

INLEIDING TOT DE MEDISCHE STATISTIEK II

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

VERHANDELING VAN HET
NEDERLANDS INSTITUUT VOOR PRAEVENTIEVE GENEESKUNDE
XLVIII

INLEIDING TOT DE
MEDISCHE STATISTIEK

DEEL 2

KLASSIEKE METHODEN

DOOR

H. DE JONGE

1980

VOORWOORD

Met het verschijnen van dit tweede deel, gewijd aan de klassieke statistische methoden, waarin tevens de antwoorden op de opgaven en het register voor beide delen zijn opgenomen, is de 'Inleiding tot de Medische Statistiek' gereed gekomen.

Bij de in deel I behandelde verdelingsvrije methoden is dankbaar gebruik gemaakt van het vele belangrijke werk, dat door de Statistische Afdeling van het Mathematisch Centrum te Amsterdam onder leiding van wijlen Prof. dr. D. van Dantzig en Prof. dr. J. Hemelrijk op dit gebied is verricht. Ook de door deze afdeling ontwikkelde terminologie en symboliek heb ik zoveel mogelijk aangehouden. Bij de klassieke methoden in deel II diende ik voornamelijk terug te grijpen op buitenlandse en wel speciaal op de Engels-Amerikaanse statistische literatuur. Hierbij zijn vrijwel alle bereikbare hand- en leerboeken en vele oorspronkelijke artikelen geraadpleegd. Wat de terminologie betreft is gestreefd naar uniformiteit met deel I, terwijl verder zo mogelijk gebruik is gemaakt van termen, die in Nederland reeds zijn ingeburgerd.

De methoden die in beide delen zijn ondergebracht vormen niet het gehele statistische arsenaal, daar de meer geavanceerde technieken niet zijn behandeld of slechts zijn geïntroduceerd. Op de mathematische achtergronden is niet (of niet diep) ingegaan, maar de stof wordt op zodanige wijze gepresenteerd, dat desondanks een uitgesproken 'leer'-boek en geen 'lees'-boek is ontstaan. Overigens zal de tijd dienen te leren, welke aanvullingen de behandelde stof behoeft, respectievelijk welke beperkingen geboden zijn. Ik zal het zeer op prijs stellen wanneer de gebruikers mij hun suggesties dien-aangaande willen geven en mij ook van andere op- en aanmerkingen in kennis willen stellen.

Gaarne wil ik nogmaals dank zeggen aan allen, die aan het tot stand komen van dit boek hebben medegewerkt. Deze dank geldt in het bijzonder de Directie van het Nederlands Instituut voor Praeventieve Geneeskunde, welke mij in staat stelde aan het boek te werken en welke tevens bereid bleek de uitgave als 'Verhandeling' van het N.I.P.G. op zich te nemen.

Leiden, 31 Januari 1960.

H. DE JONGE

INHOUDSOPGAVE

DEEL C: KLASSIEKE METHODEN

13. STEEKPROEVEN UIT NORMAAL VERDEELDE POPULATIES

13.1. INLEIDING	339
13.2. DE KANSVERDELING VAN HET GEMIDDELDE	340
13.2.1. Normaal verdeelde populaties	340
13.2.2. Niet-normaal verdeelde populaties	341
13.2.3. Opgaven	343
13.3. DE KANSVERDELING VAN DE VARIANTIE	343
13.3.1. Kleine steekproeven ($n \leq 100$)	343
13.3.2. Grote steekproeven	345
13.3.3. De simultane verdeling van \bar{x} en s_x	346
13.3.4. Opgaven	346
13.4. STUDENT-VERDELINGEN	347
13.4.1. Definitie en toelichting	347
13.4.2. Fractielen van STUDENT-verdelingen: tabel C	348
13.4.3. Opgaven	348
13.5. F -VERDELINGEN	349
13.5.1. Definitie en toelichting	349
13.5.2. Fractielen van F -verdelingen: tabel F	349
13.5.3. Opgaven	351
13.6. DE VERDELINGEN VAN \underline{F} , \underline{t} , \underline{T} EN χ^2	352
13.7. DE KANSVERDELING VAN HET VERSCHIL TUSSEN TWEE GEMIDDELDEN	354
13.7.1. Populatievarianties bekend	354
13.7.2. Populatievarianties onbekend	355
13.7.3. Opgaven	356
13.8. EEN STEEKPROEFEXPERIMENT	357
13.8.1. Inleiding en beschrijving	357
13.8.2. De waargenomen verdeling van het gemiddelde	360
13.8.3. De waargenomen verdeling van de variantie	360
13.8.4. De waargenomen simultane verdeling van het gemiddelde en de spreiding	362
13.8.5. De waargenomen verdelingen van de grootheden \underline{T} en \underline{t}	363
13.8.6. De waargenomen verdeling van de grootheid \underline{F}	363

14. TECHNIEKEN BETREFFENDE VARIANTIES

14.1. TOETS VOOR DE HYPOTHESE $\sigma^2 = \sigma_0^2$	365
14.1.1. Kleine steekproeven ($n \leq 100$)	365
14.1.2. Grote steekproeven	367
14.2. BETROUWBAARHEIDSGRENZEN VAN DE VARIANTIE EN DE SPREIDING.	367

14.2.1.	Kleine steekproeven ($n \leq 100$)	367
14.2.2.	Grote steekproeven	370
14.2.3.	Toetsen en schatten	371
14.3.	TOETS VOOR DE HYPOTHESE $\sigma_1^2 = \sigma_2^2$	371
14.4.	TOETSEN VOOR HET VERGELIJKEN VAN MEER DAN TWEE VARIANTIES	373
14.4.1.	De toets van BARTLETT	373
14.4.2.	De toets van HARTLEY	376
14.4.3.	De toets van COCHRAN	377
14.4.4.	De toets van DOORNBOS	378
14.5.	DE TOETS TEGEN VERLOOP VAN BELLINSON	378
14.6.	OPGAVEN	381

15. TECHNIEKEN BETREFFENDE HET GEMIDDELDE VAN EEN POPULATIE

15.1.	TOETS VOOR DE HYPOTHESE $\mu = \mu_0$	382
15.1.1.	Populatiespreiding bekend	382
15.1.2.	Populatiespreiding onbekend	383
15.1.3.	Onderscheidingsvermogen en steekproefomvang	385
15.1.4.	Opgaven	390
15.2.	BETROUWBAARHEIDSGRENZEN VAN HET GEMIDDELDE	390
15.2.1.	Schattingsinterval van μ , als σ bekend is	391
15.2.2.	Schattingsinterval van μ , als σ onbekend is	394
15.2.3.	De keuze van de steekproefomvang	395
15.2.4.	Opgaven	396
15.3.	TOEPASSING OP HET GEMIDDELDE VAN VERSCHILLEN TUSSEN GEPAARDE WAARNEMINGEN	396
15.3.1.	De toets van STUDENT op verschillen	396
15.3.2.	Betrouwbaarheidsgrenzen van het gemiddelde verschil	398
15.3.3.	Opgaven	398
15.4.	DE NAUWKEURIGHEID VAN WAARNEMINGEN	399
15.4.1.	Metingen	399
15.4.2.	Tellingen	403
15.4.3.	Opgaven	408
15.5.	HET SIGNALEREN VAN UITBIJTERS	408
15.5.1.	Inleiding	408
15.5.2.	De uitbijtertoets van DIXON	409
15.5.3.	Andere uitbijtertoetsen	410

16. TECHNIEKEN BETREFFENDE DE GEMIDDELDEN VAN TWEE POPULATIES

16.1.	TOETS VOOR DE HYPOTHESE $\mu_1 = \mu_2$ ALS σ_1 EN σ_2 BEKEND ZIJN	411
16.1.1.	Tweezijdige toetsing	411

16.1.2.	Eénzijdige toetsing.	412
16.1.3.	De keuze van de omvang der steekproeven	412
16.2.	TOETS VOOR DE HYPOTHESE $\mu_1 = \mu_2$ ALS σ_1 EN σ_2 ONBEKEND, MAAR GELIJK ZIJN	413
16.2.1.	Tweezijdige toetsing: de toets van STUDENT. . . .	413
16.2.2.	Tweezijdige toetsing: de toets van LORD	416
16.2.3.	Eénzijdige toetsing: de toets van STUDENT. . . .	416
16.3.	TOETS VOOR DE HYPOTHESE $\mu_1 = \mu_2$ ALS σ_1 EN σ_2 ONBEKEND EN ONGELIJK ZIJN	417
16.3.1.	Kleine steekproeven ($n_1 \leq n_2 \leq 100$).	417
16.3.2.	Grote steekproeven	418
16.3.3.	Oorzaken van ongelijkheid van varianties.	419
16.4.	TOETS VOOR DE HYPOTHESE $\mu_1 = \mu_2$ EN $\sigma_1 = \sigma_2$	419
16.5.	BETROUWBAARHEIDSGRENZEN VAN $\mu_1 - \mu_2$	420
16.5.1.	Populatievarianties bekend	420
16.5.2.	Populatievarianties onbekend	421
16.5.3.	De keuze van de omvang der steekproeven	422
16.6.	OPGAVEN	423

17. TECHNIEKEN BETREFFENDE DE GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

17.1.	ENKELVOUDIGE ANALYSE: k ASELECTE STEEKPROEVEN VAN DEZELFDE OMVANG	425
17.1.1.	Onderstellingen en nulhypothese.	425
17.1.2.	De toetsingsgrootheid en haar verdeling onder H_0 .	425
17.1.3.	Uitvoering van de toetsing	427
17.1.4.	Toetsing via de spreidingsbreedte	431
17.1.5.	Voortzetting van de analyse als H_0 verworpen wordt	431
17.2.	ENKELVOUDIGE ANALYSE: k ASELECTE STEEKPROEVEN VAN VERSCHILLENDE OMVANG	436
17.2.1.	Uitvoering van de analyse	436
17.2.2.	Betrouwbaarheidsgrenzen van één gemiddelde . .	437
17.2.3.	De enkelvoudige analyse toegepast op twee steek- proeven	438
17.3.	TWEEVOUDIGE VARIANTIE-ANALYSE OP k VERWANTE STEEK- PROEVEN	439
17.3.1.	Inleiding	439
17.3.2.	Onderstellingen en nulhypothesen	440
17.3.3.	Uitvoering van de toetsingen	442
17.3.4.	De residuele variantie	447
17.3.5.	De tweevoudige analyse toegepast op gepaarde waarnemingen	449
17.4.	VERWANTE STEEKPROEVEN MET REPLICATIES	450
17.4.1.	Onderstellingen en nulhypothesen	450
17.4.2.	Berekeningen	451
17.4.3.	De interactie-variantie	454
17.4.4.	De variantie binnen de steekproeven.	455

17.4.5. Uitvoering van de toetsingen	455
17.5. OPGAVEN	460
18. REGRESSIE EN CORRELATIE	
18.1. TWEEDIMENSIONALE VERDELINGEN	464
18.2. REGRESSIE.	466
18.2.1. Regressieproblemen	466
18.2.2. Berekening van een regressielijn volgens de methode der kleinste kwadraten	473
18.2.3. Berekening van een best-passende kromme volgens de methode der kleinste kwadraten	477
18.3. CORRELATIE	479
18.3.1. Correlatieproblemen	479
18.3.2. De covariantie en de correlatiecoëfficiënt	482
18.4. OPGAVEN	486
19. REGRESSIE-ANALYSE	
19.1. LINEAIR-NORMALE REGRESSIE	488
19.1.1. Beschrijving van het model	488
19.1.2. Schatting van A , B en $\sigma_{y,x}^2$ uit een steekproef.	489
19.1.3. Toetsen en intervallschattingen	497
19.1.4. Schatting van B als $A = 0$	502
19.1.5. Toets op de rechtlijnigheid van de regressie	503
19.1.6. Vereenvoudiging van de berekeningen	508
19.2. UITBREIDING VAN DE REGRESSIE-ANALYSE	511
19.3. OPGAVEN	512
20. CORRELATIEREKENING	
20.1. TWEEDIMENSIONAAL NORMALE CORRELATIE	514
20.1.1. Beschrijving van het model	514
20.1.2. Schattingen uit een steekproef.	517
20.1.3. Toetsing van de hypothese $\rho = 0$	521
20.1.4. Toetsing en schattingen indien $\rho \neq 0$: de z -trans- formatie	524
20.1.5. Toetsing van de hypothese $\rho_1 = \rho_2$	526
20.1.6. Toetsing van de hypothese $\rho_1 = \rho_2 = \dots = \rho_k$	527
20.1.7. Berekening van r uit een correlatietabel	528
20.1.8. Enkele opmerkingen betreffende de correlatie- coëfficiënt	531
20.2. OPGAVEN	536
20.3. DE CORRELATIECOËFFICIËNT ALS TOETSINGSGROOTHEID	539
20.4. FUNCTIONELE RELATIES	543
21. VERDELINGSVRIJE EN KLASSIEKE METHODEN	
21.1. INLEIDING	546
21.2. KLASSIEKE EN VERDELINGSVRIJE METHODEN	546
21.3. DE KEUZE VAN EEN KLASSIEKE OF VERDELINGSVRIJE METHODE	549

21.4.	NORMALITEITSTOETSEN	553
21.4.1.	Normaal waarschijnlijkheidspapier	553
21.4.2.	De χ^2 -toets voor aanpassing.	553
21.4.3.	De toets van KOLMOGOROV-SMIRNOV.	553
21.4.4.	De normaliteitstoetsen van FISHER	553
21.4.5.	Combinatie van de toetsen van FISHER.	557
21.4.6.	De toets tegen kurtosis van GEARY	558
21.4.7.	Scheefheidstoets gebaseerd op kwartielen	558
21.5.	TRANSFORMATIES	558
21.5.1.	De logaritmische transformatie.	559
21.5.2.	De vierkantswortel transformatie	559
21.5.3.	De reciproke transformatie	559
21.5.4.	De bgsin transformatie.	559
21.5.5.	Transformatie van χ^2 , F - en t -verdelingen	561
22.	STATISTIEK EN PROEFOPZET	
22.1.	INLEIDING.	562
22.2.	STATISTISCHE PROEFSHEMA'S BIJ EXPERIMENTEN	566
22.2.1.	Klassieke en factoriële proefopzetten	567
22.2.2.	Een 2^3 factorieel proefschema zonder replicaties	568
22.2.3.	Een 2^3 factorieel proefschema met replicaties	571
22.2.4.	Een $3 \times 4 \times 5$ factorieel schema	576
22.2.5.	Latijnse en Grieks-Latijnse vierkanten	579
22.2.6.	Opmerkingen	583
22.3.	SPECIALE STEEKPROEFTECHNIEKEN BIJ ONDERZOEKINGEN	585
22.3.1.	Laagsgewijs aselechte steekproeven	585
22.3.2.	Steekproeven bestaande uit groepen	586
23.	STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE	
23.1.	INLEIDING.	588
23.2.	DIRECTE STANDAARDISATIE	589
23.3.	INDIRECTE STANDAARDISATIE MET KWANTITATIEVE RES- PONSIES	593
23.3.1.	Inleiding	593
23.3.2.	Een 'parallel line assay'	594
23.3.3.	Een 'slope ratio assay'	600
23.4.	INDIRECTE STANDAARDISATIE GEBASEERD OP HET AL DAN NIET REAGEREN: ED_{50} BEPALINGEN	604
23.4.1.	Inleiding	604
23.4.2.	De probit-analyse	605
23.4.3.	De grafische methode van LITCHFIELD en WILCOXON	605
23.4.4.	De regressie-analyse na toepassing van de bgsin transformatie	606
23.4.5.	De methode van SPEARMAN en KÄRBER	607
23.4.6.	De methode der voortschrijdende gemiddelden van THOMPSON	611
23.4.7.	De sequente methode van DIXON en MOOD	613

24. ENKELE BIJZONDERE TECHNIEKEN

24.1. DE TOETS VAN YATES	618
24.1.1. Toepassing op een $2 \times r$ tabel	618
24.1.2. Vergelijking met de toets van WILCOXON	620
24.1.3. Toepassing op een $k \times r$ tabel	621
24.1.4. Vergelijking met de toets van KRUSKAL en WALLIS	623
24.1.5. De 2×2 tabel	624
24.2. VERDERE ANALYSE VAN EEN $K \times R$ TABEL	624
24.3. TOETS VOOR DE HYPOTHESE $\mu_A - \mu_B = \mu_C - \mu_D$	629

BIJLAGEN

I. TABELLENVERZAMELING (<i>vervolg</i>)	635
III. LITERATUUROPGAVE (<i>vervolg</i>)	650
IV. ANTWOORDEN BIJ DE OPGAVEN (<i>deel I en II</i>)	655
V. REGISTER	676

C

KLASSIEKE METHODEN

De mathematicus, die zich in een latere levensphase aan de statistiek is gaan wijden, kan soms met weemoed terugdenken aan de goede oude tijd, toen hij nog slechts één geweten had. Zijn wiskundig geweten blijft onbelast als hij maar zorgt volgens de daarvoor vaststaande wetten te redeneren en hierbij geen fouten te maken. Iedere fout is door hem of door anderen met zekerheid aan te tonen.

Laat de mathematisch statisticus echter alleen zijn wiskundig geweten spreken, dan kan hij zijn wetenschap nooit op reële problemen toepassen. Hij moet dan nl. altijd onderstellingen maken, bv. de onderstelling dat verschillende waarnemingen onderling onafhankelijk zijn of dat bepaalde waarschijnlijkheidsverdelingen in de loop van de tijd onveranderd blijven. Vraagt hij zich dan af, zoals ook HARRIET FREEZER: 'Is dat nu wel zo?', dan luidt het antwoord dikwijls ontkennend of, in het gunstigste geval, weet hij het antwoord niet. Zijn wiskundig, of, algemener gesproken, zijn wetenschappelijk geweten zou hem dus verbieden, deze onderstellingen te maken, en zijn antwoord op de gestelde vraag zou bijna altijd moeten luiden: 'Ik weet het niet' of 'Men kan niets concluderen'. De statisticus heeft echter ook een maatschappelijk geweten, dat hem gebiedt, onderzoekers op velerlei gebied althans zo goed mogelijk te helpen, zij het zonder volstrèkte zekerheid, dat zijn conclusies juist zijn. Het maatschappelijk geweten zegt hem dus, niet al te kritisch te zijn, maar genoeg te nemen met een redelijke mate van zekerheid. Hij moet een compromis sluiten tussen zijn wetenschappelijk en zijn maatschappelijk geweten; de statisticus is dus in de kern gespleten: hij lijdt aan kernsplitsing.

Prof. Dr D. VAN DANTZIG (199)

Tabel 13.1. Overzicht van de voornaamste klassieke methoden in de hoofdstukken 14 t/m 17

Methoden gebaseerd op*	Methoden betreffende				
	Gemiddelden	Spreidingen (varianties)			
Eén aselechte steekproef	15.1.	Toets van STUDENT voor de hypothese $\mu = \mu_0$	14.1.	Toets voor de hypothese $\sigma^2 = \sigma_0^2$	
	15.2.	Betrouwbaarheidsgrenzen van μ	14.2.	Betrouwbaarheidsgrenzen van σ (σ^2)	
Aselechte steekproeven	2	16.2.	Toets van STUDENT voor de hypothese $\mu_1 = \mu_2$, indien $\sigma_1 = \sigma_2$	14.3.	F -toets van FISHER voor de hypothese $\sigma_1^2 = \sigma_2^2$
		16.3.	Toets voor de hypothese $\mu_1 = \mu_2$, indien $\sigma_1 \neq \sigma_2$		
		16.4.	Toets voor de hypothese $\mu_1 = \mu_2$ en $\sigma_1 = \sigma_2$		
		16.5.	Betrouwbaarheidsgrenzen van $\mu_1 - \mu_2$		
	k	Enkelvoudige variantie-analyse: toets voor de hypothese $\mu_1 = \mu_2 = \dots = \mu_k$		Toetsen voor de hypothese $\sigma_1 = \sigma_2 = \dots = \sigma_k$	
		17.1.	Steekproeven van dezelfde omvang	14.4.2.	Steekproeven van dezelfde omvang — toets van HARTLEY
	17.2.	Steekproeven van verschillende omvang	14.4.1.	Steekproeven van verschillende omvang — toets van BARTLETT	
			14.4.3.	Toets van COCHRAN	
			14.4.4.	Toets van DOORNBOS	
Verwante steekproeven	2	15.3.1.	Toets van STUDENT voor de hypothese $\mu_v = 0$	Zie 2 aselechte steekproeven	
		15.3.2.	Betrouwbaarheidsgrenzen van μ_v		
	k	17.3.	Tweevoudige variantie-analyse	Zie k aselechte steekproeven	
		17.4.	Tweevoudige variantie-analyse met replicaties		

* Zie voor deze indeling 8.2.2, blz. 170.

STEEKPROEVEN UIT NORMAAL VERDEELDE POPULATIES

13.1. Inleiding

In deel I van dit boek zijn eerst een aantal statistische begrippen en vervolgens verschillende verdelingsvrije methoden behandeld. Dit deel is gewijd aan de zg. *klassieke* statistische methoden, dat zijn toetsings- en schattingstechnieken betreffende het gemiddelde en de spreiding (standaarddeviatie) van de populatieverdelingen van (pseudo) normale grootheden.

Bij de afleiding van de klassieke methoden wordt ondersteld, dat de bestudeerde grootheden op de populatie exact normaal verdeeld zijn. De normale verdeling is echter een mathematische conceptie en precies normale verdelingen treft men dan ook in werkelijkheid weinig (of wellicht in 't geheel niet) aan. Vele grootheden blijken echter een verdeling te bezitten, die betrekkelijk onbelangrijke afwijkingen van het normale type vertoont. Andere grootheden blijken na het toepassen van een transformatie (bv. door het nemen van logaritmen) bij benadering normaal te worden. In dergelijke gevallen kan men het 'normale model' zonder veel bezwaar gebruiken en dus de klassieke methoden toepassen. Wanneer er echter duidelijke afwijkingen van de normaliteit aanwezig zijn, kan men gewoonlijk beter een verdelingsvrije methode kiezen, omdat anders het gevaar niet denkbeeldig is, dat men een exact lijkende, maar verre van juiste uitkomst verkrijgt. Om de lezer behulpzaam te zijn bij een verantwoorde keuze tussen een klassieke en een verdelingsvrije methode komen wij later nog terug op de factoren, die hierbij een rol spelen.

In dit hoofdstuk behandelen wij de kansverdelingen van verschillende steekproeffuncties bij aselechte trekking van een steekproef van n elementen uit een normaal verdeelde populatie met parameters μ_{α} en σ_{α} . Wij zien hierbij af van het geven van bewijzen¹, maar vermelden in plaats daarvan de uitkomsten van een steekproefexperiment, dat zich leent voor klassikale uitvoering en waarmee wij de lezer hopen te demonstreren, dat deze theoretisch verkregen kansverdelingen een reële betekenis bezitten. De hierop volgende hoofdstukken

¹ Deze bewijzen vergen vrij veel plaatsruimte, terwijl de ervaring ons heeft geleerd dat zij het merendeel der lezers niet interesseren. Daar zij in elk leerboek der mathematische statistiek te vinden zijn [zie bv. BENNETT & FRANKLIN (4), HOEL (169), MOOD (24), WILKS (208), KENDALL (173)], hebben wij gemeend deze ruimte beter te kunnen vullen met voorbeelden en opgaven, die zoveel mogelijk aan de praktijk zijn ontleend.

(14 t/m 20) bevatten de belangrijkste technieken voor het toetsen van hypothesen en het geven van schattingsintervallen betreffende gemiddelden en spreidingen. Tabel 13.1 geeft een overzicht van de methoden, die in de hoofdstukken 14 t/m 17 behandeld worden (men bestudere deze tabel nogmaals, wanneer men deze hoofdstukken heeft doorgewerkt: zij geeft dan weer de mogelijkheid om snel een bepaalde methode te localiseren). Hoofdstuk 18 bevat een inleiding tot de klassieke methoden betreffende de analyse van waarnemingsparen, die vervolgens in de hoofdstukken 19 (regressie-analyse) en 20 (correlatierekening) verder worden uitgewerkt. Hoofdstuk 21 betreft de reeds genoemde keuze tussen verdelingsvrije en klassieke methoden en in hoofdstuk 22 beschouwen wij enkele meer gecompliceerde proefopzetten, gebaseerd op de variantie-analyse, die het mogelijk maken om bij een experiment op efficiënte wijze een maximum aan informatie te verkrijgen. Tenslotte wordt in hoofdstuk 23 een overzicht gegeven van het gebruik van statistische methoden bij de biologische standaardisatie van geneesmiddelen, insecticiden e.d.

