M. G. (1959/60), Gedachten over statistiek, geneeskunde en gezondheidszorg, T. v. Sociale Geneeskunde, 747–752 / 93–96, 447–456 en 633–655.
NEWMAN, D. (1939), The distribution of the range in samples from a normal population, expressed in terms of an independent estimate of standard deviation, Biometrika 31, 20–30.
NEYMAN, J. (1950), First course in probability and statistics, Henry Holt and Company, New York.
NEYMAN, J. en E. S. PEARSON (1933), On the problem of the most efficient tests of statistical hypotheses, Ph. Trans. Royal Society, Ser. A, Vol. 231, London.
NORMALISATIECOMMISSIE 70 (1951), Weergeven van waarnemingsreeksen (ontwerp), Waltman, Delft.
OFFICE OF CHIEF OF ORDNANCE (1952), Tables of the cumulative binomial probabilities, Ordnance Corps Pamphlet ORD P 201, Washington.
OLDS, E. G. (1949), The 5% significance levels for sums of squares of rank differences and a correction, Ann. Math. Stat. 20, 117–118.
PANNE, C. VAN DE en H. LANGE (1958), Scheefheidstoetsen, Statistica Neerlandica 12, 255–264.
PANNEKOEK, J. H. (1956), Critiek bij de beoordeling van de werkzaamheid van geneesmiddelen, Medisch Contact, 26/1/56.

PAULSON, E. (1942), An approximate normalization of the analysis of variance distribution, Ann. Math. Stat. 13, 233–235.

PEARL, R. (1941), Introduction to medical biometry and statistics, W. B. Saunders Company, 3rd ed.

PEARSON, E. S. (1929), The distribution of frequency constants in small samples from non-normal symmetrical and skew populations, Biometrika 21, 259–286.

PEARSON, E. S. (1931), The analysis of variance in cases of non-normal variation, Biometrika 23, 114–133.

PEARSON, E. S. en C. CHANDRA SEKAR (1936), The efficiency of statistical tools and a criterion for the rejection of outlying observations, Biometrika 28, 308–320.

PEARSON, E. S. en H. O. HARTLEY (1954), Biometrika tables for statisticians, Vol. I, Cambridge University Press.

PEARSON, K. (1900), On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it cannot be reasonably supposed to have arisen from random sampling, Philosophical Magazine 5th series 50, 339–357.

PEARSON, K. (1948), Karl Pearson's early statistical papers, Cambridge University Press, Cambridge, Mass.

PERRY, W. L. M. (1950), The design of toxicity tests, Report no. 270, Medical Research Council, H. M. Stationery Office, London.

PITMAN, E. J. G. (1937), Significance tests which may be applied to samples from any populations, Suppl. J. R. Stat. Soc. 4, 119–129.

PROOSDIJ-HARTZEMA, E. G. VAN (1953), Indirecte bloeddrukmeting bij de rat, Dissertatie, Amsterdam.

PROOSDIJ-HARTZEMA, E. G. VAN (1958), De graphische bepaling van de ED_{50} volgens Litchfield-Wilcoxon, Verslag van het symposion over ED_{50} bepalingen, Utrecht.

QUENOUILLE, M. H. (1952), Associated measurements, Butterworths Scientific Publications, London.

RAND CORPORATION, THE (1955), A million random digits with 100.000 normal deviates, Free Press, Glencoe, Ill.

RIJKEN VAN OLST, H. (1958), Inleiding tot de sociale statistiek, van Gorcum & Comp. N.V., Assen, deel I, (1959) deel II.

RIJKOORT, P.J. (1952), A generalisation of Wilcoxon's test, Proc. Kon. Ned. Akad. v. Wet. A 55, 394.

RIJKOORT, P. J. en M. E. WISE (1953), Simple approximations and nomograms for two ranking tests, Proc. Kon. Ned. Akad. v. Wet. A 56, 294–302 en Indag. Math. 15, nr. 3.