Ook bij de behandeling van de klassieke methoden hebben wij ons beperkingen moeten opleggen. Deze bestaan vnl. hierin, dat de variantie-, regressie- en correlatie-analyse slechts worden ingeleid en niet in al hun vertakkingen worden besproken. Verder wijzen wij erop, dat een aantal paragrafen (of delen daarvan) met * zijn aangeduid. Zij hebben betrekking op onderwerpen of technieken, die bij eerste lezing zonder bezwaar kunnen worden overgeslagen.

13.2. De kansverdeling van het gemiddelde

13.2.1. NORMAAL VERDEELDE POPULATIES

Beschouw een populatie, waarvan elk element een waarde bezit van de grootheid x en onderstel dat deze grootheid op de populatie normaal verdeeld is met gemiddelde μ_x en spreiding σ_x . Trekt men uit deze populatie aselekt een steekproef van n elementen, dan blijkt de steekproeffunctie $\bar{x} = \sum_{i=1}^n x_i/n$ een kansverdeling te volgen, die eveneens normaal is met gemiddelde

$$(13.1.1) \quad \mu_{\bar{x}} = \mu_x, \quad \text{en met variantie}$$

$$(13.2.1) \quad \sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n}, \quad \text{dus met spreiding}$$

$$(13.3.1) \quad \sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}.$$

Uit het voorgaande volgt, dat de grootheid

$$(13.4) \quad T = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu_x}{\sigma_x/\sqrt{n}} = \frac{(\bar{x} - \mu_x)\sqrt{n}}{\sigma_x}$$

standaardnormaal verdeeld is.

Men noemt de kansverdeling van de grootheid \bar{x} ook wel de steekproefverdeling van het gemiddelde (Eng. sampling distribution of the mean) en de spreiding van deze verdeling, $\sigma_{\bar{x}}$, de *standaardfout* (Eng. standard error) van het gemiddelde. Uit (13.3.1) blijkt, dat deze spreiding afneemt, naarmate n toeneemt. Illustratief is in dit verband de volgende tabel, waarin voor een populatieverdeling met $\sigma_x = 10$ de waarde van $\sigma_{\bar{x}}$ is opgenomen bij verschillende waarden van n :

n	2	3	4	9	16	25	64	100	256	1024	2500	4096
$\sigma_{\bar{x}}$	7,07	5,77	5	3,33	2,5	2	1,25	1	0,625	0,3125	0,2	0,15625

Voorbeeld

13.1. Beschouw een grootheid x , waarvan de populatieverdeling normaal is met gemiddelde $\mu_x = 20$ en spreiding $\sigma_x = 4$. Hoe groot is de kans, dat één aselekt getrokken element een waarde van x aanneemt tussen a) 18 en 22, b) 17 en 23?

$$\begin{array}{l}
 \left. \begin{array}{l}
 a) \quad x_1 = 18, T_1 = \frac{18 - 20}{4} = -0,5 \\
 \quad \quad x_2 = 22, T_2 = \frac{22 - 20}{4} = +0,5
 \end{array} \right\} \begin{array}{l}
 \text{Uit tabel A volgt:} \\
 P(18 < x < 22) \\
 = P(-0,5 < T < 0,5) \\
 = 1 - 2(0,3085) = 0,3830.
 \end{array} \\
 \\
 \left. \begin{array}{l}
 b) \quad x_1 = 17, T_1 = \frac{17 - 20}{4} = -0,75 \\
 \quad \quad x_2 = 23, T_2 = \frac{23 - 20}{4} = +0,75
 \end{array} \right\} \begin{array}{l}
 \text{Uit tabel A volgt:} \\
 P(17 < x < 23) \\
 = P(-0,75 < T < 0,75) \\
 = 1 - 2(0,2266) = 0,5468.
 \end{array}
 \end{array}$$

Uit deze populatie trekt men aselekt een steekproef van 16 elementen. Hoe groot is de kans, dat het gemiddelde van deze steekproef een waarde aanneemt tussen c) 18 en 22, d) 17 en 23?

Voor $n = 16$ volgt de grootheid \bar{x} een normale verdeling met $\mu_{\bar{x}} = \mu_x = 20$ en $\sigma_{\bar{x}} = \sigma_x/\sqrt{n} = 4/\sqrt{16} = 1$, zodat

$$\begin{array}{l}
 \left. \begin{array}{l}
 c) \quad \bar{x}_1 = 18, T_1 = \frac{18 - 20}{1} = -2 \\
 \quad \quad \bar{x}_2 = 22, T_2 = \frac{22 - 20}{1} = +2
 \end{array} \right\} \begin{array}{l}
 \text{Uit tabel A blijkt, dat} \\
 P(18 < \bar{x} < 22) \\
 = P(-2 < T < 2) \\
 = 1 - 2(0,0228) = 0,9544.
 \end{array} \\
 \\
 \left. \begin{array}{l}
 d) \quad \bar{x}_1 = 17, T_1 = \frac{17 - 20}{1} = -3 \\
 \quad \quad \bar{x}_2 = 23, T_2 = \frac{23 - 20}{1} = +3
 \end{array} \right\} \begin{array}{l}
 \text{Uit tabel A blijkt, dat} \\
 P(17 < \bar{x} < 23) \\
 = P(-3 < T < 3) \\
 = 1 - 2(0,00135) = 0,9973.
 \end{array}
 \end{array}$$

13.2.2. NIET-NORMAAL VERDEELDE POPULATIES

Wij beschouwen vervolgens de verdeling van een grootheid x met gemiddelde μ_x en spreiding σ_x op een populatie van N elementen, die *niet normaal* is. Trekt men uit deze populatie aselekt een steekproef

van n elementen, dan blijkt voor de kansverdeling van \bar{x} te gelden:

bij trekking met teruglegging

bij trekking zonder teruglegging

$$(13.1.2)$$

$$\mu_{\bar{x}} = \mu_{\alpha}$$

$$(13.2.2) \quad \sigma_{\bar{x}}^2 = \frac{\sigma_{\alpha}^2}{n}$$

$$(13.2.3) \quad \sigma_{\bar{x}}^2 = \frac{\sigma_{\alpha}^2}{n} \cdot \frac{N-n}{N-1}$$

of

$$(13.3.2) \quad \sigma_{\bar{x}} = \frac{\sigma_{\alpha}}{\sqrt{n}}$$

$$(13.3.3) \quad \sigma_{\bar{x}} = \frac{\sigma_{\alpha}}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Wanneer de steekproefomvang n klein is met betrekking tot de populatie-omvang N , is $N-n \approx N-1$, zodat $\sqrt{(N-n)/(N-1)} \cong 1$. In dat geval kan men dus, ook als zonder teruglegging getrokken wordt, gebruik maken van (13.3.2), resp. (13.2.2). Daar in de praktijk bij het trekken zonder teruglegging gewoonlijk aan de voorwaarde $n \ll N$ voldaan wordt, laten wij (13.2.3) en (13.3.3) verder buiten beschouwing.

Als de populatieverdeling niet normaal is, volgt de grootheid \bar{x} geen normale verdeling. Bij *toenemende steekproefomvang* nadert echter de verdeling van \bar{x} tot een normale verdeling, zodat ook dan de verdeling van

$$(13.5) \quad \underline{T}' = \frac{(\bar{x} - \mu_{\alpha}) \sqrt{n}}{\sigma_{\alpha}}$$

tot de standaardnormale verdeling nadert. Bij pseudo-normaal verdeelde populaties blijkt de grootheid \underline{T}' reeds bij kleine steekproefomvang ($n > 10$) vrijwel standaardnormaal te zijn. Voor verdelingen, die wel ééntoppig en ongeveer symmetrisch, maar 'spitser' of 'vlakker' zijn dan de normale verdeling en voor scheve verdelingen verloopt de benadering trager, terwijl zij het sterkst wordt vertraagd als uitbijters in één richting voorkomen. WEST (207) voerde een experiment uit, waarbij hij een groot aantal steekproeven van 100 elementen trok uit de volgende populatie met een uitgesproken scheefheid naar rechts en uitbijters in dezelfde richting:

x	0-29	30-63	64-97	98-131	132-165	166-199	200-233	234-267	268-301	302-335	336-369	370-403	404-437	438-471	472-505	Totaal
f	47	143	154	82	62	33	13	6	4	6	2	—	2	—	2	556

Bij deze steekproefomvang bleek de waargenomen verdeling van \bar{x} niet significant van een normale verdeling (met gemiddelde μ_{α} en spreiding $\sigma_{\alpha}/10$) te verschillen.¹

¹ COCHRAN (51) geeft aan, dat voor ééntoppige, scheve verdelingen de verdeling van \bar{x} bij benadering normaal is, indien $n > 25g_1$, waarin g_1 een door FISHER ingevoerde maat voor de scheefheid is, die wij in 21.2 bespreken. Volgens dit criterium vindt hij voor de door WEST gebruikte populatieverdeling $n > 90$, een uitkomst die met de empirische bevinding van laatstgenoemde overeenstemt.

13.2.3. OPGAVEN

13.1. In de voorbeelden 3.17 en 3.18 (blz. 74) is voor een populatie met $N = 6$ en $x = 1, 2, 3, 4, 6$ en 8 de kansverdeling van \bar{x} berekend bij aselechte trekking van een steekproef met $n = 2$ en bij trekking zonder, resp. met teruglegging. Verifieer voor deze kansverdelingen de formules (13.1) en (13.2).

13.2. De verdeling van het brutogewicht van een populatie van vliegtuig-passagiers is enigszins scheef naar rechts met een gemiddelde van 70 kg en een spreiding van 2,5 kg. Hoe groot is bij benadering de kans, dat de volledige passagiersbezetting van een bepaald type vliegtuig, bestaande uit 25 tot deze populatie behorende personen, een totaal brutogewicht van 1782 kg overschrijdt?

13.3. Ontleend aan het examen statistisch analyst, 1955:

In een magazijn ligt een groot aantal platte schijven, waarvan de dikte (ongeveer) normaal verdeeld is met een gemiddelde van 12 cm en een standaarddeviatie van 2 cm. Het magazijn is binnenwerks 320 cm hoog. Men heeft de gewoonte 25 schijven op elkaar te stapelen. Bij welk gedeelte van de stapels zal het niet gelukken deze compleet te maken, aangenomen dat de schijven vóór het stapelen grondig door elkaar liggen?

13.3. De kansverdeling van de variantie

13.3.1. KLEINE STEEKPROEVEN ($n \leq 100$)

Trekt men aselekt een steekproef uit een willekeurig verdeelde populatie, dan blijkt

$$s_x^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

een zuivere schatting te zijn van de populatievariantie σ_x^2 . D.w.z. het gemiddelde van de kansverdeling van de grootheid s_x^2 is gelijk aan σ_x^2 . Wij volstaan ermee, de juistheid van deze bewering met een voorbeeld te illustreren en beschouwen daartoe de populatie in voorbeeld 3.17, bestaande uit 6 fiches met de nummers 1, 2, 3, 4, 6 en 8. Zoals men zelf kan uitrekenen, is hiervoor $\sigma^2 = 17/3$. Op blz. 74 zijn de 36 steekproeven van 2 elementen, die met teruglegging uit deze populatie kunnen worden getrokken, in een tabel opgenomen. Berekent men van elke steekproef s^2 , dan blijkt de kansverdeling van s^2 te zijn:

s^2	0	0,5	2	4,5	8	12,5	18	24,5	Totaal
f	6	6	8	4	4	4	2	2	36
$P (s^2 = \sigma^2)$	6/36	6/36	8/36	4/36	4/36	4/36	2/36	2/36	1
fs^2	0	3	16	18	32	50	36	49	204

Zoals uit de laatste regel van deze tabel blijkt, is het gemiddelde van de kansverdeling gelijk aan $204/36 = 17/3 = \sigma^2$. Het voorgaande gaat alleen op, als men bij de berekening van s^2 deelt door $n - 1$; deelt men door n , dan is de verkregen steekproefvariantie geen zuivere schatting meer van σ^2 , maar systematisch te laag. Men noemt $n - 1$

het aantal *vrijheidsgraden* van de variantie. De betekenis van deze uitspraak is duidelijk als men bedenkt, dat de n verschillen $x_i - \bar{x}$ niet volledig onafhankelijk van elkaar zijn, daar $\sum (x_i - \bar{x}) = 0$. Zijn er dus $n - 1$ verschillen bepaald, dan ligt ook het n -de verschil vast.

Tussen de variantie van de kansverdeling van s^2 en de populatievariantie σ^2 bestaat *geen* eenvoudige relatie zoals bij het gemiddelde, waar immers $\sigma_{\bar{x}}^2 = \sigma^2/n$. Voor aselechte steekproeven uit een *normaal verdeelde* populatie volgt echter de grootheid

$$(13.6) \quad \chi^2 = \frac{(n-1) \underline{s}_x^2}{\sigma_x^2} = \frac{\sum (x_i - \bar{x})^2}{\sigma_x^2}$$

een χ^2 -verdeling met $n - 1$ vrijheidsgraden. Is de populatieverdeling van de grootheid x pseudonormaal, dan geldt (13.6) bij benadering. Deze benadering is beter, naarmate de afwijkingen van de normaliteit kleiner zijn en de steekproefomvang groter is. Hier is het dus een functie van de grootheid \underline{s}_x^2 , nl. $\frac{n-1}{\sigma^2} \cdot \underline{s}_x^2$, die een bekende theoretische verdeling volgt en wel de χ^2 -verdeling die wij reeds in hoofdstuk 9 hebben geïntroduceerd. De kansverdeling van deze functie is dus *niet* normaal, ook al is de populatieverdeling dit wel. Zoals echter in 13.3.2 zal blijken, nadert de kansverdeling van de spreiding \underline{s}_x bij toenemende $\nu = n - 1$ tot een normale verdeling.

Ten behoeve van de in hoofdstuk 14 behandelde toetsings- en schattingstechnieken betreffende varianties zijn in dit deel als tabel T een aantal fractielen van χ^2 -verdelingen met $\nu = 1$ (1) 30 en $\nu = 30$ (10) 100 opgenomen. Verder vermelden wij, dat men voor het berekenen van kansen betreffende s^2 soms nuttig gebruik kan maken van de WILSON-HILFERTY transformatie. Deze auteurs hebben be-

wezen, dat de grootheid $\sqrt[3]{\frac{\chi^2}{\nu}}$ bij benadering normaal verdeeld is

met $\mu = 1 - \frac{2}{9\nu}$ en $\sigma^2 = \frac{2}{9\nu}$, zodat de grootheid

$$(13.7) \quad \underline{T} = \frac{\sqrt[3]{\frac{\chi^2}{\nu}} - \left(1 - \frac{2}{9\nu}\right)}{\sqrt{\frac{2}{9\nu}}} = 3 \left\{ \sqrt[3]{\frac{\chi^2}{\nu}} + \frac{2}{9\nu} - 1 \right\} \sqrt{\frac{\nu}{2}}$$

bij benadering standaardnormaal verdeeld is. Deze transformatie levert gewoonlijk een zeer nauwkeurige benadering op. Dit blijkt, als men haar uitkomsten vergelijkt met exact bekende kansen van χ^2 -verdelingen. Zo vindt men bv. in tabel T, dat $\chi^2_{0.10} (\nu = 9) = 4,168$, d.w.z.

$P_L(\chi^2 = 4,168) = 0,10$. Met de transformatie vindt men

$$T = 3 \{ \sqrt[3]{4,168/9} + 2/(9 \times 9) - 1 \} \sqrt{9/2} = 3(0,7737 + 0,0247 - 1) \sqrt{4,5} \\ = 3(-0,2016)(2,1213) = -1,28,$$

en tabel A geeft dat $P_L(T = -1,28) = 0,1003$.

Voorbeeld

13.2. Beschouw een normaal verdeelde populatie met $\sigma^2 = 12$. Men trekt hieruit aselekt een steekproef van 10 elementen. Hoe groot is de kans, dat de variantie van deze steekproef, s^2 , groter is dan 20? Door substitutie in (13.6) vindt men: $\chi^2(\nu = 9) = \frac{(10-1)20}{12} = 15$. Uit

tabel T blijkt, dat voor de χ^2 -verdeling met 9 vrijheidsgraden $\chi^2_{0,90} = 14,684$ en $\chi^2_{0,95} = 16,919$, zodat de gevraagde overschrijdingskans is:

$$0,10 > P_R(\chi^2 = 15) > 0,05.$$

Deze kans kan met behulp van de transformatie van WILSON en HILFERTY nauwkeuriger worden benaderd. Men vindt dan met (13.7):

$$T = 3 \{ \sqrt[3]{15/9} + 2/(9 \times 9) - 1 \} \sqrt{9/2} \\ = 3(1,1855 + 0,0247 - 1) \sqrt{4,5} = 1,34,$$

en daar $P_R(T = 1,34) = 0,0901$, is de gevraagde overschrijdingskans dus ongeveer 0,09. Men kan van deze methode ook gebruik maken, als men de kans wil bepalen, dat s^2 tussen twee waarden in ligt.

13.3.2. GROTE STEEKPROEVEN

In 9.1 (blz. 179) is reeds medegedeeld, dat voor $\nu > 30$ de grootheid $\underline{T} = \sqrt{2\chi^2} - \sqrt{2\nu - 1}$ bij benadering standaardnormaal verdeeld is (de benadering blijkt reeds bevredigend te zijn als $\nu = 30$ en zij wordt beter naarmate ν toeneemt). In combinatie met (13.6) volgt hieruit, dat voor $n > 100$ de grootheid

$$\underline{T} = \sqrt{\frac{2(n-1)s_x^2}{\sigma_x^2} - 2(n-1) - 1} = \frac{s_x}{\sigma_x} \sqrt{2(n-1)} - \sqrt{2(n-1) - 1}$$

bij benadering standaardnormaal verdeeld is. Maar voor grote steekproeven geldt ook, dat $\sqrt{2(n-1)} \cong \sqrt{2(n-1) - 1}$, zodat men dan zonder bezwaar kan schrijven

$$\underline{T} = \frac{s_x}{\sigma_x} \sqrt{2(n-1)} - \sqrt{2(n-1)} \\ = \left(\frac{s_x}{\sigma_x} - 1 \right) \sqrt{2(n-1)} = \frac{s_x - \sigma_x}{\sigma_x} \sqrt{2(n-1)}, \quad \text{of}$$

$$(13.8) \quad T = \frac{\frac{\underline{s}_x - \sigma_x}{\sigma_x}}{\frac{\sigma_x}{\sqrt{2(n-1)}}} = \frac{\underline{s}_x - \sigma_x}{\sigma_x}$$

waarin

$$(13.9) \quad \sigma_x = \frac{\sigma_x}{\sqrt{2(n-1)}}.$$

Voor *grote steekproeven* is de verdeling van de grootheid \underline{s} dus bij benadering normaal met gemiddelde σ_x en spreiding σ_x . Men noemt σ_x wel de standaardfout van de spreiding (standaarddeviatie).

Voorbeeld

13.3. Uit een normaal verdeelde populatie met $\sigma^2 = 12$ trekt men aselekt een steekproef van 51 elementen. Hoe groot is de kans, dat de steekproefvariantie kleiner is dan 7,13?

Om te demonstrenen, dat de normale benadering voor $30 < n < 100$ reeds bevredigend nauwkeurig is, hebben wij met opzet $n = 51$ gekozen. Met (13.8) vindt men dan:

$$T = \frac{(\sqrt{7,13} - \sqrt{12}) \sqrt{2(51-1)}}{\sqrt{12}} = \frac{(2,6702 - 3,4641) 10}{3,4641} = -2,29,$$

met een linkse kans van 0,0110 (tabel A). De exacte toets via (13.6) levert op

$$\chi^2(\nu = 50) = \frac{(51-1)(7,13)}{12} = 29,71,$$

een uitkomst die volgens tabel T een linkse kans van 0,01 bezit. De benaderende kans is dus iets te hoog.

13.3.3. DE SIMULTANE VERDELING VAN $\underline{\bar{x}}$ EN \underline{s}_x

Wij vermelden, dat bij aselekte trekking uit een normaal verdeelde populatie de steekproeffuncties $\underline{\bar{x}}$ en \underline{s}_x (\underline{s}_x^2) stochastisch onafhankelijk zijn.

13.3.4. OPGAVEN

13.4. Men trekt uit een normaal verdeelde populatie met $\sigma^2 = 10$ aselekt een steekproef van 20 elementen. Hoe groot is de kans, dat de steekproefvariantie: a. Groter is dan 13, b. Kleiner is dan 6, c. Tussen 8 en 12 ligt?

13.5. Men trekt uit een normaal verdeelde populatie met $\sigma^2 = 100$ aselekt een steekproef van 99 elementen. Bepaal bij benadering: a. $P(\underline{s} > 12)$, b. $P(\underline{s} < 7)$, c. $P(9 < \underline{s} < 11)$.

— 13.4. Student-verdelingen

13.4.1. DEFINITIE EN TOELICHTING

Wanneer de grootheden u en v stochastisch onafhankelijk zijn, terwijl de grootheid u standaardnormaal verdeeld is en de grootheid v een χ^2 -verdeling volgt met ν vrijheidsgraden, dan volgt de grootheid

$$(13.10) \quad t = \frac{u}{\sqrt{\frac{v}{\nu}}}$$

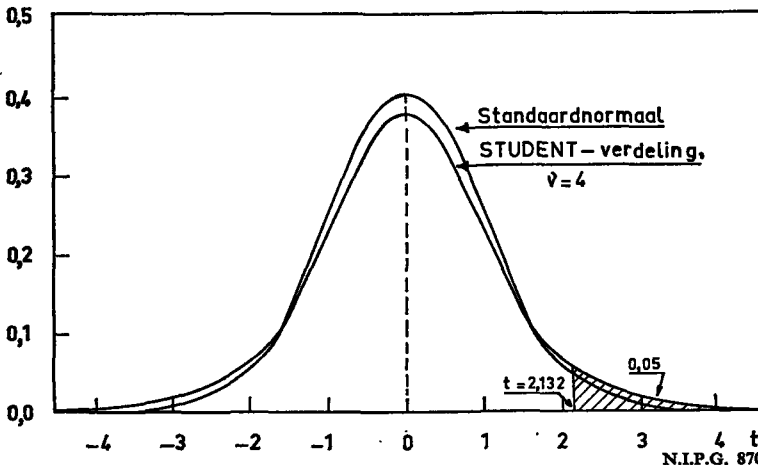
een zg. STUDENT-verdeling met ν vrijheidsgraden.¹

Een STUDENT-verdeling wordt — evenals een χ^2 -verdeling — gespecificeerd door één parameter, het aantal vrijheidsgraden ν . Het is een symmetrische verdeling met $\mu_t = 0$ en $\sigma_t^2 = \nu/(\nu - 2)$, die qua vorm veel op een normale verdeling lijkt, maar een lagere top en dikkere 'staarten' vertoont, zoals blijkt uit figuur 13.1, waarin de STUDENT-verdeling met $\nu = 4$ met de standaardnormale verdeling wordt vergeleken. Bij een toenemend aantal vrijheidsgraden nadert σ_t tot 1 en nadert de STUDENT-verdeling tot de standaardnormale verdeling.

Bij aselechte trekking van een steekproef van n elementen uit een normaal verdeelde populatie is volgens (13.4) de grootheid

$$\underline{T} = (\bar{x} - \mu_x) \sqrt{n}/\sigma_x$$

standaardnormaal verdeeld, en volgens (13.6) volgt de grootheid $\underline{x}^2 = (n - 1) \underline{s}_x^2/\sigma_x^2$ een χ^2 -verdeling met $\nu = n - 1$. Daar deze groot-



Figuur 13.1. De standaardnormale verdeling en de STUDENT-verdeling met 4 vrijheidsgraden.

¹ Deze verdeling is in 1908 gevonden door de Engelse statisticus W. S. GOSSET, die verbonden was aan de Guinness' Brouwerij te Dublin. Daar hij als employé van deze zaak niet onder zijn eigen naam mocht publiceren, koos hij het pseudoniem STUDENT.

heden stochastisch onafhankelijk zijn, volgt dus volgens (13.10) de grootheid

$$(13.11) \quad t = \frac{\bar{T}}{\sqrt{\frac{\chi^2}{n-1}}} = \frac{(\bar{x} - \mu_a) \sqrt{n} / \sigma_a}{\sqrt{(n-1) \frac{s_a^2}{\sigma_a^2}}} = \frac{(\bar{x} - \mu_a) \sqrt{n}}{s_a}$$

een STUDENT-verdeling met $\nu = n - 1$ vrijheidsgraden.

13.4.2. FRACTIELEN VAN STUDENT-VERDELINGEN: TABEL C

In tabel C zijn van verschillende STUDENT-verdelingen een aantal fractielen gegeven. Voor de STUDENT-verdeling met $\nu = 4$ leest men bv. in de kolom $t_{0,95}$ de waarde 2,132 af. Men kan dus kortweg schrijven: $t_{0,95}(\nu = 4) = 2,132$, of nog korter $t_{0,95}(4) = 2,132$. Het fractiel 0,95 van de STUDENT-verdeling met 4 vrijheidsgraden is dus 2,132, d.w.z. links van de ordinaat op dit punt ligt 95% van het oppervlak onder de curve van deze verdeling, zodat $P_L(2,132) = 0,95$ en $P_R(2,132) = 0,05$ (zie figuur 13.1). Wegens de symmetrie van de verdeling t.o.v. $\mu_t = 0$ volgt hieruit, dat $P_L(-2,132) = P_R(2,132) = 0,05$ en dat $P_D(2,132) = 2(0,05) = 0,10$.

Wij hebben reeds opgemerkt, dat de STUDENT-verdeling bij een toenemend aantal vrijheidsgraden tot de standaardnormale verdeling nadert. Uit tabel C blijkt, dat reeds bij $\nu = 120$ de STUDENT-verdeling bevredigend kan worden benaderd door de standaardnormale verdeling, waarvan de fractielen achter $\nu = \infty$ zijn opgenomen.

Voorbeeld

13.4. Verifieer de volgende kansen betreffende STUDENT-verdelingen met tabel C:

$$\begin{aligned} P_L(7,173 | \nu = 4) &= 0,999 & P(-3 < t < +3 | \nu = 7) &= 0,98 \\ P_L(-7,173 | \nu = 4) &= 0,001 & P(-2,179 < t < 2,179 | \nu = 12) &= 0,95 \\ P_D(7,173 | \nu = 4) &= 0,002 & 0,98 < P(-2,6 < t < 2,6 | \nu = 20) &< 0,99 \\ P_R(1,397 | \nu = 8) &= 0,10 & P_D(1,980 | \nu = 120) &= 0,05. \end{aligned}$$

13.4.3. OPGAVEN

13.6. Bepaal met tabel C de volgende kansen betreffende STUDENT-verdelingen:

a. $P_L(3,747 \nu = 4)$	f. $P_R(2,81 \nu = 15)$
b. $P_L(-0,684 \nu = 25)$	g. $P_D(1,43 \nu = 35)$
c. $P_D(2,306 \nu = 8)$	h. $P(-2,06 < t < 2,06 \nu = 25)$
d. $P_D(2,262 \nu = 9)$	i. $P(-1,8 < t < 1,8 \nu = 6)$
e. $P(-2,228 < t < 2,228 \nu = 10)$	j. $P(-1,8 < t < 1,8 \nu = 18)$.

13.7. Bepaal de waarden van t waarvoor geldt $P_D(t | \nu) = 0,05$ voor:

a. $n = 3$ b. $n = 6$ c. $n = 12$ d. $n = 24$ e. $n = 120$ f. $n = 500$.

— 13.5. *F*-verdelingen

13.5.1. DEFINITIE EN TOELICHTING

Als de grootheden v_1 en v_2 elk een χ^2 -verdeling volgen, resp. met v_1 en v_2 vrijheidsgraden en als zij stochastisch onafhankelijk zijn, volgt de grootheid

$$(13.12) \quad \underline{F} = \frac{\underline{\chi}_1^2 / v_1}{\underline{\chi}_2^2 / v_2}$$

een zg. *F*-verdeling met (v_1, v_2) vrijheidsgraden.

In 13.3 hebben wij gezien, dat voor aselechte steekproeven uit een normaal verdeelde populatie met spreiding σ de grootheid $(n - 1)s^2/\sigma^2$ een χ^2 -verdeling met $n - 1$ vrijheidsgraden volgt. Trekt men dus aselechte een steekproef van n_1 elementen uit deze populatie, dan bezit de grootheid

$$\underline{\chi}_1^2 = (n_1 - 1) s_1^2 / \sigma^2$$

een χ^2 -verdeling met $n_1 - 1$ vrijheidsgraden. Voor een aselechte steekproef van n_2 elementen uit dezelfde populatie volgt de grootheid

$$\underline{\chi}_2^2 = (n_2 - 1) s_2^2 / \sigma^2$$

een χ^2 -verdeling met $n_2 - 1$ vrijheidsgraden. Daar aselechte getrokken wordt zijn de grootheden $\underline{\chi}_1^2$ en $\underline{\chi}_2^2$ stochastisch onafhankelijk, zodat volgens (13.12) de grootheid:

$$(13.13) \quad \underline{F} = \frac{\underline{\chi}_1^2 / (n_1 - 1)}{\underline{\chi}_2^2 / (n_2 - 1)} = \frac{s_1^2}{s_2^2}$$

een *F*-verdeling met $(v_1 = n_1 - 1, v_2 = n_2 - 1)$ vrijheidsgraden volgt.¹

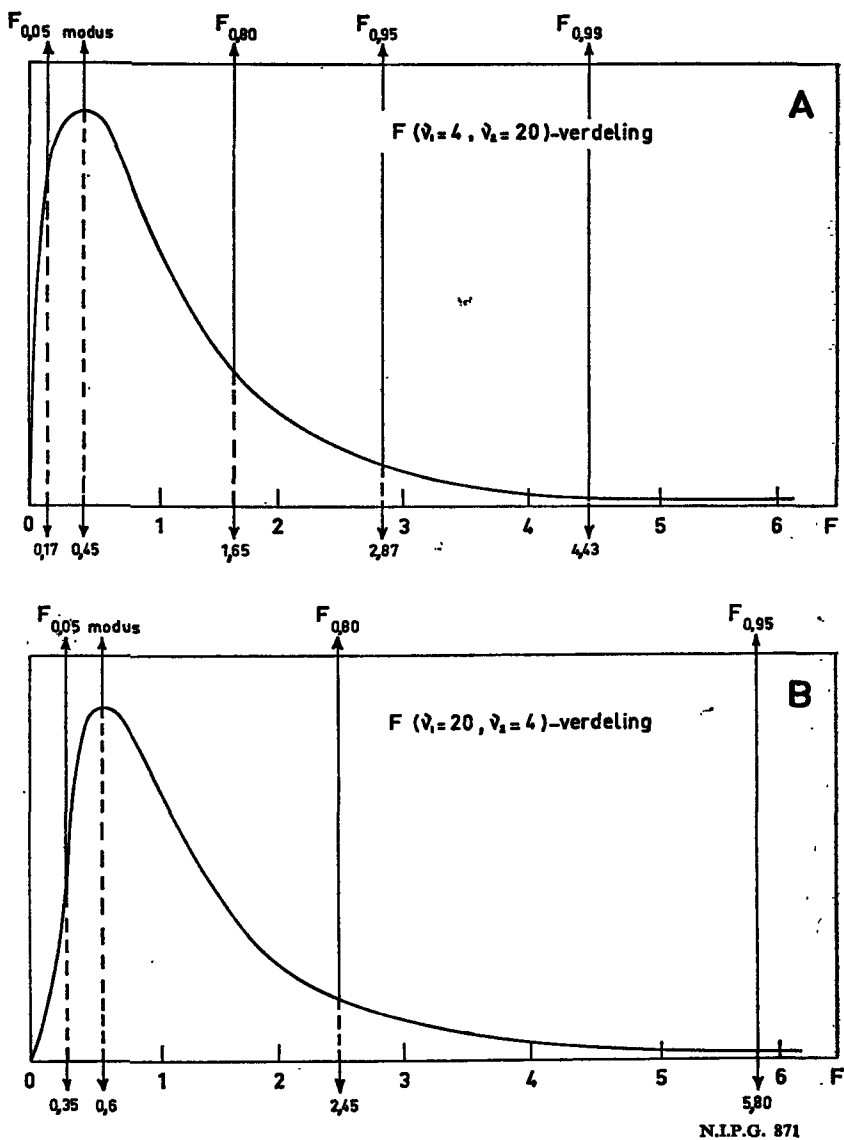
Een *F*-verdeling bezit twee parameters, v_1 en v_2 . Daar $\underline{\chi}_1^2$ en $\underline{\chi}_2^2$ (s_1^2 en s_2^2) de waarde nul of een positieve waarde aannemen, kan hun quotiënt waarden aannemen van nul (als $s_1^2 = 0$ en $s_2^2 > 0$) tot oneindig (als $s_1^2 > 0$ en $s_2^2 = 0$). Een *F*-verdeling heeft voor elk paar verschillende parameters (v_1, v_2) een andere vorm en is steeds scheef. De modus van zo'n verdeling ligt bij $v_2 (v_1 - 2) / v_1 (v_2 + 2)$ en

$$\mu_F = \frac{v_2}{v_2 - 2}.$$

13.5.2. FRACTIELEN VAN *F*-VERDELINGEN: TABEL F

Een tabel van een *F*-verdeling verschilt van die van een χ^2 -verdeling en *t*-verdeling, omdat twee parameters moeten worden opgenomen.

¹ R. A. FISHER bepaalde de exacte verdeling van de grootheid $\underline{F} = \frac{1}{2} \log_e \underline{F} = \log_e s_1 / s_2$. SNEDECOR (49) tabelleerde de verdeling van de 'variance ratio' s_1^2 / s_2^2 en duidde deze grootheid aan met de letter *F* van FISHER.



Figuur 13.2. F -verdelingen: A. $\nu_1 = 4, \nu_2 = 20$, B. $\nu_1 = 20, \nu_2 = 4$.

Men geeft daarom gewoonlijk een aantal fractielen $F_{1-\alpha}$ en per fractieel verschillende waarden van ν_1 en ν_2 . Onze tabel F bevat de vier fractielen, die men het meest nodig heeft, nl. $F_{0,95}$, $F_{0,975}$, $F_{0,99}$ en $F_{0,995}$. In FISHER & YATES (66) vindt men ook $F_{0,80}$ en $F_{0,90}$, in BENNETT & FRANKLIN (4) $F_{0,75}$ en $F_{0,90}$ en de BIOMETRIKA TABLES (74) bevatten nog enkele andere fractielen.

Beschouw nu bv. de F ($\nu_1 = 4, \nu_2 = 20$)-verdeling in figuur 13.2A (wij duiden deze in het vervolg kortweg aan met $F(4, 20)$). In tabel F-1 leest men voor deze verdeling in de kolom $\nu_1 = 4$ en op de regel $\nu_2 = 20$ af: $F_{0,95}(4, 20) = 2,87$, d.w.z. links van de ordinaat op dit punt ligt 95% van het oppervlak onder de $F(4, 20)$ -curve, of: $P_L(2,87) = 0,95$, $P_R(2,87) = 0,05$, $P_D(2,87) = 0,10$. En uit tabel F-3 leest men op deze wijze af: $F_{0,99}(4, 20) = 4,43$, dus $P_L(4,43) = 0,99$ en $P_R(4,43) = 0,01$. Verder geldt voor deze curve: modus $= 20(4 - 2)/4(20 + 2) = 10/22 = 0,45$ en $\mu_F = 20/(20 - 2) = 1,11$.

Er zijn geen tabellen gegeven voor $F_{0,05}$, $F_{0,025}$, $F_{0,01}$ en $F_{0,005}$, omdat men deze fractielen direct kan berekenen door te nemen

$$(13.14) \quad F_\alpha(\nu_1, \nu_2) = \frac{1}{F_{1-\alpha}(\nu_2, \nu_1)}.$$

Onderstel dat men $F_{0,05}(4, 20)$ wil bepalen (dus: $\nu_1 = 4, \nu_2 = 20$). In tabel F-1 zoekt men dan eerst op $F_{1-0,05}(20, 4) = F_{0,95}(20, 4) = 5,80$ en uit (13.14) volgt dan:

$$F_{0,05}(4, 20) = \frac{1}{F_{0,95}(20, 4)} = \frac{1}{5,80} = 0,172.$$

Op overeenkomstige wijze vindt men:

$$F_{0,01}(4, 20) = \frac{1}{F_{0,99}(20, 4)} = \frac{1}{14,0} = 0,071.$$

Tenslotte vermelden wij, dat in de 'STATISTISCHE TABELLEN EN NOMOGRAMMEN' (76a) een nomogram is opgenomen, waarmee men de rechtse overschrijdingskans van een waarde F_0 van $\underline{F}(\nu_1, \nu_2)$ kan aflezen.

13.5.3. OPGAVEN

13.8. Bepaal voor de $F(20, 4)$ -verdeling (figuur 13.2B): $F_{0,95}$, $F_{0,99}$, $F_{0,05}$, $F_{0,01}$, de modus en het gemiddelde.

13.9. Uit een normaal verdeelde populatie trekt men aselekt een steekproef van 10 elementen met variantie s_1^2 en een steekproef van 6 elementen met variantie s_2^2 . Verifieer met tabel F, dat

$$P(\underline{s}_1^2/\underline{s}_2^2 > 10,2) = 0,01 \text{ en } P(\underline{s}_2^2/\underline{s}_1^2 > 6,06) = 0,01.$$

Daar $1/10,2 = 0,098$ en $1/6,06 = 0,165$, volgt hieruit:

$$P(0,165 < \underline{s}_1^2/\underline{s}_2^2 < 10,2) = 0,98 \text{ en } P(0,098 < \underline{s}_2^2/\underline{s}_1^2 < 6,06) = 0,98.$$

Bepaal op overeenkomstige wijze de ontbrekende getallen in de volgende vergelijkingen:

$$a. \quad P(\dots < \underline{s}_1^2/\underline{s}_2^2 < \dots) = 0,95 \qquad b. \quad P(\dots < \underline{s}_2^2/\underline{s}_1^2 < \dots) = 0,95.$$

13.10. Men trekt uit een normaal verdeelde populatie met parameters μ_1 en σ_1 aselekt een steekproef van n_1 elementen. Uit een tweede normaal verdeelde populatie met parameters μ_2 en σ_2 trekt men aselekt een steekproef van n_2

elementen. Geef voor onderstaande waarden van n_1 en n_2 de gevraagde fractiën van de verdeling van $\underline{F} = \underline{s}_1^2 / \underline{s}_2^2$:

	n_1	n_2	$F_{0,01}$	$F_{0,05}$	$F_{0,95}$	$F_{0,99}$
a.	4	8
b.	8	4
c.	8	8
d.	8	16
e.	8	61
f.	61	8
g.	61	25
h.	25	1000
i.	1000	61

niet 13.6.* De verdelingen van \underline{F} , \underline{t} , \underline{T} en $\underline{\chi}^2$

Bestudeer de volgende relaties tussen de verdelingen van

$$\underline{F} (\nu_1 = 1, \nu_2 = n_2 - 1), \underline{t} (\nu = n_2 - 1) \text{ en } \underline{t}^2:$$

$\nu_2 = n_2 - 1$	$F_{0,95} (\nu_1 = 1)$	$t_{0,975}$	$t^2_{0,975}$	$F_{0,99} (\nu_1 = 1)$	$t_{0,995}$	$t^2_{0,995}$
5	6,61	2,571	6,6100	16,3	4,032	16,2570
10	4,96	2,228	4,9640	10,0	3,169	10,0426
15	4,54	2,131	4,5412	8,68	2,947	8,6848
30	4,17	2,042	4,1698	7,56	2,750	7,5625
∞	3,84	1,960	3,8416	6,63	2,576	6,6358

Uit deze tabel zien wij, dat steeds $F_{0,95} = t^2_{0,975}$ en $F_{0,99} = t^2_{0,995}$. In het algemeen blijkt te gelden

$$(13.15) \quad F_{1-\alpha} (\nu_1 = 1, \nu_2 = n_2 - 1) = t^2_{1-\frac{1}{2}\alpha} (\nu = n_2 - 1).$$

De verdeling van \underline{t}^2 met $n - 1$ vrijheidsgraden kan dus beschouwd worden als een F -verdeling met $\nu_1 = 1$ en $\nu_2 = n - 1$.

Bestudeer vervolgens de onderstaande relaties betreffende de verdelingen van $\underline{F} (\nu_1 = 1, \nu_2 = \infty)$ en \underline{T} , resp. \underline{T}^2 :

$F_{0,80} = 1,64$	$F_{0,90} = 2,71$	$F_{0,95} = 3,84$	$F_{0,99} = 6,63$	$F_{0,999} = 10,83$
$T_{0,80} = 1,282$	$T_{0,95} = 1,645$	$T_{0,975} = 1,96$	$T_{0,995} = 2,576$	$T_{0,9995} = 3,291$
$T^2_{0,80} = 1,644$	$T^2_{0,95} = 2,706$	$T^2_{0,975} = 3,842$	$T^2_{0,995} = 6,636$	$T^2_{0,9995} = 10,831$

In het algemeen blijkt te gelden

$$(13.16) \quad F_{1-\alpha} (\nu_1 = 1, \nu_2 = \infty) = T^2_{1-\frac{1}{2}\alpha}.$$

Uit (13.15) en (13.16) volgt (en in de eerste tabel, betreffende F en \underline{t}^2 , is reeds te zien) dat

$$(13.17) \quad t^2_{1-\frac{1}{2}\alpha} (\nu = \infty) = T^2_{1-\frac{1}{2}\alpha}.$$

of, zoals reeds in 13.4 is opgemerkt: de \underline{t} -verdeling nadert tot de standaard-normale verdeling, wanneer ν nadert tot oneindig.

Bestudeer tenslotte de volgende relaties tussen de verdelingen van $\underline{F} (\nu_1 = n_1 - 1, \nu_2 = \infty)$ en $\underline{\chi}^2 (\nu_1 = n_1 - 1)$, resp. $\underline{\chi}^2 / \nu_1$:

$\nu_1 = n_1 - 1$	$F_{0,95}$ ($\nu_2 = \infty$)	$\chi^2_{0,95}$	$\chi^2_{0,95}/\nu_1$	$F_{0,99}$ ($\nu_2 = \infty$)	$\chi^2_{0,99}$	$\chi^2_{0,99}/\nu_1$
3	2,60	7,815	2,605	3,78	11,345	3,782
7	2,01	14,067	2,010	2,64	18,475	2,639
10	1,83	18,307	1,831	2,32	23,209	2,321
30	1,46	43,773	1,459	1,70	50,892	1,696

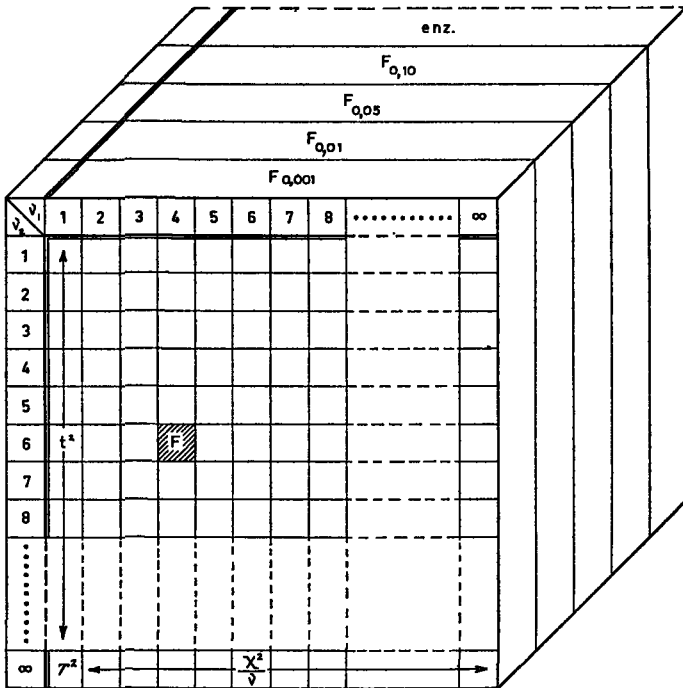
In het algemeen blijkt te gelden, dat

(13.18) $F_{1-\alpha}(\nu_1 = n_1 - 1, \nu_2 = \infty) = \chi^2_{1-\alpha}/\nu_1.$

[Deze relatie is duidelijk: $\underline{F}(\nu_1, \nu_2) = \underline{s}_1^2/\underline{s}_2^2$ en als ν_2 nadert tot oneindig, nadert \underline{s}_2^2 tot σ_2^2 , zodat \underline{F} nadert tot $\underline{s}_1^2/\sigma_2^2 = \chi^2/\nu_1$ — zie formule (13.6).] Uit (13.16), (13.17) en (13.18) volgt dan, dat

(13.19) $F_{1-\alpha}(\nu_1 = 1, \nu_2 = \infty) = T^2_{1-\alpha} = \chi^2_{1-\alpha}(\nu = 1) = t^2_{1-\alpha}(\nu = \infty).$

De relaties, die wij in het voorgaande hebben gedemonstreerd, zijn in figuur 13.3 weergegeven. Elk vierkantje aan de voorzijde van de kubus kan beschouwd worden als de voorkant van een lade, waarin men fractielen van een F -verdeling aantreft. De gearceerde 'lade' bevat dus de $F(4, 6)$ -verdeling, de lade links onder de $F(1, \infty)$ -verdeling = de verdeling van T^2 , enz. Trekt men alle laden even ver uit, dan kan men dus voor alle F -verdelingen een bepaald fractiel $F_{1-\alpha}$ aflezen.



N.I.P.G. 980

Figuur 13.3. De relatie tussen de grootheden \underline{F} , \underline{t} , \underline{T} en $\underline{\chi^2}/\nu$. Naar MATHER (47).

13.7. De kansverdeling van het verschil tussen twee gemiddelden

- 13.7.1. POPULATIEVARIANTIES BEKEND

Beschouw twee normaal verdeelde grootheden, t.w. de verdeling van de grootheid x_1 met gemiddelde μ_1 en spreiding σ_1 op populatie I en de verdeling van de grootheid x_2 met gemiddelde μ_2 en spreiding σ_2 op populatie II. De grootheid \bar{x}_1 , het gemiddelde van een aselechte steekproef van n_1 elementen uit populatie I, is dan normaal verdeeld met gemiddelde $\mu_{\bar{x}_1} = \mu_1$ en variantie $\sigma^2_{\bar{x}_1} = \sigma_1^2/n_1$. De grootheid \bar{x}_2 , het gemiddelde van een aselechte steekproef van n_2 elementen uit populatie II, is eveneens normaal verdeeld, met gemiddelde $\mu_{\bar{x}_2} = \mu_2$ en variantie $\sigma^2_{\bar{x}_2} = \sigma_2^2/n_2$. Het verschil tussen deze gemiddelden, $\bar{x}_1 - \bar{x}_2$, blijkt eveneens normaal verdeeld te zijn, met volgens (3.5.2) als gemiddelde

$$(13.20) \quad \mu_{\bar{x}_1 - \bar{x}_2} = \mu_1 - \mu_2$$

en volgens (3.6.2) als variantie

$$(13.21) \quad \sigma^2_{\bar{x}_1 - \bar{x}_2} = \sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Hieruit volgt dat de grootheid

$$(13.22) \quad \underline{T} = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

standaardnormaal verdeeld is.

Wanneer de populaties I en II hetzelfde gemiddelde bezitten, is $\mu_1 - \mu_2 = 0$, zodat (13.22) overgaat in

$$(13.23) \quad \underline{T} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

Als de populatieverdelingen tevens dezelfde spreiding bezitten, zodat $\sigma_1 = \sigma_2 = \sigma$, gaat (13.21) over in

$$(13.24) \quad \sigma^2_{\bar{x}_1 - \bar{x}_2} = \sigma^2 \left[\frac{1}{n_1} + \frac{1}{n_2} \right] = \sigma^2 \left(\frac{n_1 + n_2}{n_1 n_2} \right),$$

zodat men voor (13.23) kan schrijven

$$(13.25) \quad \underline{T} = \frac{\bar{x}_1 - \bar{x}_2}{\sigma} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}.$$

Voorbeelden

13.5. Van twee normaal verdeelde populaties is gegeven:

Populatie I: $\mu_1 = 25$, $\sigma_1 = 3$; Populatie II: $\mu_2 = 20$, $\sigma_2 = 4$.

Men trekt uit populatie I een aselechte steekproef van 10 elementen (met gemiddelde \bar{x}_1) en uit populatie II een aselechte steekproef van 20 elementen (met gemiddelde \bar{x}_2). Hoe groot is de kans, dat het verschil tussen de steekproefgemiddelden $\bar{x}_1 - \bar{x}_2$ kleiner dan 2 is? Volgens (13.22) is:

$$T = \frac{2 - \frac{(25 - 20)}{20}}{\sqrt{\frac{(3)^2}{10} + \frac{(4)^2}{20}}} = \frac{-3}{\sqrt{\frac{17}{10}}} = \frac{-3}{1,3038} = -2,30.$$

Uit tabel A blijkt, dat $P_L(T = -2,30) = 0,0107$. Dit is dus de gevraagde kans.