ROBBINS, H. en S. MUNRO (1951), A stochastic approximation method, Ann. Math. Stat. 22, 400–407.

ROMIG, H. G. (1953), 50–100 Binomial tables, John Wiley & Sons, New York, Chapman & Hall, London.

RÜMKE, Chr. L. (1953), Kwantificering in medisch onderzoek, Statistica 7, 223–232.

RÜMKE, Chr. L. (1954), De fouten bij de telling van bloedcellen in een telkamer, N.T.v.G. 98, 3480–3485.

RÜMKE, Chr. L. (1958*a*), De fouten in de uitkomsten van de differentiatie van cellen in een uitstrijkpreparaat, N.T.v.G. 102, 2505–2508.

RÜMKE, CHR. L. (1958*b*), De toepassing van aselecteringsprocedures, Mededelingen Medisch-Biologische Sectie V.v.S. 6.

RÜMKE, CHR. L. (1958*c*), Toepassing van statistiek bij experimenten in een farmacologisch laboratorium, Statistica Neerlandica 12, 149–158.
RÜMKE, CHR. L. (1958*d*), De taak van de medische statistiek, Openbaar College.
RÜMKE, CHR. L. (1961*a*), Een algemene inleiding tot de sequente analyse, Statistica Neerlandica 15, 47–54.
RÜMKE, CHR. L. (1961*b*), De nauwkeurigheid van leucocytentellingen, N.T.v.G. 105, 2016–2022.
RÜMKE, CHR. L. (1961*c*), Generalisatie in de farmacologie, Openbare les.
RÜMKE, CHR. L. en A. R. BLOEMENA (1958), De bloedproef statistisch bezien, Statistica Neerlandica 12, 75–82.
RÜMKE, CHR. L. en C. VAN EEDEN (1961), Statistiek voor medici, Stafleu en Zoon, Leiden.

SATTERTHWAITE, F. E. (1946), An approximate distribution of estimates of variance components, Biometrics Bull. 2, 110–114.
SCHEFFÉ, H. (1959), The analysis of variance, John Wiley and Sons, New York-London.
SCHEIJDE, E. H. (1960), Orde en wanorde in de medische statistiek in ziekenhuizen, Statistica Neerlandica 14, 295–307.
SCHWARTZ, D. en G. ANGUERA (1957), Une cause de biais dans certaines enquêtes médicales: le temps de séjour des malades à l'hôpital, Proc. 30th Session I.S.I., Stockholm.
SIEGEL, S. (1956), Nonparametric statistics, McGraw-Hill Book Company, New York.
SIEGEL, S. en J. W. TUKEY (1960), A nonparametric sum of ranks procedure for relative spread in unpaired samples, J. Am. Stat. Ass. 55, 429–445.
SMID, L. J. (1956), On the distribution of the test statistic of Kendall and Wilcoxon's two sample test when ties are present, Statistica Neerlandica 10, 205–214.
SMID, L. J. (1959), De symmetrietoets van Wilcoxon, indien 'gelijken' optreden, Statistica Neerlandica 13, 463–464.
SMIRNOV, N.V. (1948), Table for estimating the goodness of fit of empirical distributions, Ann. Math. Stat. 19, 279–281.
SMITH, H. F. (1936), The problem of comparing the results of experiments with unequal errors, J. Sc. Ind. Research (India) 9, 211–212.
SNEDECOR, G. W. (1950), Statistical methods, Iowa State College Press, Ames, Iowa, 4th ed., 4th printing.
SPEARMAN, CH. (1904), The proof and measurement of association between two things, Am. J. of Psychology 15, 72–101.
SPEARMAN, CH. (1908), The method of 'right and wrong cases' ('constant stimuli') without Gauss's formulae, Br. J. of Psychology 2, 227–242.
SPITZ, J. C. (1956/57), Syllabus 'Inleiding variantie-analyse', Psych. Lab. Soc.-Paed. Instituut, Universiteit van Amsterdam.
SPITZ, J. C. (1961), De l-toets en l'-toets, volwaardige vervangers van enkele gebruikelijke χ^2-toetsen, Ned. T. v. Psychologie, Nieuwe reeks 16, 68–88.
SPROWLS, R. C. (1955), Elementary statistics, McGraw-Hill Book Company, New York.
STATISTICAL RESEARCH GROUP, COLUMBIA UNIVERSITY (1945), Sequential analysis of statistical data, Applications, Columbia University Press, New York, rev. ed.
STATISTISCHE TABELLEN EN NOMOGRAMMEN, H. E. Stenfert Kroese N.V., Leiden, (uitgave in samenwerking met de Vereniging voor Statistiek).
STEEL, R. G. D. en J. H. TORRIE (1960), Principles and procedures of statistics, McGraw-Hill Book Company, New York-Toronto-London.