13.6. Van een normaal verdeelde populatie is gegeven dat $\mu = 40$ en $\sigma = 5$. Men trekt aselekt een steekproef van $n_1 = 16$ exemplaren en vervolgens een tweede aselechte steekproef van $n_2 = 25$ exemplaren. Hoe groot is de kans, dat het verschil tussen de gemiddelden van deze steekproeven meer dan 3 van nul verschilt?

Als getrokken wordt uit dezelfde populatie geldt (13.25), zodat:
 $T = \frac{3}{5} \sqrt{\frac{16 \times 25}{16 + 25}} = 1,87$. Daar gevraagd wordt naar de kans op een verschil van 3 in positieve of negatieve richting dient van deze uitkomst de tweezijdige overschrijdingskans te worden bepaald. Tabel A levert op: $P_D(T = 1,87) = 2(0,0307) = 0,0614$.

13.7.2. POPULATIEVARIANTIES ONBEKEND

Als men beschikt over twee steekproeven uit een normaal verdeelde populatie, resp. van de omvang n_1 en n_2 , dan kan men uit de steekproefvarianties s_1^2 en s_2^2 een zuivere schatting verkrijgen van de populatievariantie σ^2 door te nemen:

$$(13.26) \quad s^2 = \frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2}.$$

Daar $(n_1 - 1) s_1^2 = \Sigma (x_1 - \bar{x}_1)^2 = \Sigma x_1^2 - (\Sigma x_1)^2/n_1$ en $(n_2 - 1) s_2^2 = \Sigma (x_2 - \bar{x}_2)^2 = \Sigma x_2^2 - (\Sigma x_2)^2/n_2$ kan men s^2 berekenen met

$$(13.27) \quad s^2 = \frac{\Sigma x_1^2 - \frac{(\Sigma x_1)^2}{n_1} + \Sigma x_2^2 - \frac{(\Sigma x_2)^2}{n_2}}{n_1 + n_2 - 2} = \frac{\Sigma x^2 - \left[\frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} \right]}{n_1 + n_2 - 2}.$$

Voor $n_1 = n_2 = n$ gaat (13.26) over in

$$(13.28) \quad s^2 = \frac{(n-1)s_1^2 + (n-1)s_2^2}{2(n-1)} = \frac{s_1^2 + s_2^2}{2}.$$

Analoog aan (13.26) geldt voor k steekproeven

$$(13.29) \quad s^2 = \frac{\sum_{j=1}^k (n_j - 1) s_j^2}{\sum_{j=1}^k (n_j - 1)} = \frac{\sum (n_j - 1) s_j^2}{\sum n_j - k}.$$

Het is duidelijk dat (13.29) voor k steekproeven van dezelfde omvang overgaat in:

$$(13.30) \quad s^2 = \frac{\sum_{j=1}^k s_j^2}{k}.$$

Volgens (13.6) volgt de grootheid

$$(13.31) \quad \chi^2 = \frac{(n_1 + n_2 - 2) s^2}{\sigma^2}$$

een χ^2 -verdeling met $n_1 + n_2 - 2$ vrijheidsgraden. Uit (13.25), (13.31) en (13.10) volgt, dat de grootheid

$$(13.32) \quad t = \frac{\bar{x}_1 - \bar{x}_2}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

een STUDENT-verdeling met $n_1 + n_2 - 2$ vrijheidsgraden bezit [vergelijk met formule (13.11)]. De lezer verifiëre dat men, als $n_1 = n_2 = n$, voor (13.32) kan schrijven:

$$(13.33) \quad t = \frac{(\bar{x}_1 - \bar{x}_2) \sqrt{n}}{\sqrt{s_1^2 + s_2^2}}.$$

Het voorgaande geldt uiteraard eveneens als men beschikt over 2 (k) aselechte steekproeven uit verschillende normaal verdeelde populaties, die *dezelfde variantie* hebben.

13.7.3. OPGAVEN

13.11. Van twee normaal verdeelde populaties is gegeven: $\mu_1 = 150,2$, $\sigma_1^2 = 18$, $\mu_2 = 142,0$, $\sigma_2^2 = 4$. Men trekt aselekt een steekproef van 10 elementen uit populatie I en een steekproef van 25 elementen uit populatie II. Hoe groot is de kans, dat het gemiddelde van de eerste steekproef meer dan 11 groter is dan dat van de tweede?

13.12. Van twee normaal verdeelde populaties is bekend dat $\mu_1 = 35$ en $\sigma_1 = 4$, resp. dat $\mu_2 = 30$ en $\sigma_2 = 2$. Men trekt aselekt uit elke populatie een steekproef van n elementen. Bepaal de waarde van n waarvoor geldt, dat de kans dat het gemiddelde van de steekproef uit de eerste populatie zeven of meer groter is dan het gemiddelde van de steekproef uit de tweede populatie, ongeveer gelijk is aan 0,05.

13.13. Men beschikt over vijf aselechte steekproeven uit een normaal verdeelde populatie, waarvan de volgende gegevens ter beschikking staan: $n_1 = 10$, $s_1 = 2,11$; $n_2 = 21$, $s_2 = 2,85$; $n_3 = 8$, $s_3 = 2,44$; $n_4 = 25$, $s_4 = 3,02$; $n_5 = 33$, $s_5 = 2,75$. Schat hieruit de populatiespreiding.

— 13.8. Een steekproefexperiment

13.8.1. INLEIDING EN BESCHRIJVING

In de voorgaande paragrafen hebben wij de kansverdelingen van een aantal steekproeffuncties bij aselechte trekking uit normaal verdeelde populaties de revue laten passeren. Deze verdelingen zijn alle langs wiskundige weg afgeleid. Zij kunnen echter ook — zij het bij benadering — langs empirische weg worden gevonden en wel door inderdaad een groot aantal aselechte steekproeven van bepaalde omvang te trekken en daaruit de waargenomen frequentieverdeling van de te beschouwen steekproeffunctie op te stellen. Het is wel duidelijk dat de mathematische oplossing, die algemeen geldende, exacte kansverdelingen oplevert, superieur is aan de empirische methode. Het komt dan ook betrekkelijk weinig voor, dat men de empirische weg (die men tegenwoordig de Monte Carlo methode noemt) in werkelijkheid gebruikt. Soms zijn er echter problemen, waarbij de wiskundige moeilijkheden vrijwel onoplosbaar zijn, terwijl bij het volgen van de Monte Carlo methode de enorme hoeveelheid (reken)werk door het gebruik van elektronische machines niet langer onoverkomelijk is.

Behalve voor het oplossen van zeer moeilijke vraagstukken wordt de Monte Carlo methode ook voor opleidingsdoeleinden gebruikt, bv. om de steekproeftheorie en de verschillende daarbij optredende kansverdelingen ook bij de niet wiskundig geschoolde lezer tot een levend begrip te maken. Hiertoe beschrijven wij nu een experiment, waarbij met behulp van aselechte getallen een groot aantal steekproeven van kleine omvang ($n = 5$) wordt getrokken uit de pseudo-normaal verdeelde populatie van 201 getallen in tabel Q, met $\mu = 20$ en $\sigma = 3,470$. De hierbij waargenomen verdelingen van de steekproeffuncties in 13.2 t/m 13.5 worden dan stuk voor stuk vergeleken met hun kansverdelingen. Dit experiment leent zich uitstekend voor uitvoering in groepsverband: zijn er bv. 20 deelnemers, dan kan men iedere deelnemer 10 steekproeven laten trekken en dus beschikken over 200 steekproeven. Men dient echter de trekking zo te regelen, dat de deelnemers niet (gedeeltelijk) dezelfde uitkomsten krijgen.

De uitkomsten van de trekking en de daarop volgende berekeningen kunnen het best worden vastgelegd op een werkstaat, zoals deze in tabel 13.2 is opgenomen. Hierbij zijn de volgende stappen te onderscheiden:

1. Trek met teruglegging uit een tabel met aselechte getallen vijf

steekproeven van vijf getallen tussen 000 en 202 en noteer deze achter (1) in de kolommen 1 t/m 5 onder 'Nr'. (Als men het experiment met een groep uitvoert, laat men tevoren iedere deelnemer een werkstaat met het juiste aantal kolommen maken.)

2. Lees de bij de verkregen getallen behorende waarde van x in tabel Q af en noteer deze. Bepaal vervolgens x^2 .

3. Voer de berekeningen op de regels 2 t/m 12 uit. Deze leveren per steekproef van 5 elementen: \bar{x} , s^2 , s , T , t en χ^2 .

In de volgende punten van deze paragraaf geven wij nu successievelijk de bij 250 steekproeven van 5 elementen waargenomen verdelingen van verschillende steekproeffuncties. Als men het experiment zelf heeft uitgevoerd, kunnen de uitkomsten daarvan op dezelfde wijze worden beschouwd.

Tabel 13.2. Werkstaat

Omschrijving	Aselecte steekproeven (met teruglegging) van 5 elementen*																																																																										
	1			2			3			4			5																																																														
	Nr	x	x ²	Nr	x	x ²	Nr	x	x ²	Nr	x	x ²	Nr	x	x ²																																																												
1. Aselect gekozen nummers en daarmee corresponderende waarden van x en van x^2	173	17	289	201	19	361	019	17	289	079	18	324	010	17	289	074	19	361	029	22	484	124	21	441	050	24	576	162	22	484	005	20	400	110	15	225	104	22	484	146	12	144	050	24	576	138	21	441	076	23	529	097	21	441	138	21	441	194	17	289	166	18	324	040	23	529	050	24	576	062	17	289	133	20	400
2. Σx	95			102			105			92			100																																																														
3. $\bar{x} = \frac{\Sigma x}{5}$	19,0			20,4			21,0			18,4			20,0																																																														
4. Σx^2	1815			2128			2231			1774			2038																																																														
5. $\frac{(\Sigma x)^2}{5} = (0,2) (\Sigma x)^2$	1805			2080,8			2205			1692,8			2000																																																														
6. $\Sigma x^2 - \frac{(\Sigma x)^2}{5}$	10			47,2			26			81,2			38																																																														
7. $s^2 = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{5}}{4}$	2,5			11,8			6,5			20,3			9,5																																																														
8. $s = \sqrt{s^2}$	1,58			3,44			2,55			4,51			3,08																																																														
9. $\bar{x} - \mu = \bar{x} - 20$	-1,0			+0,4			+1,0			-1,6			0																																																														
10. $T = \frac{(\bar{x} - \mu)\sqrt{n}}{\sigma} = \frac{(\bar{x} - 20)(2,236)}{3,470}$	-0,64			+0,26			+0,64			-1,03			0																																																														
11. $t = \frac{(\bar{x} - \mu)\sqrt{n}}{s} = \frac{(\bar{x} - 20)(2,236)}{s}$	-1,42			+0,26			+0,88			-0,79			0																																																														
12. $\chi^2 = \frac{\Sigma(x - \bar{x})^2}{\sigma^2} = \frac{\Sigma x^2 - \frac{(\Sigma x)^2}{5}}{12,041}$...	0,83			3,92			2,16			6,74			3,16																																																														

* Trekking uit tabel Q, te beginnen bij kolom 3, regel 6, kolomsgewijs omlaaggaande.

Tabel 13.3. Verdeling van de gemiddelden van 250 steekproeven van 5 elementen uit een pseudo-normaal verdeelde populatie

\bar{x}	Ongegroepeerd f	Gegroepeerd f	Berekeningen		
			u	fu	fu^2
25,0					
24,8		1	+8	+ 8	64
24,6	1				
24,4					
24,2	1	1	+7	+ 7	49
24,0					
23,8	1				
23,6	1	3	+6	+18	108
23,4	1				
23,2	1				
23,0	2	5	+5	+25	125
22,8	2				
22,6	3				
22,4	3	12	+4	+48	192
22,2	6				
22,0	6				
21,8	7	19	+3	+57	171
21,6	6				
21,4	5				
21,2	9	23	+2	+46	92
21,0	9				
20,8	7				
20,6	10	35	+1	+35	35
20,4	18				
20,2	19				
20,0	14	47	0	0	0
19,8	14				
19,6	9				
19,4	14	36	-1	-36	36
19,2	13				
19,0	8				
18,8	10	28	-2	-56	112
18,6	10				
18,4	11				
18,2	5	18	-3	-54	162
18,0	2				
17,8	6				
17,6	1	10	-4	-40	160
17,4	3				
17,2	3				
17,0	1	6	-5	-30	150
16,8	2				
16,6	3				
16,4	1	4	-6	-24	144
16,2					
16,0	1				
15,8		2	-7	-14	98
15,6	1				
Totaal	250	250		-10 Σfu	1698 Σfu^2

13.8.2. DE WAARGENOMEN VERDELING VAN HET GEMIDDELDE

De verdeling van de 250 steekproefgemiddelden, die het experiment heeft opgeleverd, is in tabel 13.3 onder 'ongegroepeerd' opgenomen. Deze verdeling is vervolgens gegroepeerd, waarna men duidelijk een pseudo-normale verdeling met als modale klasse 19,8 t/m 20,2 ziet verschijnen. Berekent men via een hulpgrootheid u het gemiddelde \bar{x} en de spreiding $s_{\bar{x}}$ van deze verdeling van steekproefgemiddelden, dan is:

$$\bar{u} = \frac{-10}{250} = -0,04, \quad \bar{x} = 20 + (0,6)(-0,04) = 19,98,$$

$$s_u^2 = \frac{1698 - \frac{(-10)^2}{250}}{249} = 6,8177, \quad s_u = 2,611, \quad s_{\bar{x}} = (0,6)(2,611) = 1,567.$$

Hieruit blijkt, dat $\bar{x} = 19,98$ vrijwel gelijk is aan de verwachting $\mu_{\bar{x}} = \mu_u = 20$ en dat $s_{\bar{x}} = 1,567$ slechts iets hoger is dan

$$\sigma_{\bar{x}} = \sigma_u / \sqrt{n} = 3,47 / \sqrt{5} = 1,552.$$

13.8.3. DE WAARGENOMEN VERDELING VAN DE VARIANTIE

De verdeling van de 250 waargenomen waarden van

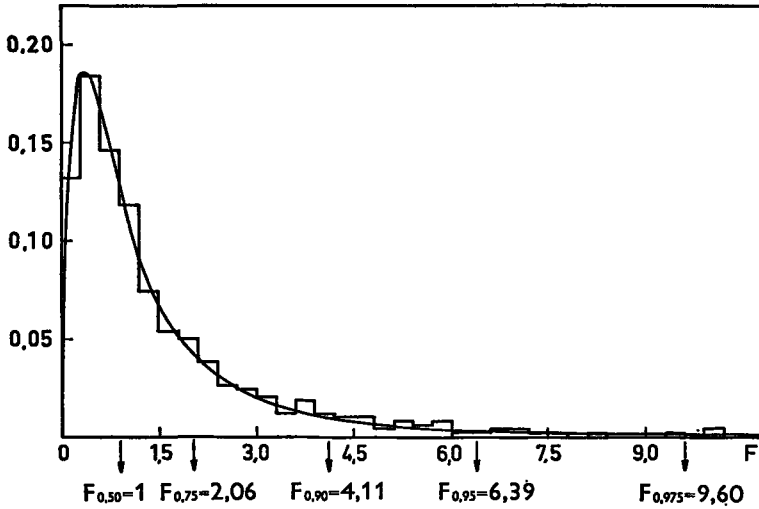
$$\chi^2 = \frac{(n-1) s^2}{\sigma^2} = \frac{4s^2}{(3,47)^2}$$

is in tabel 13.4 opgenomen. In figuur 13.4 is deze empirisch verkregen verdeling vergeleken met de theoretische, d.i. met de χ^2 -verdeling met $n - 1 = 4$ vrijheidsgraden. De overeenstemming tussen deze verdelingen blijkt reeds zeer bevredigend te zijn: de empirische verdeling

Tabel 13.4. Waargenomen verdeling van $\chi^2 = (n - 1) s^2 / \sigma^2$ voor 250 steekproeven van 5 elementen uit een pseudo-normaal verdeelde populatie met $\sigma = 3,47$

Klasse	Frequentie	Rel. freq.	Klasse	Frequentie	Rel. freq.
			Transport	229	0,916
0,00 - 0,99	24	0,096	9,00 - 9,99	6	0,024
1,00 - 1,99	37	0,148	10,00 - 10,99	4	0,016
2,00 - 2,99	45	0,180	11,00 - 11,99	2	0,008
3,00 - 3,99	37	0,148	12,00 - 12,99	1	0,004
4,00 - 4,99	30	0,120	13,00 - 13,99	3	0,012
5,00 - 5,99	19	0,076	14,00 - 14,99	2	0,008
6,00 - 6,99	13	0,052	15,00 - 15,99	2	0,008
7,00 - 7,99	15	0,060	16,00 - 16,99	—	0,000
8,00 - 8,99	9	0,036	17,00 - 17,99	1	0,004
Transporteren	229	0,916	Totaal	250	1,000

heeft iets te weinig uitkomsten tussen 1 en 2 en iets te veel hoge waarden.



Figuur 13.4. Waargenomen verdeling van de grootheid $\chi^2 = (n-1)s^2/\sigma^2$ voor 250 steekproeven van 5 elementen uit een pseudo-normaal verdeelde populatie met $\sigma = 3,47$ en de χ^2 -verdeling met 4 vrijheidsgraden.

Tabel 13.5. Simultane verdeling van het gemiddelde en de spreiding voor 250 steekproeven van 5 elementen uit een pseudo-normaal verdeelde populatie, $r = -0,007$

		Gemiddelden (\bar{x})														Totaal		
		15,6-16,0	16,2-16,6	16,8-17,2	17,4-17,8	18,0-18,4	18,6-19,0	19,2-19,6	19,8-20,2	20,4-20,8	21,0-21,4	21,6-22,0	22,2-22,6	22,8-23,2	23,4-23,8		24,0-24,4	24,6-25,0
Standaarddeviaties (s)	7,00-7,49								1									1
	6,50-6,99						1	2	1	1								5
	6,00-6,49					1		1				2						4
	5,50-5,99				1		1			1	1		1					5
	5,00-5,49				1		4		1	1	3		1					11
	4,50-4,99			2	1	1	1	1	7	3	1	2	1	2				22
	4,00-4,49				1		4	3	5	4	3	1	2		1			24
	3,50-3,99		1		3	4	2	7	5	7	4	1	3	1	1			39
	3,00-3,49	1	2	3	2	3	3	8	13	4		2		1				42
	2,50-2,99	1	1		1	5	4	5	2	1	5	4	2	1		1		33
	2,00-2,49					4	2	3	3	8	1	3	1				1	26
	1,50-1,99			1			3	3	4	3	5	2	1		1			23
	1,00-1,49						2	2	3	2		2						11
	0,50-0,99						1	1	2									4
Totaal		2	4	6	10	18	28	36	47	35	23	19	12	5	3	1	1	250

13.8.4. DE WAARGENOMEN SIMULTANE VERDELING VAN HET GEMIDDELDE EN DE SPREIDING

Elke getrokken steekproef van 5 elementen levert een waarnemingspaar (\bar{x}, s) op. In tabel 13.5 zijn de paren (\bar{x}, s) die bij 250 aselechte steekproeven van 5 elementen werden verkregen in een correlatietabel opgenomen. Deze tabel laat zien, dat bij trekking uit een (pseudo) normaal verdeelde populatie de grootheden \bar{x} en s onafhankelijk zijn. De waargenomen correlatiecoëfficiënt $r = -0,007$ ligt zeer dicht bij de verwachting ($\rho = 0$).¹

Tabel 13.6. Waargenomen verdeling van de grootheden T en t bij 250 steekproeven van 5 elementen uit een pseudo-normaal verdeelde populatie met $\mu = 20$ en $\sigma = 3,47$

Klassemidden (1)	Frequenties	
	$T = \frac{(\bar{x} - 20) \sqrt{5}}{3,47}$ (2)	$t = \frac{(\bar{x} - 20) \sqrt{5}}{s}$ (3)
5,03		1
4,64		2
4,26		—
3,87		1
3,48		1
3,10	1	1
2,71	1	3
2,32	3	3
1,93	5	10
1,55	12	11
1,16	19	13
0,77	23	24
0,39	35	34
0	47	42
−0,39	36	33
−0,77	28	23
−1,16	18	15
−1,55	10	14
−1,93	6	7
−2,32	4	4
−2,71	2	2
−3,10		2
−3,48		1
−3,87		1
−4,26		1
−4,64		1
Totaal	250	250

Zie voor een bespreking van de correlatiecoëfficiënt 18.3 en hoofdstuk 20.

13.8.5. DE WAARGENOMEN VERDELINGEN VAN DE GROOTHEDEN \underline{T} EN \underline{t}

De verdeling van de 250 waargenomen waarden van

$$\underline{T} = (\bar{x} - \mu) \sqrt{n}/\sigma = (\bar{x} - 20) \sqrt{5}/3,47$$

is in tabel 13.6 (kolom 2) opgenomen. Deze verdeling is verkregen uit de gegroepeerde verdeling van \bar{x} in tabel 13.3 en heeft dus als klassebreedte $0,6 \sqrt{5}/3,47 = 0,387$.

De 250 waargenomen waarden van $\underline{t} = (\bar{x} - \mu) \sqrt{n}/s = (\bar{x} - 20) \sqrt{5}/s$ zijn in kolom 3 van de tabel in dezelfde klassen ondergebracht als de T -waarden. Uit deze verdeling kan men berekenen: $\bar{x}_t = 0,01$ en $s_t = 1,38$, waarden die vrijwel gelijk zijn aan hun verwachting (zie

$$13.4): \mu_t = 0, \quad \sigma_t = \sqrt{\frac{n-1}{n-3}} = \sqrt{2} = 1,41.$$

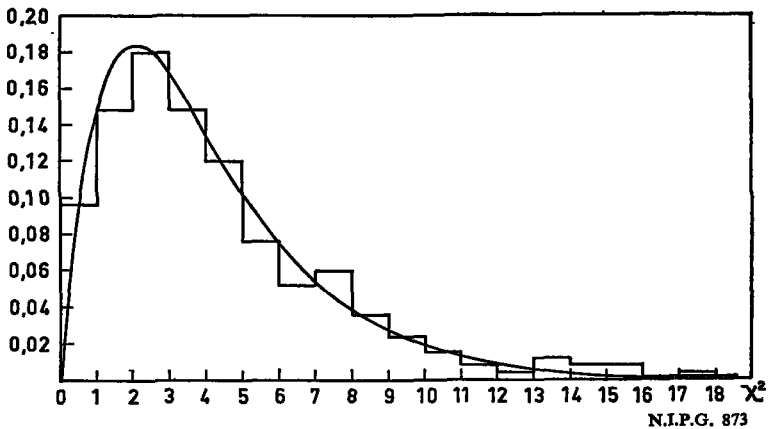
Wij laten het aan de lezer over om de waargenomen verdelingen van \underline{T} en \underline{t} grafisch te vergelijken. Uit tabel 13.6 blijkt echter reeds duidelijk, dat beide verdelingen ongeveer symmetrisch zijn t.o.v. nul, maar dat de t -verdeling in de buurt van dit centrum iets lagere frequenties vertoont dan de T -verdeling, maar daartegenover langere 'staarten' te zien geeft (vergelijk met figuur 13.1).

13.8.6. DE WAARGENOMEN VERDELING VAN DE GROOTHEID \underline{F}

Bij het steekproefexperiment zijn 250 steekproefvarianties (s^2) verkregen, die stuk voor stuk onafhankelijke schattingen zijn van de populatievariantie σ^2 . Hieruit zijn 500 verschillende quotiënten $F = s_i^2/s_j^2$ gevormd (door te nemen: de variantie van steekproef 1,

Tabel 13.7. Waargenomen verdeling van $\underline{F} = s_i^2/s_j^2$ voor 500 paren steekproeven van 5 elementen uit een pseudo-normaal verdeelde populatie

Klasse	f	f/500	Klasse	f	f/500	Klasse	f	f/500
0,00 - 0,29	66	0,132	Transport	439	0,878	Transport	481	0,962
0,30 - 0,59	92	0,184	3,60 - 3,89	9	0,018	6,90 - 7,19	2	0,004
0,60 - 0,89	73	0,146	3,90 - 4,19	6	0,012	7,20 - 7,49	1	0,002
0,90 - 1,19	59	0,118	4,20 - 4,49	5	0,010	7,50 - 7,79	1	0,002
1,20 - 1,49	37	0,074	4,50 - 4,79	5	0,010	7,80 - 8,09	-	0,000
1,50 - 1,79	27	0,054	4,80 - 5,09	2	0,004	8,10 - 8,39	1	0,002
1,80 - 2,09	25	0,050	4,80 - 5,39	4	0,008	8,40 - 8,69	-	0,000
2,10 - 2,39	19	0,038	5,40 - 5,69	3	0,006	8,70 - 8,99	-	0,000
2,40 - 2,69	13	0,026	5,70 - 5,99	4	0,008	9,00 - 9,29	-	0,000
2,70 - 2,99	12	0,024	6,00 - 6,29	1	0,002	9,30 - 9,59	1	0,002
3,00 - 3,29	10	0,020	6,30 - 6,59	1	0,002	9,60 - 9,89	-	0,000
3,30 - 3,59	6	0,012	6,60 - 6,89	2	0,004	9,90 - 10,19	2	0,004
						> 10,20	11	0,022
Transporteren	439	0,878	Transporteren	481	0,962	Totaal	500	1,000



Figuur 13.5. Waargenomen verdeling van de grootheid $F = \frac{s_i^2/s_j^2}$ voor 500 paren steekproeven van 5 elementen uit een pseudo-normaal verdeelde populatie en de F -verdeling met $\nu_1 = 4$ en $\nu_2 = 4$. (Weggelaten zijn 11 F -waarden groter dan 10,20; hiervan zijn er drie groter dan 20, t.w. 33,3, 45,0 en 64,7.)

gedeeld door de variantie van steekproef 2, resp. van steekproef 3; de variantie van steekproef 2, gedeeld door de variantie van steekproef 3, resp. van steekproef 4, enz.). De verdeling van deze 500 F -waarden is in tabel 13.7 gegeven en in figuur 13.5 vergeleken met de theoretische F -curve met $\nu_1 = 4$ en $\nu_2 = 4$. De overeenstemming tussen de waargenomen en de theoretische verdeling blijkt zeer goed te zijn.

Beneden $F_{0,75} = 2,06$ verwachten wij 375 van de 500 F -waarden en beneden 2,09 vinden wij er 379; boven $F_{0,975} = 9,60$ verwachten wij $2\frac{1}{2}\%$ van de 500 F -waarden, dat zijn er dus 12 à 13, terwijl er 13 worden waargenomen. Tenslotte vermelden wij, dat het gemiddelde van de waargenomen verdeling 2,003 bedraagt en dus zeer goed overeenstemt met $\mu_F = \nu_2/(\nu_2 - 2) = 2$.

TECHNIEKEN BETREFFENDE VARIANTIES

14.1. Toets voor de hypothese $\sigma^2 = \sigma_0^2$

overeenhang.

14.1.1. KLEINE STEEKPROEVEN ($n \leq 100$)

Onderstel, dat men de hypothese H_0 wil toetsen, dat de variantie van een normaal verdeelde populatie een bepaalde waarde σ_0^2 heeft en beschikt over een aselechte steekproef van n elementen met variantie s_0^2 . De toetsing kan als volgt worden uitgevoerd:

Tweezijdige toetsing

- a. $H_0: \sigma^2 = \sigma_0^2$, alternatieve hypothese $H_1: \sigma^2 \neq \sigma_0^2$.
- b. Kies de onbetrouwbaarheidsdrempel α waarmee de toetsing wordt uitgevoerd.
- c. De toetsingsgrootheid is

$$(14.1) \quad \chi^2 = \frac{(n-1) s^2}{\sigma_0^2}$$

- d. Onder H_0 volgt deze grootheid volgens (13.6) een χ^2 -verdeling met $n - 1$ vrijheidsgraden.
- e. Bereken χ^2_0 door in (14.1) de waarden van n , s_0^2 en σ_0^2 te substitueren.
- f. Verwerp H_0 , indien $P_D(\chi^2_0) = 2P_R(\chi^2_0) \leq \alpha$ is, d.i. als:

$$\chi^2_0 \leq \chi^2_{\frac{1}{2}\alpha} (\nu = n - 1) \text{ of als } \chi^2_0 \geq \chi^2_{1-\frac{1}{2}\alpha} (\nu = n - 1)$$

De kritieke waarden $\chi^2_{\frac{1}{2}\alpha}$ en $\chi^2_{1-\frac{1}{2}\alpha}$ kunnen in tabel T worden opgezocht. Voor $\alpha = 0,05$ zijn zij resp. $\chi^2_{0,025}$ en $\chi^2_{0,975}$, voor $\alpha = 0,01$ zijn zij resp. $\chi^2_{0,005}$ en $\chi^2_{0,995}$ en voor $\alpha = 0,001$ zijn zij resp. $\chi^2_{0,0005}$ en $\chi^2_{0,9995}$.

Links éézijdige toetsing

- a. $H_0: \sigma^2 = \sigma_0^2$ (men kan ook zeggen: $\sigma^2 \geq \sigma_0^2$), $H_1: \sigma^2 < \sigma_0^2$.
Als s_0^2 groter is dan σ_0^2 kan men direct tot het niet verwerpen van H_0 besluiten. Anders verlopen de stappen b t/m e als bij de tweezijdige toetsing en dan volgt:
- f. Verwerp H_0 , indien $P_L(\chi^2_0) \leq \alpha$ is, d.i. voor $\chi^2_0 \leq \chi^2_\alpha (\nu = n - 1)$. De kritieke waarden χ^2_α zijn: voor $\alpha = 0,05$, $\chi^2_{0,005}$; voor $\alpha = 0,01$, $\chi^2_{0,01}$ en voor $\alpha = 0,001$, $\chi^2_{0,001}$ (tabel T).

Rechts éénzijdige toetsing

a. $H_0: \sigma^2 = \sigma_0^2$ (men kan ook zeggen: $\sigma^2 \leq \sigma_0^2$), $H_1: \sigma^2 > \sigma_0^2$.

Als s_0^2 kleiner is dan σ_0^2 kan men onmiddellijk tot het niet verwerpen van H_0 besluiten. Anders volgt op de stappen b t/m e:

f. Verwerp H_0 , indien $P_R(\chi^2_0) < \alpha$ is, d.i. voor $\chi^2_0 \geq \chi^2_{1-\alpha}$ ($\nu = n-1$).

Voorbeelden

14.1. Een aselechte steekproef van 15 elementen uit een pseudo-normaal verdeelde populatie levert op $s_0 = 4$. Toets met $\alpha = 0,05$ de hypothese dat de populatiespreiding gelijk is aan 6.

Er dient tweezijdig te worden getoetst en deze toetsing verloopt op de volgende wijze:

a. $H_0: \sigma^2 = (6)^2 = 36$, $H_1: \sigma^2 \neq 36$.

b. $\alpha = 0,05$.

c. De toetsingsgrootheid is $\chi^2 = (15 - 1) s^2/36$.

d. Onder H_0 volgt deze grootheid een χ^2 -verdeling met 14 vrijheidsgraden.

e. Bereken:

$$\chi^2_0(14) = \frac{14 \times 16}{36} = 6,22 .$$

f. Uit tabel T blijkt dat $0,025 < P_L(\chi^2_0 = 6,22) < 0,05$, zodat de tweezijdige overschrijdingskans van χ^2_0 tussen 0,05 en 0,10 ligt. H_0 wordt dus bij de gekozen 5% drempel niet verworpen.

14.2. Een farmaceutische industrie beschikt over een machine, die een product aflevert waarvan de gewichtsspreiding ten hoogste 8 mg mag zijn. Men trekt uit de productiestroom van deze machine aselechte steekproef van 25 exemplaren en vindt hiervoor $s_0 = 10$ mg. Geeft deze uitkomst aanleiding om de productie stop te zetten?

Nu dient rechts éénzijdig te worden getoetst en wel als volgt:

a. $H_0: \sigma^2 = (8)^2 = 64$ (men kan ook zetten: $\sigma^2 \leq 64$), $H_1: \sigma^2 > 64$.

b. Kies $\alpha = 0,10$ (men kiest deze grote waarde voor α , omdat men liever het risico loopt de productie onnodig te onderbreken, dan het risico exemplaren met te grote spreiding af te leveren).

c. De toetsingsgrootheid is: $\chi^2 = (25 - 1) s^2/64$.

d. Als wordt aangenomen, dat de gewichten van de producten pseudo-normaal verdeeld zijn, volgt deze grootheid onder H_0 bij benadering een χ^2 -verdeling met 24 vrijheidsgraden.

e. Bereken:

$$\chi^2_0 = \frac{24 (10)^2}{64} = \frac{2400}{64} = 37,5 .$$

- f. Verwerp H_0 (zet de productie stop), daar $P_R(\chi^2 = 37,5 \mid \nu = 24)$ kleiner is dan 0,10 (tabel T laat zien dat $\chi^2_{0,90}(\nu = 24) = 33,196$ en $\chi^2_{0,95}(\nu = 24) = 36,415$).

14.1.2. GROTE STEEKPROEVEN

Volgens (13.8) is voor grote steekproeven de grootheid

$$(14.2) \quad \underline{T} = \frac{s_x - \sigma}{\sigma} = \frac{(s_x - \sigma) \sqrt{2(n-1)}}{\sqrt{2(n-1)}\sigma}$$

bij benadering standaardnormaal verdeeld (de benadering wordt beter, naarmate n groter is). In dit geval kan de toetsing dus worden uitgevoerd door in (14.2) de waargenomen spreiding s_0 , de steekproefomvang n en de waarde van σ_0 in te vullen en de tweezijdige, resp. linkse of rechtse kans van de verkregen T_0 in tabel A op te zoeken.

Voorbeeld

14.3. Een aselechte steekproef van 99 elementen uit een normaal verdeelde populatie levert op: $s_0 = 3$. Is deze uitkomst te verenigen met de onderstelling, dat $\sigma = 4$?

- a. $H_0: \sigma = 4, H_1: \sigma \neq 4$.
- b. Kies $\alpha = 0,05$.
- c. De toetsingsgrootheid is $\underline{T} = (s_x - 4) \sqrt{2(99 - 1)}/4$.
- d. Onder H_0 volgt de grootheid \underline{T} bij benadering een standaardnormale verdeling.
- e. Bereken: $T_0 = (3 - 4) \sqrt{196}/4 = -3,5$.
- f. Uit tabel A blijkt, dat $P_D(T = -3,5) = 2(0,000233) \cong 0,0005$. H_0 wordt verworpen; met een onbetrouwbaarheid die ongeveer gelijk is aan 0,0005 kan men beweren, dat de populatiespreiding kleiner dan 4 is.

14.2. Betrouwbaarheidsgrenzen van de variantie en de spreiding

14.2.1. KLEINE STEEKPROEVEN ($n \leq 100$)

De redenering die bij het aangeven van betrouwbaarheidsgrenzen voor de variantie gevolgd wordt, is in principe die van 7.2. Volgens (14.1) is de kans, dat de grootheid $(n - 1) s^2/\sigma^2$ een waarde aanneemt die $\chi^2_{0,95}$ overschrijdt, gelijk aan 0,05. Evenzo is de kans van een waarde beneden $\chi^2_{0,05}$ gelijk aan 0,05. De kans dat deze grootheid een waarde tussen $\chi^2_{0,05}$ en $\chi^2_{0,95}$ aanneemt is dus $1 - 2(0,05) = 0,90$. Of in het algemeen

$$(14.3) \quad P \left[\chi^2_{1-\alpha} < \frac{(n - 1) s^2}{\sigma^2} < \chi^2_{1-\alpha} \right] = 1 - \alpha.$$

De ongelijkheid tussen de grote haken kan echter als volgt geschreven worden (bedenk, dat $6 < 8 < 10$, maar dat $\frac{1}{6} > \frac{1}{8} > \frac{1}{10}$, of in het algemeen als $a < b$ en a en b positief zijn dat $\frac{1}{a} > \frac{1}{b}$):

$$\frac{1}{\chi^2_{1-\alpha}} > \frac{\sigma^2}{(n-1)s^2} > \frac{1}{\chi^2_{1-\frac{1}{2}\alpha}}.$$

Als elk lid van deze ongelijkheid met $(n-1)s^2$ wordt vermenigvuldigd, verkrijgt men:

$$\frac{(n-1)s^2}{\chi^2_{1-\alpha}} > \sigma^2 > \frac{(n-1)s^2}{\chi^2_{1-\frac{1}{2}\alpha}}.$$

Men kan zodoende voor (14.3) ook schrijven

$$(14.4) \quad P \left[\frac{(n-1)s^2}{\chi^2_{1-\alpha}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\frac{1}{2}\alpha}} \right] = 1 - \alpha.$$

Men dient zich nu goed te realiseren, dat (14.4) geen uitspraak inhoudt betreffende de kans, dat de grootheid σ^2 een waarde tussen twee bepaalde grenzen inneemt. σ^2 is immers geen stochastische grootheid, maar een constante: de variantie van de populatieverdeling. Trekt men aselekt één steekproef en heeft deze als variantie s_0^2 , dan vindt men bij substitutie van deze waarde in de ongelijkheid tussen de grote haken in (14.4)

$$(14.5) \quad \frac{(n-1)s_0^2}{\chi^2_{1-\alpha}} < \sigma^2 < \frac{(n-1)s_0^2}{\chi^2_{\frac{1}{2}\alpha}}.$$

Deze ongelijkheid bevat in het geheel geen stochastisch element, daar het linker- en rechterlid bekende getallen zijn en σ^2 de (onbekende) parameter is. Als men slechts één geïsoleerd geval beschouwt, kan men van de ongelijkheid (14.5) slechts zeggen, dat zij juist of onjuist is, zonder dat men weet welke van deze twee mogelijkheden gerealiseerd is. Bij aselecte trekking is s^2 echter wèl een stochastische grootheid en de betekenis van (14.4) is: de kans, dat bij aselecte trekking het interval $(n-1)s^2/\chi^2_{1-\alpha}$ tot $(n-1)s^2/\chi^2_{\frac{1}{2}\alpha}$ de populatievariantie σ^2 bevat, is $1-\alpha$. Deze grenzen verschaffen ons dus een *betrouwbaarheidsinterval* met een betrouwbaarheid $1-\alpha$, of wel: een 100 $(1-\alpha)\%$ betrouwbaarheidsinterval.

Ter verduidelijking van het voorgaande beschouwen wij de vijf steekproeven van 5 elementen in tabel 13.2 die aselekt werden getrokken uit een pseudo-normaal verdeelde populatie. In tabel 14.1 is voor elk van deze steekproeven en de steekproef van 25 elementen, die zij tezamen opleveren, het 90% en het 80% betrouwbaarheidsinterval van σ^2 berekend (ga deze berekeningen en de getallen waarop zij berusten na!). De verkregen intervallen zijn in figuur 14.1 grafisch voorgesteld.

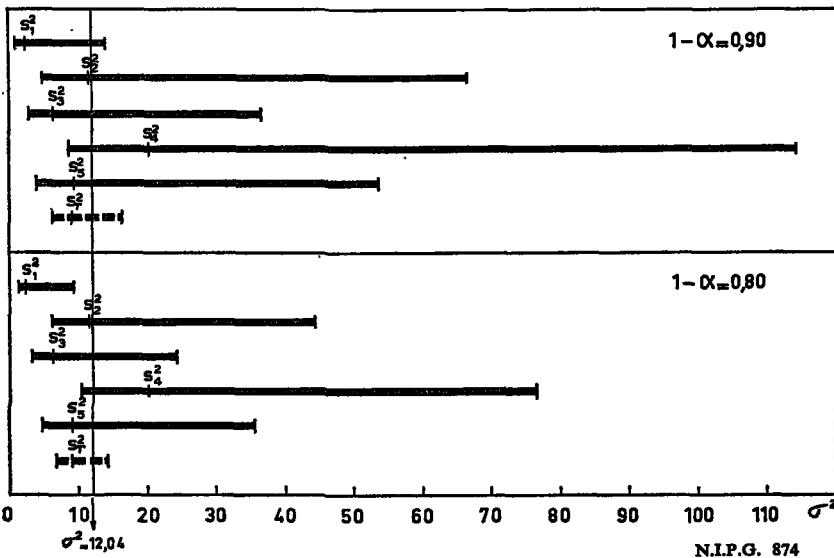
BETROUWBAARHEIDSGRENZEN VAN DE VARIANTIE EN DE SPREIDING 14.2

Tabel 14.1. Betrouwbaarheidsintervallen van σ^2 voor de steekproeven in tabel 13.2 met $n = 5$ en de gecombineerde steekproef met $n = 25$

$n = 5, n - 1 = 4, \chi^2_{0,05} = 0,711, \chi^2_{0,10} = 1,064, \chi^2_{0,90} = 7,779, \chi^2_{0,95} = 9,488$
 $n = 25, n - 1 = 24, \chi^2_{0,05} = 13,848, \chi^2_{0,10} = 15,659, \chi^2_{0,90} = 33,196, \chi^2_{0,95} = 36,415$

Steekproef	n	s^2	$(n-1) s^2$	$(n-1) s^2$		$(n-1) s^2$	
				$\chi^2_{0,95}$	$\chi^2_{0,05}$	$\chi^2_{0,90}$	$\chi^2_{0,10}$
1	5	2,5	10,0	1,05	14,06	1,29	9,40
2	5	11,8	47,2	4,97	66,39	6,07	44,36
3	5	6,5	26,0	2,74	36,57	3,34	24,44
4	5	20,3	81,2	8,56	114,21	10,44	76,32
5	5	9,5	38,0	4,01	53,45	4,88	35,71
Totaal	25	9,36	224,64	6,17 16,22 90% grenzen		6,77 14,35 80% grenzen	

Uit de tabel en de figuur blijkt in de eerste plaats, dat men bij $n = 5$ slechts een zeer grove schatting van σ^2 verkrijgt. De betrouwbaarheidsintervallen zijn breed en niet symmetrisch t.o.v. s^2 : deze steekproefvariantie ligt veel dichterbij de benedengrens dan bij de bovengrens van het interval. Bij $n = 25$ wordt reeds een veel nauwkeuriger schatting van σ^2 verkregen, terwijl de schattingsintervallen minder asymmetrisch t.o.v. s^2 zijn. Verder ziet men, dat bij een bepaalde n en s^2



Figuur 14.1. Grafiek van de betrouwbaarheidsintervallen in tabel 14.1.

de intervalbreedte toeneemt, naarmate men een grotere betrouwbaarheid wenst. Bij de kleine steekproefomvang blijkt dan bij de variantie vooral de bovengrens van het interval aanzienlijk hoger te worden. Bij het 90% interval zijn immers de benedengrenzen bij $n = 5$ ongeveer 18% lager dan de benedengrenzen van het 80% interval, maar de bovengrenzen van het 90% interval zijn bijna 50% hoger dan die van het 80% interval.

Het is duidelijk, dat men de betrouwbaarheidsgrenzen van de spreiding (standaarddeviatie) kan verkrijgen door de wortel uit de grenswaarden van het betrouwbaarheidsinterval van de variantie te nemen. Zo vindt men bv. voor steekproef 1 (tabel 14.1) de volgende betrouwbaarheidsintervallen voor de spreiding:

$$90\% \text{ betrouwbaarheid: } 1,02 < \sigma < 3,75 ,$$

$$80\% \text{ betrouwbaarheid: } 1,14 < \sigma < 3,07 .$$

In tabel 14.2 zijn nog enige voorbeelden van de berekening van betrouwbaarheidsintervallen van σ^2 gegeven, waarbij verschillende waarden van n , s^2 en $1 - \alpha$ zijn gekozen. Zoek — bij wijze van oefening — de in deze tabel gegeven waarden van $\chi^2_{\frac{1}{2}\alpha}$ en $\chi^2_{1-\frac{1}{2}\alpha}$ ook zelf in tabel T op.

Tabel 14.2. Berekening van betrouwbaarheidsintervallen van de variantie

Steekproef	n	$\nu = n-1$	s^2	α	$1-\alpha$	$\chi^2_{\frac{1}{2}\alpha}$	$\frac{(n-1)s^2}{\chi^2_{\frac{1}{2}\alpha}}$	$\chi^2_{1-\frac{1}{2}\alpha}$	$\frac{(n-1)s^2}{\chi^2_{1-\frac{1}{2}\alpha}}$	100 (1- α)% betrouwbaarheidsinterval
1	5	4	10	0,10	0,90	0,711	56,3	9,488	4,2	$4,2 < \sigma^2 < 56,3$
2	5	4	10	0,02	0,98	0,297	134,7	13,277	3,0	$3,0 < \sigma^2 < 134,7$
3	10	9	10	0,10	0,90	3,325	27,1	16,919	5,3	$5,3 < \sigma^2 < 27,1$
4	10	9	5	0,10	0,90	3,325	13,5	16,919	2,7	$2,7 < \sigma^2 < 13,5$
5	10	9	5	0,05	0,95	2,700	16,7	19,023	2,4	$2,4 < \sigma^2 < 16,7$
6	20	19	10	0,01	0,99	6,844	27,8	38,582	4,9	$4,9 < \sigma^2 < 27,8$
7	20	19	10	0,10	0,90	10,117	18,8	30,144	6,3	$6,3 < \sigma^2 < 18,8$
8	30	29	15	0,10	0,90	17,708	24,6	42,557	10,2	$10,2 < \sigma^2 < 24,6$
9	30	29	10	0,05	0,95	16,047	18,1	45,722	6,3	$6,3 < \sigma^2 < 18,1$
10	30	29	5	0,01	0,99	13,121	11,1	52,336	2,8	$2,8 < \sigma^2 < 11,1$

14.2.2. GROTE STEEKPROEVEN

Voor $n > 100$ kan het 100 (1 - α)% betrouwbaarheidsinterval van de spreiding bij benadering worden verkregen door te nemen

$$(14.6.1) \quad s + T_{\frac{1}{2}\alpha} \frac{s}{\sqrt{2(n-1)}} < \sigma < s + T_{1-\frac{1}{2}\alpha} \frac{s}{\sqrt{2(n-1)}} .$$

Men kan voor dit schattingsinterval ook schrijven

$$(14.6.2) \quad s \pm T_{1-\alpha} \frac{s}{\sqrt{2(n-1)}} .$$

Voorbeeld

14.4. Een aselechte steekproef van 129 elementen uit een normaal verdeelde populatie levert op $s_0 = 5$. Bepaal het 95% betrouwbaarheidsinterval van σ .

Daar $1 - \alpha = 0,95$, is $\alpha = 0,05$, zodat $T_{1-\alpha} = T_{0,025} = -T_{1-1-\alpha} = -T_{0,975} = -1,96$ (tabel B). Volgens (14.6) is dus het gevraagde 95% interval bij benadering:

$$5 - 1,96 \frac{5}{\sqrt{2(129-1)}} < \sigma < 5 + 1,96 \frac{5}{\sqrt{2(129-1)}} ,$$

$$4,39 < \sigma < 5,61 .$$

14.2.3. TOETSEN EN SCHATTEN

Uit de voorgaande paragrafen blijkt nog eens duidelijk de nauwe samenhang die bestaat tussen het toetsen van een hypothese en het bepalen van een betrouwbaarheidsinterval. Met behulp van de door (14.1) gegeven grootte kan men hypothesen toetsen van de vorm $\sigma^2 = \sigma_0^2$. Men kan nu gemakkelijk nagaan, dat alle waarden van σ^2 die — op grond van gegeven s_0^2 en bij gegeven onbetrouwbaarheidsdrempel α — op deze wijze *niet* verworpen kunnen worden, tezamen juist het uit s_0^2 en de ongelijkheid in (14.4) volgende schattingsinterval met betrouwbaarheid $1 - \alpha$ vormen. Omgekeerd kan men dus ook zeggen, dat bij gegeven s_0^2 en α (resp. $1 - \alpha$), alle waarden van σ^2 die buiten het betrouwbaarheidsinterval liggen voor verwerping in aanmerking komen.

Dit parallelisme tussen de toetsingstheorie en de theorie der betrouwbaarheidsintervallen geldt in zeer algemene vorm. Men zou een betrouwbaarheidsinterval dan ook kunnen definiëren als de verzameling van die hypothesen, die, op grond van een bepaalde toetsingsmethode, bepaalde waarnemingsuitkomsten en een bepaalde drempelwaarde, *niet* voor verwerping in aanmerking komen.

14.3. Toets voor de hypothese $\sigma_1^2 = \sigma_2^2$ = F-toets

Deze toets kan worden toegepast als men, op grond van twee aselechte steekproeven uit twee normaal verdeelde populaties, de hypothese wil toetsen dat de populatiespreidingen gelijk zijn.