STEPHAN, F. F. en PH. J. MCCARTHY (1958), Sampling opinions, John Wiley and Sons, New York-London.
STOUFFER, S. A. c.s. (1949), The American soldier, Vol. I, Adjustment during army life, Princeton University Press, Princeton.
STRIK, R. VAN (1958*a*), Statistiek en experiment rondom het poliomyelitis vaccin, Statistica Neerlandica 12, 159–169.
STRIK, R. VAN (1958*b*), Bepaling van de mediaan-effectieve dosis (ED_{50}) en de relatieve effectiviteit met de probit-analyse, Verslag v. h. symposion over ED_{50} bepalingen, Utrecht.
STRIK, R. VAN (1959), Toeval of regelmaat in het fluctuatiepatroon van opeenvolgende waarnemingsuitkomsten: de toepassing van toetsen gebaseerd op runs, Mededelingen Medisch-Biologische Sectie V.v.S. 7.
STRIK, R. VAN (1960), Uitschieters, Sigma 6, 10–17.
STUDENT (W. S. GOSSET) (1908), The probable error of a mean, Biometrika 6, 1.
STUDENT (W. S. GOSSET) (1909), On the distribution of the means of samples which are not drawn at random, Biometrika 7, 210–214.
SWAROOP, S. (1960), Introduction to health statistics, E. & S. Livingstone Ltd., Edinburgh en London.
SWED, F. S. en C. EISENHART (1943), Tables for testing randomness of grouping in a sequence of alternatives, Ann. Math. Stat. 14, 66–87.

TERPSTRA, T. J. (1952), The asymptotic normality and consistency of Kendall's test against trend when ties are present in one ranking, Proc. Kon. Ned. Akad. v. Wet. A 55, Indag. Math. 14, 327–333.
TERPSTRA, T. J. (1955/56), A generalization of Kendall's rank correlation statistic, Proc. Kon. Ned. Akad. v. Wet. A 58 en 59, Indag. Math. 17 en 18, 690–696 en 59–66.
TERRY, M. E. (1952), Some rank order tests, which are most powerful against parametric alternatives, Ann. Math. Stat. 23, 346–366.
THOMPSON, C. M. (1935), On a criterion for the rejection of observations and the distribution of the ratio of the deviation to the sample standard deviation, Ann. Math. Stat. 6, 214–219.
THOMPSON, C. M. (1941), Tables of the percentage points of the χ^2-distribution, Biometrika 42, 188–189.
THOMPSON, C. M. (1947), Use of moving averages and interpolation to estimate median-effective dose. I. Fundamental formulas, estimation of error, and relation to other methods, Bact. Reviews 11, 115–145.
TIPPETT, L. H. C. (1927), Random sampling numbers, Tracts for computors XV, Cambridge University Press, London.
TUKEY, J. W. (1949), Comparing individual means in the analysis of variance, Biometrics 5, 99.
TUKEY, J. W. (1951), Quick and dirty methods in statistics, part II, Simple analyses of standard designs, Proceedings Fifth Annual Convention, American Society for Quality Control, 189–197.