Onderstel, dat de steekproef uit de ene populatie bestaat uit n_1 elementen met variantie s_1^2 en die uit de tweede populatie uit n_2

elementen met variantie s_2^2 . De toetsing verloopt dan als volgt:

Tweezijdige toetsing

a. $H_0: \sigma_1^2 = \sigma_2^2 = \sigma^2$, $H_1: \sigma_1^2 \neq \sigma_2^2$.

b. Kies α .

c. De toetsingsgrootheid is $\underline{F} = s_1^2/s_2^2$.

d. Onder H_0 volgt de grootheid \underline{F} volgens (13.13) een F -verdeling met $\nu_1 = n_1 - 1$ en $\nu_2 = n_2 - 1$.

e. Bereken F_0 . Om de toets zo eenvoudig mogelijk uit te voeren, plaatst men hierbij altijd de *grootste* steekproefvariantie in de *teller*. D.w.z. men berekent

$$F_0 = s_1^2/s_2^2 \text{ als } s_1^2 > s_2^2 \text{ is en } F_0 = s_2^2/s_1^2 \text{ als } s_1^2 < s_2^2 \text{ is.}$$

f. Verwerp H_0 , als $P_D(F_0) \leq \alpha$ is, d.w.z. als $F_0 > F_{1-\alpha}$. Let er op, dat voor $s_1^2 > s_2^2$ het aantal vrijheidsgraden $\nu_1 = n_1 - 1$ en $\nu_2 = n_2 - 1$ en voor $s_1^2 < s_2^2$ het aantal vrijheidsgraden $\nu_1 = n_2 - 1$ en $\nu_2 = n_1 - 1$ is. Toetst men met $\alpha = 0,05$, dan vindt men de kritieke waarde $F_{0,975}$ in tabel F-2; is $\alpha = 0,01$ dan is de kritieke waarde $F_{0,995}$ in tabel F-4 te vinden.

Links éézijdige toetsing

a. $H_0: \sigma_1^2 = \sigma_2^2$ (men kan ook zeggen: $\sigma_1^2 \geq \sigma_2^2$), $H_1: \sigma_1^2 < \sigma_2^2$.

Men kan nu zonder meer tot het niet verwerpen van H_0 besluiten, indien $s_1^2 \geq s_2^2$ is. Is dit niet het geval, dan berekent men $F_0 = s_2^2/s_1^2$ en verwerpt H_0 , indien $P_R(F_0) \leq \alpha$ is, d.w.z. voor $F_0 \geq F_{1-\alpha}$ ($\nu_1 = n_2 - 1$ en $\nu_2 = n_1 - 1$). Toetst men dus met $\alpha = 0,05$, dan vindt men de kritieke waarde $F_{0,95}$ in tabel F-1 en toetst men met $\alpha = 0,01$ dan vindt men de kritieke waarde $F_{0,99}$ in tabel F-3.

Rechts éézijdige toetsing

a. $H_0: \sigma_1^2 = \sigma_2^2$ (of: $\sigma_1^2 \leq \sigma_2^2$), $H_1: \sigma_1^2 > \sigma_2^2$.

Men kan onmiddellijk tot het niet verwerpen van H_0 overgaan, indien $s_1^2 \leq s_2^2$ is. Is dit niet het geval, dan berekent men $F_0 = s_1^2/s_2^2$ en verwerpt H_0 , indien $P_R(F_0) \leq \alpha$ is, d.i. voor

$$F_0 \geq F_{1-\alpha} \text{ (} \nu_1 = n_1 - 1 \text{ en } \nu_2 = n_2 - 1 \text{)} .$$

Voorbeelden

14.5. Van twee steekproeven uit normaal verdeelde populaties is gegeven: $n_1 = 10$, $s_1^2 = 11,72$ en $n_2 = 31$, $s_2^2 = 3,17$. Toets met $\alpha = 0,01$ de hypothese, dat beide populaties dezelfde spreiding bezitten.

Er dient tweezijdig te worden getoetst. Men vindt: $F_0(9, 30) =$

TOETSEN VOOR HET VERGELIJKEN VAN MEER DAN TWEE VARIANTIES 14.4

= 11,72/3,17 = 3,70. In tabel F-4 vindt men: $F_{1-\alpha}(v_1, v_2) = F_{0,995}(9, 30) = 3,45$, zodat $P_D(3,45) = 0,01$. Daar F_0 groter is dan $F_{0,995}$ wordt H_0 verworpen.

14.6. Voor twee steekproeven uit normaal verdeelde populaties geldt: $n_1 = 16$, $s_1^2 = 14,25$ en $n_2 = 25$, $s_2^2 = 33,21$. Toets bij een drempelwaarde 0,05 de hypothese $H_0: \sigma_1^2 \geq \sigma_2^2$.

Nu luidt de alternatieve hypothese $H_1: \sigma_1^2 < \sigma_2^2$, zodat éézijdige toetsing plaatsvindt. Men vindt: $F_0(24, 15) = 33,21/14,25 = 2,33$. Tabel F-1 levert op: $F_{1-\alpha} = F_{0,95}(24, 15) = 2,29$. Daar F_0 groter is dan $F_{0,95}$ wordt H_0 verworpen.

14.4. Toetsen voor het vergelijken van meer dan twee varianties

Als men beschikt over k aselechte steekproeven, resp. van de omvang n_1, n_2, \dots, n_k , uit k normaal verdeelde populaties, kan de hypothese $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$ op verschillende manieren worden getoetst. Wij behandelen hier beknopt de toetsen van BARTLETT, HARTLEY, COCHRAN en DOORNBOS.

14.4.1. DE TOETS VAN BARTLETT

Dit is een toets voor de hypothese $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$, die gevoelig is voor verschillen van allerlei aard tussen de populatievarianties.

Onder H_0 zijn alle steekproefvarianties (zuivere) schattingen van de gemeenschappelijke parameter σ^2 . Een zuivere schatting s^2 van σ^2 kan dan worden verkregen met

$$(14.7) \quad s^2 = \frac{\sum (n_i - 1) s_i^2}{\sum (n_i - 1)}$$

Voor het toepassen van de toets dient men nu vervolgens te berekenen

$$(14.8.1) \quad M = \sum (n_i - 1) \cdot \log_e s^2 - \sum (n_i - 1) \log_e s_i^2,$$

of, als in plaats van de natuurlijke logaritmen met grondtal e gebruik wordt gemaakt van logaritmen met grondtal 10^1

$$(14.8.2) \quad M = 2,3026 [\sum (n_i - 1) \cdot \log_{10} s^2 - \sum (n_i - 1) \log_{10} s_i^2].$$

Noem verder:

$$(14.9) \quad C = 1 + \frac{1}{3(k-1)} \left[\sum \frac{1}{n_i - 1} - \frac{1}{\sum (n_i - 1)} \right].$$

BARTLETT heeft aangetoond, dat voor $n_i \geq 6$ de grootheid

$$(14.10) \quad \underline{B} = \underline{M}/C$$

¹ Zie bijlage 2, par. 4.

onder H_0 bij benadering een χ^2 -verdeling met $k-1$ vrijheidsgraden volgt. H_0 kan dus met een onbetrouwbaarheid α verworpen worden, indien $B_0 \geq \chi^2_{1-\alpha} (\nu = k-1)$ is. Daar C altijd groter is dan 1 kan men de berekening van B_0 achterwege laten, wanneer M_0 reeds insignificant blijkt te zijn.

In tabel 14.3 is een voorbeeld gegeven van de toepassing van de toets van BARTLETT op vier steekproeven van verschillende omvang, waaruit tevens blijkt op welke wijze men de berekeningen het eenvoudigst kan uitvoeren.

Tabel 14.3. De toets van BARTLETT toegepast op vier steekproeven van verschillende omvang

Steekproef	n_i	n_i-1	$\frac{1}{n_i-1}$	$\Sigma (x-\bar{x})^2$	s_i^2	$\log s_i^2$	$(n_i-1) \log s_i^2$
1	31	30	0,033333	164	5,47	0,73799	22,1397
2	15	14	0,071429	65	4,64	0,66652	9,3313
3	20	19	0,052632	218	11,47	1,05961	20,1326
4	42	41	0,024390	463	11,29	1,05271	43,1611
Totaal	104		0,181784	910			94,7647
	$\Sigma (n_i-1)$		$\Sigma \frac{1}{n_i-1}$	$\Sigma \Sigma (x-\bar{x})^2$			$\Sigma (n_i-1) \log s_i^2$

Berekeningen

$$(14.7) \quad s^2 = \frac{\Sigma (n_i - 1) s_i^2}{\Sigma (n_i - 1)} = \frac{\Sigma \Sigma (x - \bar{x})^2}{\Sigma (n_i - 1)} = \frac{910}{104} = 8,75,$$

$$\log s^2 = 0,94201, \quad \Sigma (n_i - 1) \cdot \log s^2 = (104) (0,94201) = 97,9690$$

$$(14.8.2) \quad M_0 = (2,3026) (97,9690 - 94,7647) = (2,3026) (3,2043) = 7,38$$

$$(14.9) \quad C = 1 + \frac{1}{3 \times 3} \left[0,181784 - \frac{1}{104} \right] = 1 + \frac{0,181784 - 0,009615}{9} = 1,0191$$

$$(14.10) \quad B_0 = 7,38/1,02 = 7,24 \quad (\nu = k-1 = 3), \quad 0,05 < P_R < 0,10.$$

Als de steekproeven minder dan 6 elementen tellen is de χ^2 -benadering niet betrouwbaar. Men kan dan gebruik maken van de door MERRINGTON en THOMPSON (180) vervaardigde tabellen, waarin voor $k = 3$ (1) 15 en voor verschillende waarden van $c_1 = 3$ ($k-1$) ($C-1$) rechter kritieke waarden van \underline{M} met $\alpha = 0,05$ en $0,01$ gegeven zijn. Deze tabellen zijn ook te vinden in de BIOMETRIKA TABLES (74, tabel 32) en in BENNETT en FRANKLIN (4, blz. 199).

TOETSEN VOOR HET VERGELIJKEN VAN MEER DAN TWEE VARIANTIES 14.4

Wanneer de k steekproeven van dezelfde omvang zijn, zodat $n_1 = n_2 = \dots = n_k = n$, gaat (14.7) over in

$$(14.11) \quad s^2 = \frac{(n-1) \sum s_i^2}{k(n-1)} = \frac{\sum s_i^2}{k}.$$

Daar dan tevens geldt, dat $\sum (n_i - 1) \cdot \log s_i^2 = (n-1) \sum \log s_i^2$ en $\sum (n_i - 1) = k(n-1)$, kan men voor (14.8.2) schrijven

$$(14.12) \quad M = (2,3026) (n-1) \left[k \log \frac{\sum s_i^2}{k} - \sum \log s_i^2 \right].$$

Tenslotte komt dan in de plaats van (14.9):

$$(14.13) \quad C = 1 + \frac{k+1}{3k(n-1)}.$$

In tabel 14.4 is een voorbeeld gegeven van de manier, waarop de berekeningen in dit geval het best kunnen worden uitgevoerd.

Tabel 14.4. De toets van BARTLETT: 5 steekproeven van dezelfde omvang

Steekproef	n_i	$n_i - 1$	$\sum (x - \bar{x})^2$	s_i^2	$\log s_i^2$
A	9	8	151	18,875	1,2759
B	9	8	227	28,375	1,4529
C	9	8	177	22,125	1,3449
D	9	8	120	15,000	1,1761
E	9	8	183	22,875	1,3594
Totaal	$k = 5$			$\sum s_i^2 = 107,250$	$6,6092 = \sum \log s_i^2$

Berekeningen

$$(14.11) \quad s^2 = \frac{\sum s_i^2}{k} = \frac{107,25}{5} = 21,45, \log s^2 = 1,3314, k \log s^2 = 6,6570$$

$$(14.12) \quad M_0 = 2,3026 (9-1) (6,6570 - 6,6092) = 0,88.$$

Deze M_0 is zo laag, dat de berekening van B_0 achterwege kan blijven. De conclusie luidt dus, dat H_0 niet wordt verworpen.

Voert men toch de volledige berekening uit, dan is:

$$(14.13) \quad C = 1 + \frac{5+1}{3 \times 5 \times 8} = 1,05, \text{ zodat: } B_0 = \frac{0,88}{1,05} = 0,84 (k-1=4).$$

Uit tabel T blijkt, dat de rechtse kans van deze uitkomst tussen 0,90 en 0,95 ligt.

14.4.2. DE TOETS VAN HARTLEY

Evenals de toets van BARTLETT is deze toets bestemd voor het toetsen van de hypothese $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$, met als alternatieve hypothese dat de varianties van de k normaal verdeelde populaties verschillen vertonen. Een belangrijk voordeel is echter, dat er vrijwel niet gerekend behoeft te worden; hiertegenover staat als nadeel, dat de k steekproeven (ongeveer) *dezelfde omvang* n moeten hebben.

De toetsingsgrootheid bij de toets van HARTLEY is

$$(14.14) \quad \underline{F}(\max) = s_{\max}^2 / s_{\min}^2 ,$$

waarin $\left\{ \begin{array}{l} s_{\max}^2 = \text{de grootste} \\ s_{\min}^2 = \text{de kleinste} \end{array} \right\}$ van de k steekproefvarianties.

Daar men H_0 alleen zal willen verwerpen, als $F(\max)$ een te grote waarde aanneemt wordt rechts éézijdig getoetst. Hiertoe zijn door HARTLEY (165) voor $k = 2$ t/m 12 en een reeks waarden van $n - 1$ bij benadering de kritieke waarden $F_{0,95}$ en $F_{0,99}$ van de verdeling van $\underline{F}(\max)$ onder H_0 berekend en deze zijn in tabel P opgenomen. De toetsing kan dus worden uitgevoerd, door $F_0(\max)$ te berekenen en H_0 te verwerpen:

met een onbetrouwbaarheid 0,05, indien $F_0(\max) \geq F_{0,95}(k, n-1)$,

met een onbetrouwbaarheid 0,01, indien $F_0(\max) \geq F_{0,99}(k, n-1)$.

Voorbeelden

14.7. Vier aselechte steekproeven uit pseudo-normaal verdeelde populaties, ieder bestaande uit 30 elementen, bezitten de volgende varianties:

$$s_1^2 = 16,72 \quad s_2^2 = 36,27 \quad s_3^2 = 14,00 \quad s_4^2 = 15,91 .$$

Men wenst met $\alpha = 0,05$ de hypothese $H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$ te toetsen.

Past men de toets van HARTLEY toe, dan is:

$$F_0(\max) = 36,27/14,00 = 2,59.$$

In tabel P vindt men bij $k = 4$ en $n - 1 = 30$: $F_{0,95}(\max) = 2,61$. Voor $n - 1 = 29$ zal deze kritieke waarde nog iets groter zijn (immers; bij bepaalde k en afnemende steekproefomvang neemt deze waarde toe: $F_{0,95}(\max)$ met $n - 1 = 20$ is bv. reeds 3,29 — zie tabel P). F_0 is dus (even) kleiner dan de kritieke waarde, zodat H_0 (nog juist) niet kan worden verworpen.

14.8. Soms kan men de toets ook toepassen als de steekproeven een verschillende omvang hebben. Onderstel bv., dat vier aselechte

TOETSEN VOOR HET VERGELIJKEN VAN MEER DAN TWEE VARIANTIES 14.4

steekproeven uit pseudo-normaal verdeelde populaties opleveren:

- | | | | |
|---------------|-----------------|---------------|-----------------|
| 1. $n_1 = 30$ | $s_1^2 = 11,22$ | 2. $n_2 = 22$ | $s_2^2 = 10,30$ |
| 3. $n_3 = 22$ | $s_3^2 = 6,20$ | 4. $n_4 = 15$ | $s_4^2 = 9,79$ |

Hier is $F_0(\max) = 11,22/6,20 = 1,81$ met $k = 4$, maar $n - 1$ varieert van 14 t/m 29. Nu hebben wij bij het voorgaande voorbeeld reeds gezien, dat bij bepaalde k en α de kritieke waarde $F_{1-\alpha}(\max)$ toeneemt, naarmate $n - 1$ lager wordt. Als dus F_0 kleiner is dan $F_{1-\alpha}$ voor de *grootste* waarde van $n - 1$ die bij de beschouwde steekproeven voorkomt, zal deze F_0 zeker niet significant zijn bij een kleinere. Nu blijkt uit tabel P, dat voor $\alpha = 0,05$ en:

$$k = 4 \text{ en } n - 1 = 30 \text{ de kritieke waarde is: } F_{0,95}(\max) = 2,61$$

en dat zelfs voor

$$k = 4 \text{ en } n - 1 = 20 \text{ de kritieke waarde is: } F_{0,95}(\max) = 3,29.$$

Het is dus duidelijk, dat H_0 niet kan worden verworpen.

Als de toets bij steekproeven van ongelijke omvang zou leiden tot het verwerpen van H_0 bij het gebruik van de grootste $n - 1$ en tot niet verwerpen bij het gebruik van de kleinste $n - 1$, dan is een ondubbeltzinnige uitspraak niet mogelijk en dient men de toets van BARTLETT toe te passen. Dit is — zoals de lezer zelf kan nagaan — bv. het geval bij de gegevens in tabel 14.3.

14.4.3.* DE TOETS VAN COCHRAN

Deze toets kan worden gebruikt, als men op basis van k aselechte steekproeven van dezelfde omvang n de hypothese

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$$

wil toetsen, met als alternatieve hypothese dat de steekproef met de grootste variantie stamt uit een populatie, waarvan de variantie groter is dan die van de overige $k - 1$ populaties.

De toetsingsgrootheid is

$$(14.15) \quad \underline{g} = \frac{s_{\max}^2}{\sum_{i=1}^k s_i^2}$$

en ook hier wordt rechts éézijdig getoetst. In DIXON & MASSEY (8, tabel 17) en EISENHART (8a, blz. 390—391) zijn de door COCHRAN (137) berekende tabellen opgenomen, waarin voor verschillende waarden van k en $n - 1$ de fractielen $g_{0,95}$ en $g_{0,99}$ van de verdeling van \underline{g} onder H_0 zijn opgenomen. Deze fractielen vormen dus kritieke waarden bij éézijdige toetsing met $\alpha = 0,05$ en $\alpha = 0,01$.

14.4.4.* DE TOETS VAN DOORNBOS

Dit is een analogon van de toets van COCHRAN, dat de kleinste steekproefvariantie betreft. Deze toets kan dus worden toegepast, als men op grond van k aselecte steekproeven van dezelfde omvang n de hypothese $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2 = \sigma^2$ wil toetsen met als alternatief, dat de steekproef met de kleinste variantie afkomstig is uit een populatie, waarvan de variantie kleiner is dan die van de overige $k - 1$ populaties. De toetsingsgrootheid is

$$(14.16) \quad \underline{a} = \frac{s^2_{min}}{\Sigma s_i^2}$$

en er wordt links éézijdig getoetst. DOORNBOS (147) geeft voor $n - 1 = 1, 2, \dots, 6$ en $k = 2, 3, \dots, 10, 12, 15$ en 20 de fractiën $a_{0,05}$ van de verdeling van \underline{a} onder H_0 .

14.5.* De toets tegen verloop van Bellinson

Deze toets tegen verloop in een reeks van n waarnemingen x_1, x_2, \dots, x_n is afkomstig van BELLINSON c.s. (126) en berust op de zg. 'mean square successive difference'

$$(14.17) \quad d^2 = \frac{\sum_{i=1}^{n-1} (x_{i+1} - x_i)^2}{n - 1}$$

Bij aselecte trekking uit een normaal verdeelde populatie volgt de grootheid $\underline{d}^2/2$ een kansverdeling met gemiddelde σ^2 , zodat deze grootheid evenals \underline{s}^2 een zuivere schatter is van de populatievariantie. De grootheid $\frac{\underline{d}^2/2}{\underline{s}^2} = \underline{d}^2/2\underline{s}^2$ blijkt dan symmetrisch verdeeld te zijn met gemiddelde 1, zodat de grootheid $\underline{k} = 1 - \underline{d}^2/2\underline{s}^2$ een symmetrische verdeling volgt met $\mu_k = 0$. De variantie van deze verdeling is

$$(14.18) \quad \sigma_k^2 = \frac{n - 2}{(n - 1)(n + 1)}$$

Treedt in de waarnemingen een geleidelijk stijgend (dalend) of een cyclisch verloop op, dan wordt de waarde van \underline{d}^2 hierdoor veel minder beïnvloed dan die van \underline{s}^2 . De waarde van $\underline{d}^2/2\underline{s}^2$ zal dan klein zijn en \underline{k} neemt dan in het algemeen een relatief grote *positieve* waarde aan. Bij het optreden van oscillatie ontstaat het tegengestelde effect, zodat \underline{k} een relatief grote *negatieve* waarde zal aannemen.

HART (164) heeft voor een aantal waarden van n tot en met 25 enkele fractiën van de verdeling van $\underline{d}^2/\underline{s}^2$ berekend, welke voor een normaal verdeelde populatie gelden bij aselecte trekking. Wij hebben deze

omgezet in de fractielen $k_{0,95}$ ($= -k_{0,05}$) en $k_{0,99}$ ($= -k_{0,01}$) van de grootheid k , die in tabel 14.5 zijn opgenomen. Deze fractielen kunnen voor $n \leq 25$ worden gebruikt bij éénzijdige toetsing van de hypothese, dat de waarnemingen een aselechte volgorde vertonen en wel met $\alpha = 0,05$ resp. $\alpha = 0,01$. Luidt de alternatieve hypothese, dat een stijgende, dalende of cyclische trend aanwezig is, dan wordt H_0 verworpen, als $k_0 \geq k_{1-\alpha}$ is. Voor de alternatieve hypothese, dat oscillatie optreedt wordt H_0 verworpen, wanneer $k_0 \leq k_\alpha$ is.

Voor $n > 25$ kan men gebruik maken van het feit, dat de grootheid k dan bij benadering normaal verdeeld is, zodat de grootheid

$$(14.19) \quad \underline{T} = \frac{k}{\sigma_k} = k \sqrt{\frac{(n-1)(n+1)}{n-2}}$$

bij benadering standaardnormaal verdeeld is.

Tabel 14.5. Fractielen van de verdeling van $k = 1 - d^2/2s^2$

n	$k_{0,95}$	$k_{0,99}$
4	0,61	0,685
5	0,59	0,73
6	0,555	0,72
7	0,53	0,695
8	0,51	0,67
9	0,49	0,645
10	0,47	0,625
11	0,45	0,605
12	0,435	0,585
15	0,395	0,54
20	0,35	0,48
25	0,315	0,435
	$-k_{0,05}$	$-k_{0,01}$

Naar HART (164).

Voorbeelden

14.9. De toets tegen verloop is toegepast op de waarnemingsreeks in tabel 14.6. De getoetste hypothese luidt, dat de waarnemingen een aselechte volgorde bezitten, terwijl de alternatieve hypothese is, dat deze volgorde niet aselekt is, zodat tweezijdig wordt getoetst. Men kiest $\alpha = 0,02$.

Uit de gegevens in tabel 14.6 volgt:

$$\Sigma (x_i - \bar{x})^2 = (n - 1) s^2 = 7714 - (330)^2/15 = 454,$$

$$\Sigma (x_{i+1} - x_i)^2 = (n - 1)d^2 = 207,$$

zodat

$$k_0 = 1 - \frac{207}{2(454)} = 1 - 0,228 = 0,772 .$$

Uit tabel 14.5 blijkt, dat voor $n = 15$ geldt, dat $k_{0,99} = 0,54$. Daar k_0 deze waarde ver overschrijdt wordt de getoetste hypothese verworpen ten gunste van het alternatief, dat een verloop optreedt.

Tabel 14.6. Toepassing van de toets tegen verloop op een reeks van 15 waarnemingen

Volgorde der waarnemingen	x_i	x_i^2	$x_{i+1} - x_i$	$(x_{i+1} - x_i)^2$
1	25	625		
2	30	900	+5	25
3	28	784	-2	4
4	24	576	-4	16
5	26	676	+2	4
6	29	841	+3	9
7	25	625	-4	16
8	23	529	-2	4
9	24	576	+1	1
10	17	289	-7	49
11	18	324	+1	1
12	20	400	+2	4
13	12	144	-8	64
14	13	169	+1	1
15	16	256	+3	9
Totaal	330	7714		207

14.10. Vervolgens passen wij de toets toe op de 50 waarnemingen in opgave 2.1 op blz. 42 en wij nemen aan dat deze waarnemingen verkregen zijn in de volgorde 14,0, 13,4, 13,6, enz. Hierbij wordt als H_0 genomen, dat de waarnemingen een aselechte volgorde bezitten, met als alternatief, dat een dalend verloop optreedt, zodat rechts éénzijdige toetsing plaatsvindt. Men kiest $\alpha = 0,01$.

Voor deze waarnemingen blijkt te gelden: $(n - 1) s^2 = 12,88$, $(n - 1) d^2 = 10,62$, zodat $k_0 = 1 - \frac{10,62}{2 \times 12,88} = 0,5877$. Het aantal waarnemingen is groter dan 25, zodat de normale benadering dient te worden gebruikt. Men vindt dan met (14.19)

$$T_0 = 0,5877 \sqrt{\frac{49 \times 51}{48}} = (0,5877) (7,215) = 4,24 .$$

De rechtse overschrijdingskans van deze uitkomst is aanzienlijk kleiner dan 0,01, zodat H_0 wordt verworpen.