VAART, H. R. VAN DER (1950), Gebruiksaanwijzing voor de toets van Wilcoxon, Rapport S 32 (M 4), Math. Centrum, Amsterdam.
VEEN, B. (1956), Normaal waarschijnlijkheidspapier, Sigma 2, 80–83.

WABEKE, D. en C. VAN EEDEN (1955), Handleiding voor de toets van Wilcoxon, Rapport S 176 (M 65–65A), Math. Centrum, Amsterdam.
WAERDEN, B. L. VAN DER (1957), Mathematische Statistik, Springer Verlag, Berlin-Göttingen-Heidelberg.
WALD, A. (1947), Sequential analysis, John Wiley and Sons, New York.

WALKER, H. M. en J. LEV (1953), Statistical inference, Henry Holt and Company, New York.
WALLIN, J. F. E. en F. VAN DRIEL (1960), Het quotiënt van Von Neumann, Statistica Neerlandica 14, 51–56.
WALLIS, W. A. (1952), Rough-and-ready statistical tests, Industrial Quality Control 8, 35–40.
WALLIS, W. A. en H. V. ROBERTS (1956), Statistics, a new approach, The Free Press, Glencoe, Ill.
WALSH, J. E. (1949), Applications of some significance tests for the median which are valid under very general conditions, J. Am. Stat. Ass. 44, 342–355.
WELCH, B. L. (1947), The generalization of 'Student's' problem when several different population variances are involved, Biometrika 34, 28–35.
WEST, Q. M. (1951), The results of applying a simple random sampling process to farm management data, Cornell University.
WILCOXON, F. (1945), Individual comparisons by ranking methods, Biometrics Bull. 1, 80–83.
WILCOXON, F. (1949), Some rapid approximate statistical procedures, Stamford, Conn.: American Cyanamid Co.
WILKS, S. S. (1943), Mathematical statistics, Princeton University Press, Princeton.
WILKS, S. S. (1949), Elementary statistical analysis, Princeton University Press, Princeton, New Jersey.
WILSON Jr, E. B. (1952), An introduction to scientific research, McGraw-Hill Book Company, New York.
WILSON, E. B. en M. M. HILFERTY (1932), The distribution of chi-square, Nat. Acad. Sc. Proc. 24, 428–440.
WITTS, L. J. (1959), Medical surveys and clinical trials, Oxford University Press, London-New York-Toronto.
WOLFENDEN, H. H. (1954), Population statistics and their compilation, The University of Chicago Press, Chicago.
WOLFF, P. DE, Bedrijfsstatistiek, Samson N.V., Alphen.
WIJNGAARDEN, A. VAN (1950), Table of the cumulative symmetric binomial distribution, Proc. Kon. Ned. Akad. v. Wet. 53, 857–868 en Indag. Math. 12, 301–312.
WIJVEKATE, M. L. (1960), Verklarende statistiek, Aulaboeken, Het Spectrum, Utrecht.

YATES, F. (1933), The analysis of replicated experiments when the field results are incomplete, Empire J. Exp. Agr. 1, 129–142.
YATES, F. (1934), Contingency tables involving small numbers and the χ^2-test, Suppl. J. R. Stat. Soc. 1, 217–235.
YATES, F. (1948), The analysis of contingency tables with groupings based on quantitative characters, Biometrika 35, 176–181.
YATES, F. (1951), The influence of 'Statistical methods for research workers' on the development of the science of statistics, J. Am. Stat. Ass. 46, 19–34.
YATES, F. (1953), Sampling methods for censuses and surveys, Ch. Griffin & Comp., London, 2nd ed.
YERUSHALMY, J. et al. (1950), An evaluation of the role of serial chest roentgenograms in estimating the progress of disease in patients with pulmonary tuberculosis, American Review of Tuberculosis 61, 443.
YOUDEN, W. J. (1951), The Fisherian revolution in methods of experimentation, J. Am. Stat. Ass. 46, 47–50.
YULE G. U. en M. G. KENDALL (1946), An introduction to the theory of statistics, Ch. Griffin & Comp., London, 13th ed.