14.6. Opgaven

14.1. Onderstaande gegevens betreffen aselechte steekproeven uit normaal verdeelde populaties:

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>k</i>	<i>l</i>
<i>n</i>	3	6	9	16	25	36	10	10	10	10	10	10
<i>s</i> ²	20	20	20	20	20	20	5	8	13	16	22	26

a. Toets voor elke steekproef de hypothese: $\sigma^2 = 10$. Kies $\alpha = 0,05$.

b. Bepaal voor elke steekproef het 95% betrouwbaarheidsinterval van σ^2 .

14.2 Bepaal het 95% betrouwbaarheidsinterval van σ^2 op basis van de volgende aselechte steekproeven uit normaal verdeelde populaties:

<i>a.</i> $n = 8, s^2 = 15$	<i>b.</i> $n = 16, s^2 = 8$	<i>c.</i> $n = 50, s^2 = 15$
<i>d.</i> $n = 100, s^2 = 10$	<i>e.</i> $n = 200, s^2 = 10$	<i>f.</i> $n = 500, s^2 = 10$

14.3. Onderstel normaliteit en toets de hypothese ($\alpha = 0,05$), dat onderstaande steekproefparen stammen uit populaties met dezelfde variantie:

<i>a.</i> $n_1 = 12, s_1^2 = 14,25$ en $n_2 = 10, s_2^2 = 27,93$
<i>b.</i> $n_1 = 21, s_1^2 = 30$ en $n_2 = 13, s_2^2 = 11$
<i>c.</i> $n_1 = 21, s_1^2 = 11$ en $n_2 = 13, s_2^2 = 30$
<i>d.</i> $n_1 = 121, s_1^2 = 6,32$ en $n_2 = 61, s_2^2 = 11,98$

14.4 Uit de cursus: Summarizing Experimental Data, Columbia University, U.S.A. Measurements of oral and rectal temperatures ($^{\circ}\text{C}$) were taken on ten normal individuals in a fasting state under environmental conditions of 32°C and 90% relative humidity. Does the variability of the individuals with respect to oral temperature differ significantly (at a 5% level) from that with respect to rectal temperature?

Oral: 36,89, 36,61, 36,44, 37,00, 37,11, 36,78, 37,22, 37,11, 36,89, 37,00

Rectal: 37,00, 36,83, 36,83, 37,22, 37,33, 36,83, 37,50, 37,22, 37,11, 37,28

14.5. Pas op onderstaande steekproefuitkomsten de toets van BARTLETT en zo mogelijk de toets van HARTLEY toe ($\alpha = 0,05$):

<i>a.</i>	Steekpr.	1	2	3	4
	n_i	21	29	41	61
	s_i^2	8,45	6,22	11,75	15,20

<i>b.</i>	Steekpr.	1	2	3	4	5	6
	n_i	15	15	15	15	15	15
	s_i^2	5,1	6,3	4,2	8,9	10,2	3,9

<i>c.</i>	Steekpr.	1	2	3	4	5	6	7
	n_i	22	20	30	25	19	24	31
	s_i^2	8,33	6,53	6,12	8,10	10,95	8,58	3,45

14.6. Pas de toets tegen verloop van BELLINSON toe op de waarnemingsreeks in opgave 11.6 (blz. 277).

**TECHNIEKEN BETREFFENDE HET GEMIDDELDE
VAN EEN POPULATIE**

*** 15.1. Toets voor de hypothese $\mu = \mu_0$**

15.1.1. POPULATIESPREIDING BEKEND

Onderstel, dat men beschikt over een aselechte steekproef van n elementen met gemiddelde \bar{x}_0 uit een normaal verdeelde populatie met bekende spreiding σ en de hypothese wil toetsen, dat het populatie-gemiddelde een bepaalde waarde μ_0 bezit. Men kan dan gebruik maken van de volgende toets:

Tweezijdige toetsing

a. $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$.

b. Kies α .

c. De toetsingsgrootheid is

$$(15.1) \quad \underline{T} = \frac{(\bar{x} - \mu_0) \sqrt{n}}{\sigma}.$$

d. Onder H_0 volgt de grootheid \underline{T} volgens (13.4) de standaardnormale verdeling.

e. Bereken T_0 , door in (15.1) \bar{x}_0 , μ_0 , n en σ te substitueren.

f. Verwerp H_0 , wanneer $P_D(T_0) < \alpha$, d.w.z. als: $T_0 \leq T_{1-\alpha}$ of als $T_0 \geq T_{1-\alpha}$. De kritieke waarden van \underline{T} volgen uit tabel B (of uit tabel 6.10).

Links éézijdige toetsing

a. $H_0: \mu = \mu_0$ (of: $\mu \geq \mu_0$), $H_1: \mu < \mu_0$.

Besluit H_0 niet te verwerpen, als $\bar{x}_0 \geq \mu_0$ is. Wanneer dit niet het geval is, bereken dan T_0 en:

f. Verwerp H_0 , indien $P_L(T_0) < \alpha$ is, d.i. als $T_0 < T_\alpha$ is.

Rechts éézijdige toetsing

a. $H_0: \mu = \mu_0$ (of: $\mu \leq \mu_0$), $H_1: \mu > \mu_0$.

Besluit H_0 niet te verwerpen, als $\bar{x}_0 \leq \mu_0$ is. Bereken anders T_0 en:

f. Verwerp H_0 , indien $P_R(T_0) < \alpha$ is, d.i. voor $T_0 > T_{1-\alpha}$.

Deze toets zal men in de praktijk bij problemen op medisch-biologisch gebied slechts zelden kunnen gebruiken, omdat daarbij ook de populatiespreiding vaak onbekend is. Wij hebben hem volledigheidshalve opgenomen, maar zien af van het geven van voorbeelden.

15.1.2. POPULATIESPREIDING ONBEKEND

De nu volgende toets kan worden toegepast, als men beschikt over een aselechte steekproef van n elementen met gemiddelde \bar{x}_0 en spreiding s_0 en de hypothese wil toetsen, dat het populatiegemiddelde gelijk is aan μ_0 . Deze toets wordt gewoonlijk de toets van STUDENT voor de hypothese $\mu = \mu_0$ genoemd. Deze situatie doet zich in de praktijk wél dikwijls voor en de toetsing kan dan op de volgende wijze worden uitgevoerd:

Tweezijdige toetsing

- a. $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$.
 b. Kies α .
 c. De toetsingsgrootheid is

$$(15.2) \quad t = \frac{(\bar{x} - \mu_0) \sqrt{n}}{s}.$$

- d. Onder H_0 volgt deze grootheid een STUDENT-verdeling met $n - 1$ vrijheidsgraden.
 e. Bereken t_0 , door in (15.2) \bar{x}_0 , s_0 , μ_0 en n te substitueren.
 f. Verwerp H_0 , als $P_D(t_0) \leq \alpha$, d.w.z. wanneer $t_0 \leq t_{1-\alpha/2}(\nu = n - 1)$ of als $t_0 \geq t_{1-\alpha/2}(\nu = n - 1)$. De kritieke waarden van t kunnen in tabel C worden opgezocht.

Links éézijdige toetsing

- a. $H_0: \mu = \mu_0$ (of: $\mu \geq \mu_0$), $H_1: \mu < \mu_0$.
 H_0 kan niet worden verworpen, als $\bar{x}_0 \geq \mu_0$ is. Voor $\bar{x}_0 < \mu_0$ berekent men t_0 en:
 f. Verwerpt H_0 , wanneer $P_L(t_0) \leq \alpha$ is, d.i. voor $t_0 \leq t_\alpha(\nu = n - 1)$.

Rechts éézijdige toetsing

- a. $H_0: \mu = \mu_0$ (of: $\mu \leq \mu_0$), $H_1: \mu > \mu_0$.
 H_0 kan niet worden verworpen, als $\bar{x}_0 \leq \mu_0$ is. Voor $\bar{x}_0 > \mu_0$ berekent men t_0 en:
 f. Verwerpt H_0 , wanneer $P_R(t_0) \leq \alpha$ is, d.i. voor $t_0 \geq t_{1-\alpha}(\nu = n - 1)$.

Voorbeelden

15.1. Men laat een analyst vier onafhankelijke bepalingen verrichten van het hemoglobinegehalte van een standaard-bloed monster. Zijn uitkomsten zijn (in g%): 14,1, 14,4, 14,2 en 14,2. Het gehalte van het monster is 14,0 g%. Kan men concluderen, dat deze analyst bepalingen verricht, die systematisch van dit gehalte verschillen?

Verricht men onder identieke omstandigheden een groot aantal metingen aan een zelfde object, dan blijken deze gewoonlijk ongeveer normaal verdeeld te zijn. Men mag dus in dit geval wel aannemen, dat de verrichte bepalingen beschouwd kunnen worden als een (aselecte)

steekproef uit een (denkbeeldige) pseudo-normaal verdeelde populatie, zodat de zojuist behandelde t -toets kan worden toegepast:

- a. $H_0: \mu = 14,0$, $H_1: \mu \neq 14,0$. Er is a priori geen aanleiding om te onderstellen, dat de eventuele niet-toevallige afwijkingen een bepaalde richting zullen bezitten, zodat tweezijdig wordt getoetst.
- b. Kies $\alpha = 0,05$. Bij deze zeer kleine steekproefomvang komt een lagere drempelwaarde niet in aanmerking (om de toets een groter onderscheidingsvermogen te verschaffen, zou men zelfs kunnen overwegen, een hogere drempel, zoals 0,10, te kiezen).
- c. De toetsingsgrootte is $t = (\bar{x} - 14,0) \sqrt{4}/s$.
- d. Onder H_0 volgt t bij benadering een STUDENT-verdeling met $\nu = 3$.
- e. Bereken: $\bar{x}_0 = 14,225$ en $s_0 = 0,12583 = 0,126$ (controleer deze uitkomsten) en vervolgens:

$$t_0 = \frac{(14,225 - 14,000) \sqrt{4}}{0,126} = \frac{0,450}{0,126} = 3,57.$$

- f. Uit tabel C blijkt, dat bij $\nu = 3$ en $\alpha = 0,05$ ($\frac{1}{2}\alpha = 0,025$) de kritieke waarden bij tweezijdige toetsing zijn: $t_{0,025} = -3,182$ en $t_{0,975} = 3,182$. Daar t_0 de rechter kritieke waarde overschrijdt wordt H_0 verworpen (merk op, dat $t_{0,99} = 4,541$, zodat

$$0,02 < P_D(t_0) < 0,05).$$

De conclusie luidt dus, dat de analyst uitkomsten produceert, die weliswaar weinig, maar toch systematisch te hoog zijn.¹

15.2. Bij een bepaald chemisch proces is het van veel belang, dat de pH van een oplossing, die daarbij wordt gebruikt, niet hoger wordt dan 8,30. Men verricht op een bepaald moment zes onafhankelijke bepalingen aan een monster uit deze oplossing en vindt daarbij de volgende uitkomsten: 8,30, 8,35, 8,31, 8,32, 8,33, 8,31. Tot welke conclusie komt men, als $\alpha = 0,05$ wordt gekozen?

Onderstelt men, dat de gemeten grootte pseudo-normaal is verdeeld, dan kan als volgt worden getoetst:

- a. $H_0: \mu \leq 8,30$, $H_1: \mu > 8,30$. Er wordt rechts éézijdig getoetst, daar men uitsluitend geïnteresseerd is in een te hoge pH van de oplossing.
- b. $\alpha = 0,05$.
- c. De toetsingsgrootte is $t = (\bar{x} - 8,30) \sqrt{6}/s$.
- d. Onder H_0 volgt deze grootte bij benadering een STUDENT-verdeling met 5 vrijheidsgraden.

¹ In werkelijkheid zal men bij een dergelijk onderzoek een groter aantal waarnemingen verrichten. Wij hebben dit kleine aantal gekozen, om te tonen dat men soms ook met weinig waarnemingen tot een conclusie kan komen.

e. Bereken: $\bar{x}_0 = 8,32$ en $s_0 = 0,0178885$, waaruit volgt:

$$t_0 = \frac{(8,32 - 8,30) \sqrt{6}}{0,01789} = \frac{0,04899}{0,01789} = 2,74.$$

f. Tabel C levert op: $t_{0,95} (\nu = 5) = 2,015$ en daar t_0 groter is dan deze kritieke waarde, wordt H_0 verworpen. De conclusie luidt dus, dat de pH van de oplossing significant hoger is dan 8,30.

15.1.3. ONDERSCHIEDINGSVERMOGEN EN STEEKPROEFOMVANG

Beschouw een normaal verdeelde populatie met standaarddeviatie $\sigma = 10$, waarvan μ onbekend is. Onderstel, dat men op basis van een aselechte steekproef van 25 elementen de hypothese $H_0: \mu_0 = 50$ toetst met als alternatieve hypothese $H_1: \mu \neq 50$ en $\alpha = 0,05$ kiest. Onder H_0 volgt de grootheid \bar{x} een normale verdeling met gemiddelde $\mu_{\bar{x}} = \mu = 50$ en spreiding $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 10/\sqrt{25} = 2$ (zie 13.2). De kritieke zone van de toets, waarbij men tot het verwerpen van H_0 overgaat, bestaat dus uit waarden van \bar{x} die ten hoogste gelijk zijn aan $50 - 1,96 \times 2 = 46,08$ en tenminste gelijk aan $50 + 1,96 \times 2 = 53,92$.

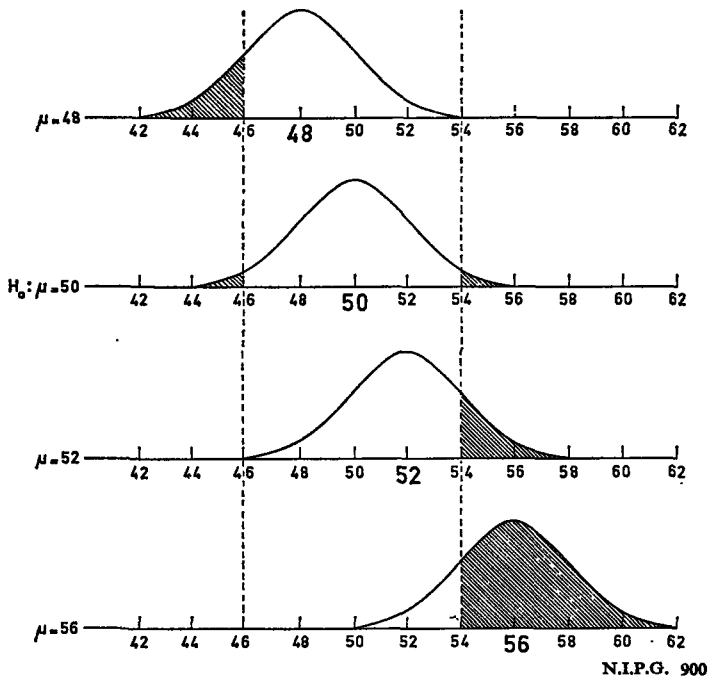
Hoe is het nu gesteld met het onderscheidingsvermogen van deze toets, d.w.z. hoe groot is de kans, dat \bar{x} een waarde in deze zone aanneemt voor verschillende alternatieve waarden van $\mu (\neq 50)$? Hiervan geeft figuur 15.1 een beeld voor drie alternatieve waarden, t.w. $\mu = 48$, 52 en 56. Deze figuur bevat — behalve de verdeling van \bar{x} onder H_0 — tevens de verdeling van \bar{x} voor deze alternatieven, met verschillende gemiddelden ($\mu_{\bar{x}} = \mu$), maar met dezelfde spreiding ($\sigma_{\bar{x}} = 2$). Het onderscheidingsvermogen van de toets voor deze drie waarden van μ is door het gearceerde oppervlak onder de krommen aangeduid. Dit onderscheidingsvermogen kan eenvoudig worden berekend, door voor elke verdeling te sommeren: $P_L(\bar{x} = 46,08) + P_R(\bar{x} = 53,92)$ en op deze wijze vindt men bv. voor de verdeling van \bar{x} onder $H_1: \mu = 52$ (met $\mu_{\bar{x}} = 52$ en $\sigma_{\bar{x}} = 2$):

$$P_L(\bar{x} = 46,08) = P_L \left[\underline{T} = \frac{46,08 - 52}{2} \right] = P_L(\underline{T} = -2,96) = 0,0015$$

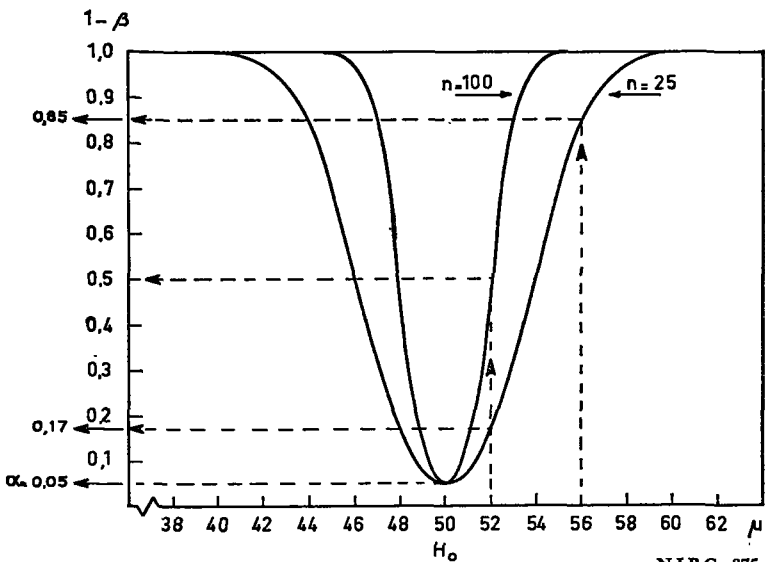
$$P_R(\bar{x} = 53,92) = P_R \left[\underline{T} = \frac{53,92 - 52}{2} \right] = P_R(\underline{T} = +0,96) = 0,1685$$

Onderscheidingsvermogen van $Z_D = 0,1700$

Als men het onderscheidingsvermogen van de toets voor verschillende alternatieve waarden van μ berekent, kan de werkingskromme van de toets voor $H_0: \mu = 50$ met $n = 25$ en $\alpha = 0,05$ worden getekend. Deze is in figuur 15.2 opgenomen en toont bv., dat de toets voor $\mu = 56$ een onderscheidingsvermogen van ongeveer 0,85 bezit en dat voor $\mu < 40$



Figuur 15.1. Verdeling van \bar{x} voor $n = 25, \sigma = 10$ en $\mu = 48, 50, 52$ en 56 . De kritieke zone Z_D is onder $H_0 : \mu = 50 : \bar{x} < 46,08$ en $\bar{x} > 53,92$.



Figuur 15.2. Werkingskrommen van de toets $H_0 : \mu = 50$, met $\sigma = 10$ en $\alpha = 0,05$ voor $n = 25$ en $n = 100$ (tweezijdige toetsing).

en $\mu > 60$ dit onderscheidingsvermogen ongeveer gelijk is aan 1. In de figuur is tevens de werkingskromme voor $n = 100$ getekend. Men ziet, dat bij deze grotere steekproefomvang bij dezelfde $\alpha = 0,05$ een veel groter onderscheidingsvermogen wordt verkregen. Zo is bv. voor $\mu = 52$ en $n = 25$ het onderscheidingsvermogen van de toets 0,17 en voor $n = 100$ is dit reeds 0,50.

Dit voorbeeld demonstreert nogmaals, dat het van groot belang is de steekproefomvang niet van allerlei min of meer toevallige factoren te laten afhangen, maar deze bewust te kiezen. Bij de in deze paragraaf behandelde toetsen is dit mogelijk, wanneer men kan aangeven:

1. Hoe ver de werkelijke waarde van μ van μ_0 moet afwijken, wil het verschil *praktisch* van voldoende betekenis zijn. Beschouw in dit verband voorbeeld 15.2. Hier is het heel goed denkbaar, dat pas een verhoging van de zuurgraad van 0,03 of meer van voldoende praktische betekenis is (bv. tot een ingrijpen in het proces moet leiden), zodat men tegenover $H_0: \mu = 8,30$ het alternatief $H_1: \mu_1 \geq 8,33$ stelt.

2. Hoe groot de populatiespreiding (ongeveer) is. Tenzij het terrein van onderzoek geheel nieuw is, kan men van deze spreiding (op grond van voorafgaand eigen onderzoek, publicaties van anderen) meestal wel een aanvaardbare schatting geven. Zo nodig kan men met twee schattingen werken, t.w. de minimale en de maximale spreiding die men verwacht. Bij voorbeeld 15.2 was bekend, dat de gebruikte meettechniek gewoonlijk een spreiding van ongeveer 0,02 vertoont.

3. Hoe groot het risico is dat men wil lopen op het ten onrechte verwerpen van H_0 . Dit is dus de keuze van α .

4. Hoe groot het risico is dat men wil lopen op het ten onrechte verwerpen van H_1 . Dit is β , de kans op een fout van de tweede soort, het complement van het onderscheidingsvermogen $1 - \beta$.

Onderstel nu, dat men de hypothese $H_0: \mu = \mu_0$ (of: $\mu \leq \mu_0$) toetst met als alternatief $H_1: \mu = \mu_1$ (of: $\mu \geq \mu_1$), zodat rechts éézijdige toetsing plaatsvindt. Onder H_0 volgt het steekproefgemiddelde \bar{x} (gebaseerd op n aselekt getrokken elementen) een normale verdeling met gemiddelde μ_0 en spreiding σ/\sqrt{n} . Onder H_1 volgt \bar{x} eveneens een normale verdeling, met gemiddelde μ_1 , maar met dezelfde spreiding σ/\sqrt{n} (zie fig. 15.3). Als men de waarde van \bar{x} , waarbij tot het verwerpen van H_0 wordt besloten \bar{x}_R noemt, is (zie de figuur)

$$\alpha = P[\bar{x} \geq \bar{x}_R | H_0], \quad \beta = P[\bar{x} < \bar{x}_R | H_1].$$

Hieruit volgt dat

$$T_{1-\alpha} = \frac{(\bar{x}_R - \mu_0) \sqrt{n}}{\sigma}, \quad T_{\beta} = \frac{(\bar{x}_R - \mu_1) \sqrt{n}}{\sigma}.$$

Dit zijn twee vergelijkingen met twee onbekenden, \bar{x}_R en n , waaruit n

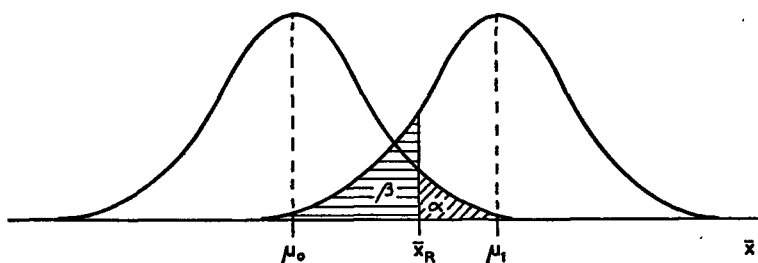
als volgt kan worden opgelost:

$$\begin{aligned} (\bar{x}_R - \mu_0) \sqrt{n} &= T_{1-\alpha} \sigma \\ (\bar{x}_R - \mu_1) \sqrt{n} &= T_{\beta} \sigma \\ \hline (\mu_1 - \mu_0) \sqrt{n} &= \sigma (T_{1-\alpha} - T_{\beta}) \\ &= \sigma (T_{1-\alpha} + T_{1-\beta}) \end{aligned}$$

$$(15.3) \quad n = \left[\frac{\sigma (T_{1-\alpha} + T_{1-\beta})}{\mu_1 - \mu_0} \right]^2.$$

Als $\mu_1 < \mu_0$ is (dus bij links éézijdige toetsing) verandert in deze formule alleen de noemer: deze wordt dan $\mu_0 - \mu_1$.

Toetst men tweezijdig, dan heeft men als alternatief $\mu_1 = \mu_0 - a$ en $\mu_1' = \mu_0 + a$. De noemer van (15.3) wordt dan $\pm a$, terwijl in de teller $T_{1-\alpha}$ door $T_{1-\frac{1}{2}\alpha}$ dient te worden vervangen.



N.I.P.G. 876

Figuur 15.3. Verdeling van \bar{x} onder H_0 en H_1 ($\mu_1 > \mu_0$).

Als n berekend is, kan ook de kritieke waarde \bar{x}_R worden vastgesteld. Bij rechts éézijdige toetsing vindt men dan:

$$\begin{aligned} (\bar{x}_R - \mu_0) \sqrt{n} &= T_{1-\alpha} \sigma \\ \bar{x}_R \sqrt{n} &= \mu_0 \sqrt{n} + T_{1-\alpha} \sigma \end{aligned}$$

$$(15.4.1) \quad \bar{x}_R = \mu_0 + T_{1-\alpha} \frac{\sigma}{\sqrt{n}}.$$

Op overeenkomstige wijze vindt men bij links éézijdige toetsing

$$(15.4.2) \quad \bar{x}_L = \mu_0 + T_{\alpha} \frac{\sigma}{\sqrt{n}}$$

en bij tweezijdige toetsing

$$(15.4.3) \quad \bar{x}_L = \mu_0 + T_{\frac{1}{2}\alpha} \frac{\sigma}{\sqrt{n}} \quad \text{en} \quad \bar{x}_R = \mu_0 + T_{1-\frac{1}{2}\alpha} \frac{\sigma}{\sqrt{n}}.$$

Overigens toont formule (15.3) duidelijk aan, dat men een grotere steekproef nodig heeft, naarmate: 1. Het verschil tussen μ_0 en μ_1 kleiner is, 2. σ groter is, en 3. α en/of β kleiner worden gekozen.

Als men σ niet kent, maar wel een minimale en maximale waarde kan aangeven, kan men voor elk dezer waarden n met (15.3) berekenen. Zodoende krijgt men een indruk van de steekproefomvang, die op zijn gunstigst, resp. op zijn ongunstigst noodzakelijk is en wordt de keuze van n in ieder geval geen volledige slag in de lucht.

Voorbeeld

15.3. Wij beschouwen nogmaals voorbeeld 15.2 en onderstellen, dat men n wil vaststellen voor $\mu_0 = 8,30$, $\mu_1 = 8,33$ en $\sigma = 0,02$. Men kiest $\alpha = 0,01$, zodat $T_{1-\alpha} = T_{0,99} = 2,326$ en $\beta = 0,05$, zodat $T_{1-\beta} = T_{0,95} = 1,645$. Met (15.3) vindt men dan, dat het aantal te verrichten bepalingen is:

$$n = \left[\frac{0,02 (2,326 + 1,645)}{8,33 - 8,30} \right]^2 = 7 .$$

Bepaalt men \bar{x}_R met (15.4.1), dan vindt men:

$$\bar{x}_R = 8,30 + 2,326 \frac{0,02}{\sqrt{7}} = 8,30 + 0,0176 = 8,3176 = 8,318.$$

Men zal dus tot het verwerpen van H_0 overgaan, als bij $n = 7$ wordt gevonden: $\bar{x}_0 \geq 8,318$.

Opmerking

Als de in (15.3) gebruikte waarde van σ een schatting is en als men bij de uitvoering van de toets gebruik maakt van de steekproefspread s_0 , dan kan men, tenzij men veiligheidshalve σ aan de hoge kant heeft geschat, beter als steekproefomvang kiezen:

$$n' = n [\text{volgens (15.3)}] + T^2_{1-\alpha/2}$$

(bij tweezijdige toetsing komt hierin weer $T^2_{1-\alpha/2}$ in de plaats van $T^2_{1-\alpha}$). Men toetst dan immers via (15.2) en niet volgens (15.1), zodat men met een STUDENT-verdeling (met $n - 1$ vrijheidsgraden) en niet met de normale verdeling te maken krijgt.

Beschouw bv. het geval, dat men 7 bepalingen verricht en vindt: $\bar{x}_0 = 8,32$ en $s_0 = 0,02$. Hoewel \bar{x}_0 groter is dan de in voorbeeld 15.3 bepaalde kritieke waarde ($\bar{x}_R = 8,318$) en s_0 gelijk is aan de geschatte waarde van σ , vindt men bij gebruik van (15.2):

$$t_{(6)} = \frac{8,32 - 8,30}{0,02} \sqrt{7} = 2,646,$$

zodat H_0 niet wordt verworpen, daar volgens tabel C geldt: $t_{0,99}(6) = 3,14$.

Zou men echter $n' = 7 + (2,33)^2/2 = 9,71 = 10$ bepalingen hebben verricht en leveren deze $\bar{x}_0 = 8,32$ en $s_0 = 0,02$ op, dan volgt uit (15.2):

$$t_{(9)} = \frac{8,32 - 8,30}{0,02} \sqrt{10} = 3,162,$$

terwijl $t_{0,99}(9) = 2,82$, zodat H_0 nu wél wordt verworpen.

In de praktijk zal men echter slechts zelden met deze verfijning bij het bepalen van de steekproefomvang te maken krijgen, omdat de betekenis van de 'correctie' sterk afneemt, naarmate n groter wordt. Overigens zal men dan de met (15.3) gevonden waarde van n gewoonlijk als een richtlijn en niet als een exact geldende waarde beschouwen.

15.1.4. OPGAVEN

15.1. Een aselechte steekproef van 16 elementen uit een normaal verdeelde populatie met $\sigma = 4$ levert op: $\bar{x}_0 = 10$. Toets de hypothese $\mu = 12$ (kies $\alpha = 0,05$).

15.2. Een aselechte steekproef van 36 elementen uit een normaal verdeelde populatie levert op $\bar{x}_0 = 23,5$ en $s_0 = 2,75$. Toets de hypothese $\mu = 25$ met $\alpha = 0,01$.

15.3. Men wil voor een normaal verdeelde populatie met $\sigma = 4$ de hypothese $\mu = 25$ toetsen met als alternatief $\mu = 23$. Bepaal de steekproefomvang voor de volgende waarden van α en β en bestudeer deze uitkomsten:

- | | | |
|----------------------------------|----------------------------------|-----------------------------------|
| a. $\alpha = \beta = 0,05$ | d. $\alpha = \beta = 0,01$ | g. $\alpha = \beta = 0,001$ |
| b. $\alpha = 0,05, \beta = 0,10$ | e. $\alpha = 0,01, \beta = 0,05$ | h. $\alpha = 0,001, \beta = 0,05$ |
| c. $\alpha = 0,05, \beta = 0,20$ | f. $\alpha = 0,01, \beta = 0,20$ | i. $\alpha = 0,001, \beta = 0,20$ |

15.4. Beschouw de volgende plannen voor tweezijdige toetsing van de hypothese $\mu = 50$: 1. $n = 40, \alpha = 0,01$, 2. $n = 100, \alpha = 0,01$, 3. $n = 200, \alpha = 0,05$, 4. $n = 500, \alpha = 0,02$, 5. $n = 500, \alpha = 0,05$. Welke proefopzet geeft een toets met: a. Het grootste onderscheidingsvermogen, b. Het kleinste onderscheidingsvermogen, c. De kleinste kans op een fout van de eerste soort?

15.5. Toets met $\alpha = 0,01$ de hypothese, dat $\mu = 85$ op basis van de volgende aselechte steekproef: 85,3, 85,7, 86,1, 85,1, 85,6, 85,9, 85,4, 84,9.

15.6. Een bepaald rattenras vertoont gedurende de eerste drie levensmaanden na het spenen een gemiddelde gewichtstoename van 65 gram. Zestien aselekt gekozen jonge ratten van dit ras worden in deze periode met een speciaal diëet gevoed en vertonen na drie maanden de volgende gewichtstoename in grammen: 65, 67, 68, 66, 70, 67, 69, 66, 63, 66, 65, 72, 67, 69, 64, 68. Kan men op grond van deze uitkomsten concluderen, dat het diëet een grotere gewichtstoename veroorzaakt dan de gebruikelijke voeding, als men $\alpha = 0,05$ kiest?
a. Pas een verdelingsvrije toets toe.
b. Onderstel normaliteit en pas de t -toets toe.

15.7. Onderstaand zijn de uterusgewichten gegeven van 20 ratten, die aselekt gekozen werden uit een groep van 500 ratten. Toets de hypothese, dat het gemiddelde uterusgewicht van deze groep 24 mg bedraagt. Kies $\alpha = 0,05$.

Uterusgewichten	9,	14,	15,	15,	16,	18,	18,	19,	19,	20,
(mg)	21,	22,	22,	24,	24,	26,	27,	29,	30,	32.

15.2. Betrouwbaarheidsgrenzen van het gemiddelde

Het gemiddelde $\mu_{\bar{x}}$ van de verdeling van \bar{x} is gelijk aan het populatie-gemiddelde μ en de spreiding van de verdeling van \bar{x} wordt kleiner, naarmate de steekproefomvang n toeneemt. \bar{x} is dus een zuivere en een betrouwbare schatter van μ . Voor normaal verdeelde populaties is \bar{x} tevens de meest doeltreffende schatter van μ .

15.2.1. SCHATTINGSINTERVAL VOOR μ , ALS σ BEKEND IS

Uit (13.4) volgt, dat voor een normaal verdeelde populatie met parameters μ en σ geldt

$$(15.5) \quad P [\mu + T_{\frac{1}{2}\alpha}\sigma_{\bar{x}} < \bar{x} < \mu + T_{1-\frac{1}{2}\alpha}\sigma_{\bar{x}}] = 1 - \alpha.$$

Beschouw nu de ongelijkheid

$$\mu + T_{\frac{1}{2}\alpha}\sigma_{\bar{x}} < \bar{x} < \mu + T_{1-\frac{1}{2}\alpha}\sigma_{\bar{x}}.$$

Deze kan als volgt worden herleid:

$$\begin{aligned} T_{\frac{1}{2}\alpha}\sigma_{\bar{x}} < \bar{x} - \mu < T_{1-\frac{1}{2}\alpha}\sigma_{\bar{x}}, \\ T_{\frac{1}{2}\alpha}\sigma_{\bar{x}} - \bar{x} < -\mu < T_{1-\frac{1}{2}\alpha}\sigma_{\bar{x}} - \bar{x}, \\ \bar{x} - T_{\frac{1}{2}\alpha}\sigma_{\bar{x}} > \mu > \bar{x} - T_{1-\frac{1}{2}\alpha}\sigma_{\bar{x}}, \end{aligned}$$

Maar $T_{\frac{1}{2}\alpha} = -T_{1-\frac{1}{2}\alpha}$, zodat

$$\bar{x} + T_{1-\frac{1}{2}\alpha}\sigma_{\bar{x}} > \mu > \bar{x} + T_{\frac{1}{2}\alpha}\sigma_{\bar{x}},$$

en men kan dus ook schrijven

$$(15.6) \quad P [\bar{x} + T_{\frac{1}{2}\alpha}\sigma_{\bar{x}} < \mu < \bar{x} + T_{1-\frac{1}{2}\alpha}\sigma_{\bar{x}}] = 1 - \alpha.$$

De ongelijkheid in (15.6) geeft men ook wel de vorm

$$(15.7) \quad \bar{x} \pm T_{\frac{1}{2}\alpha}\sigma_{\bar{x}}, \text{ of: } \bar{x} \pm T_{\frac{1}{2}\alpha} \frac{\sigma}{\sqrt{n}}.$$

Volgens een redenering, die analoog verloopt aan die in 14.2.1 geldt nu, dat men bij substitutie van één waargenomen gemiddelde \bar{x}_0 in (15.7) het *schattingsinterval*

$$(15.8) \quad \bar{x}_0 \pm T_{\frac{1}{2}\alpha} \frac{\sigma}{\sqrt{n}}$$

van μ verkrijgt met een betrouwbaarheidscoëfficiënt $1 - \alpha$ (of met een onbetrouwbaarheid α). Men noemt dit gewoonlijk het 100 $(1 - \alpha)$ % betrouwbaarheidsinterval van μ .

Wij zullen dit aan de hand van enkele figuren verduidelijken. Beschouw een normaal verdeelde populatie met gemiddelde μ en spreiding $\sigma = 3,47$. Trekt men hieruit aselekt een steekproef van 5 elementen, dan volgt het steekproefgemiddelde \bar{x} eveneens een normale verdeling en wel met gemiddelde $\mu_{\bar{x}} = \mu$ en spreiding $\sigma_{\bar{x}} = 3,47/\sqrt{5} = 1,55$. Uit (15.5) volgt bv. dat

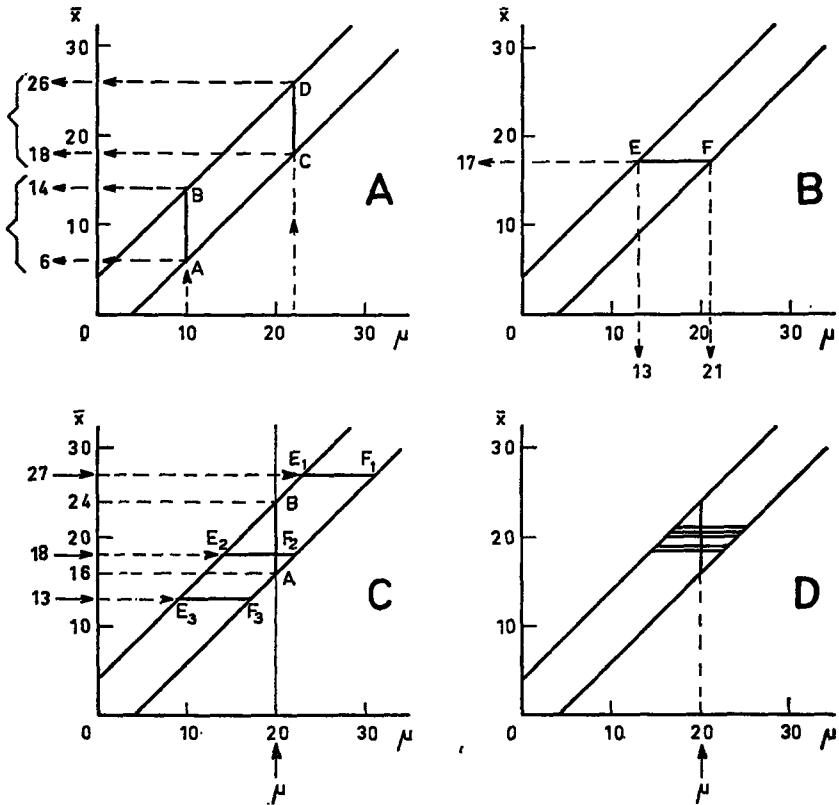
$$P [\mu - 2,58 \times 1,55 < \bar{x} < \mu + 2,58 \times 1,55] = 0,99$$

of

$$P [\mu - 4,0 < \bar{x} < \mu + 4,0] = 0,99$$

Wanneer $\mu = 10$ geldt dus: $P (6 < \bar{x} < 14) = 0,99$; is $\mu = 22$, dan geldt: $(18 < \bar{x} < 26) = 0,99$.

Deze grenzen van \bar{x} kan men voor elke waarde van μ uitrekenen en zij kunnen eenvoudig grafisch worden voorgesteld. Zie figuur 15.4 A waarin op de X-as de waarden van μ en op de Y-as de waarden van \bar{x} zijn afgezet. De verticale lijnstukken AB boven $\mu = 10$ en CD boven $\mu = 22$ vertegenwoordigen hierin de zojuist berekende intervallen, terwijl de twee hellende lijnen de boven- resp. de ondergrenzen verbinden van alle intervallen die men (bij alle mogelijke waarden van μ tussen 0 en 30) kan verkrijgen.



N.I.P.G. 877

Figuur 15.4. Betrouwbaarheidsgrenzen van het gemiddelde ($\sigma = 3,47$, $n = 5$, $1 - \alpha = 0,99$).

Beschouw vervolgens figuur 15.4B, waarin dezelfde grenslijnen zijn getekend, maar nu zonder de verticale lijnen AB en CD die aangeven hoe zij tot stand zijn gekomen. Onderstel dat een aselechte steekproef van 5 elementen $\bar{x}_0 = 17,0$ oplevert. Volgens (15.8) is dan het 99% be-

trouwbaarheidsinterval van μ : $17,0 \pm 2,58 \times 1,55 = 17 \pm 4$. Dit interval kan ook uit figuur 15.4B worden afgelezen, als men door het punt $\bar{x} = 17,0$ op de Y -as een lijn trekt evenwijdig aan de X -as en uit de snijpunten E en F van deze lijn met de grenslijnen loodlijnen op deze as neerlaat.

Eenvoudig is in te zien, dat men op deze wijze een 99% betrouwbaarheidsinterval verkrijgt. Onderstel bv., dat $\mu = 20$. De kans op een steekproefgemiddelde tussen 16 en 24 is dan 0,99 (zie figuur 15.4A). Voor elk steekproefgemiddelde tussen deze grenzen geldt echter, dat het horizontale lijnstuk EF in figuur 15.4B de loodlijn op $\mu = 20$ snijdt, d.w.z. dat de waarde $\mu = 20$ — terecht — binnen het 99% betrouwbaarheidsinterval ligt. Voor $\mu = 20$ is de kans op een steekproefgemiddelde boven 24 of beneden 16 gelijk aan 0,01. Bij zo'n gemiddelde behoort echter een lijnstuk EF , dat geheel rechts of geheel links van de loodlijn op $\mu = 20$ ligt, zodat deze waarde *niet* in het betrouwbaarheidsinterval ligt. Figuur 15.4C verduidelijkt dit: voor $\bar{x} = 27$ (> 24) snijdt E_1F_1 de loodlijn op $\mu = 20$ niet; voor $\bar{x} = 13$ (< 16) snijdt E_3F_3 deze loodlijn evenmin; voor $\bar{x} = 18$ ($16 < 18 < 24$) snijdt E_2F_2 deze loodlijn wèl.

In tabel 13.1 zijn de gemiddelden berekend van 5 aselechte steekproeven uit een pseudo-normaal verdeelde populatie met $\mu = 20$ en $\sigma = 3,47$, t.w.: 19,0, 20,4, 21,0, 18,4 en 20,0. In figuur 15.4D zijn de lijnstukken EF getekend, die bij deze waarden van \bar{x} behoren; zij snijden alle de loodlijn op $\mu = 20$, zodat de vijf 99% betrouwbaarheidsintervallen, die met deze steekproefgemiddelden corresponderen, alle $\mu = 20$ als mogelijke waarde van μ insluiten. Deze intervallen kunnen (bij benadering) uit de figuur worden afgelezen. Zij volgen ook uit (15.8) en zijn respectievelijk ¹

$$15,0 < \mu < 23,0, \quad 16,4 < \mu < 24,4, \quad 17,0 < \mu < 25,0, \\ 14,4 < \mu < 22,4 \quad \text{en} \quad 16,0 < \mu < 24,0.$$

Wat wij in het voorgaande over $\mu = 20$ hebben gezegd, geldt voor *elke* waarde van μ . De kans dat een betrouwbaarheidsinterval de werkelijke waarde van μ bevat is dus steeds 0,99, welke deze waarde ook is.

Het is duidelijk, dat bv. voor:

$$\begin{aligned} 1 - \alpha = 0,90, \quad \alpha = 0,10, \quad -T_{\frac{1}{2}\alpha} = T_{1-\frac{1}{2}\alpha} = 1,65, \\ 1 - \alpha = 0,95, \quad \alpha = 0,05, \quad -T_{\frac{1}{2}\alpha} = T_{1-\frac{1}{2}\alpha} = 1,96, \\ 1 - \alpha = 0,99, \quad \alpha = 0,01, \quad -T_{\frac{1}{2}\alpha} = T_{1-\frac{1}{2}\alpha} = 2,58, \\ 1 - \alpha = 0,999, \quad \alpha = 0,001, \quad -T_{\frac{1}{2}\alpha} = T_{1-\frac{1}{2}\alpha} = 3,29. \end{aligned}$$

¹ De hier verkregen intervallen zijn zeer breed, daar de steekproefomvang slechts klein is en $1 - \alpha = 0,99$ werd gekozen, om duidelijke figuren te verkrijgen.

15.2.2. SCHATTINGSINTERVAL VAN μ , ALS σ ONBEKEND IS

In dit geval kan op basis van het gemiddelde \bar{x}_0 van een aselechte steekproef van n elementen met spreiding s_0 een 100 $(1-\alpha)\%$ betrouwbaarheidsinterval van het populatiegemiddelde μ worden verkregen door te nemen

$$(15.9) \quad \bar{x}_0 + t_{\frac{1}{2}\alpha} \frac{s_0}{\sqrt{n}} < \mu < \bar{x}_0 + t_{1-\frac{1}{2}\alpha} \frac{s_0}{\sqrt{n}}, \quad \text{of: } \bar{x}_0 \pm t_{\frac{1}{2}\alpha} \frac{s_0}{\sqrt{n}},$$

waarbij de waarde van de volgens STUDENT verdeelde grootheid t met $n-1$ vrijheidsgraden in tabel C kan worden opgezocht.

In tabel 15.1 is een reeks voorbeelden van de berekening van betrouwbaarheidsgrenzen van μ gegeven, eerst voor $1-\alpha = 0,95$ (A) en vervolgens voor $1-\alpha = 0,99$ (B). Soms is σ hierbij bekend ondersteld, om het verschil tussen de uitkomsten van (15.8) en (15.9) duidelijk te laten uitkomen. Om het vergelijken van de verschillende verkregen intervallen te vergemakkelijken is steeds $\bar{x}_0 = 20$ genomen.

Tabel 15.1. Bepaling van betrouwbaarheidsgrenzen van μ
(voor alle steekproeven is $\bar{x}_0 = 20$)

A: $1-\alpha = 0,95, \alpha = 0,05$		$\frac{1}{2}\alpha = 0,025$		$T_{1-\frac{1}{2}\alpha} = T_{0,975}$		$t_{1-\frac{1}{2}\alpha} = t_{0,975}$		100 $(1-\alpha)\%$ betrouwbaarheidsinterval	
Steekproef nr	n	Spreiding		$\frac{\sigma}{\sqrt{n}}$ of $\frac{s_0}{\sqrt{n}}$	$T_{0,975}$	$t_{0,975}$	Berekening		Interval
		σ	s_0						
1	4	3		1,5	1,96		$20 \pm (1,96) (1,5)$		$17,1 < \mu < 22,9$
2	4		3	1,5		3,18	$20 \pm (3,18) (1,5)$		$15,2 < \mu < 24,8$
3	9	3		1,0	1,96		$20 \pm (1,96) (1,0)$		$18,0 < \mu < 22,0$
4	9		3	1,0		2,31	$20 \pm (2,31) (1,0)$		$17,7 < \mu < 22,3$
5	25	6		1,2	1,96		$20 \pm (1,96) (1,2)$		$17,6 < \mu < 22,4$
6	25		6	1,2		2,06	$20 \pm (2,06) (1,2)$		$17,5 < \mu < 22,5$
7	121	6		0,55	1,96		$20 \pm (1,96) (0,55)$		$18,9 < \mu < 21,1$
8	121		6	0,55		1,98	$20 \pm (1,98) (0,55)$		$18,9 < \mu < 21,1$
B: $1-\alpha = 0,99, \alpha = 0,01$		$\frac{1}{2}\alpha = 0,005$		$T_{1-\frac{1}{2}\alpha} = T_{0,995}$		$t_{1-\frac{1}{2}\alpha} = t_{0,995}$			
9	4	3		1,5	2,58		$20 \pm (2,58) (1,5)$		$16,1 < \mu < 23,9$
10	4		3	1,5		5,84	$20 \pm (5,84) (1,5)$		$11,2 < \mu < 28,8$
11	9	3		1,0	2,58		$20 \pm (2,58) (1,0)$		$17,4 < \mu < 22,6$
12	9		3	1,0		3,36	$20 \pm (3,36) (1,0)$		$16,6 < \mu < 23,4$
13	16	3		0,75	2,95		$20 \pm (2,95) (0,75)$		$17,8 < \mu < 22,2$
14	16		6	1,5	2,95		$20 \pm (2,95) (1,5)$		$15,6 < \mu < 24,4$
15	25	6		1,2	2,58		$20 \pm (2,58) (1,2)$		$16,9 < \mu < 23,1$
16	25		6	1,2		2,80	$20 \pm (2,80) (1,2)$		$16,6 < \mu < 23,4$
17	25	3		0,6		2,80	$20 \pm (2,80) (0,6)$		$18,3 < \mu < 21,7$

In de 'Statistische tabellen en nomogrammen' (76a) is een nomogram opgenomen, waaruit men de waarde van t/\sqrt{n} bij $\nu = n - 1$ vrijheidsgraden en verschillende waarden van α kan aflezen.

15.2.3. DE KEUZE VAN DE STEEKPROEFOMVANG

Wanneer de populatiespreiding bekend is (of ongeveer geschat kan worden) kan men vaststellen, hoe groot de steekproefomvang dient te zijn om een $100(1 - \alpha)\%$ betrouwbaarheidsinterval van een bepaalde breedte i te verkrijgen. Uit (15.8) volgt immers:

$$i = 2 T_{1-i\alpha} \frac{\sigma}{\sqrt{n}},$$

$$i^2 = 4 T_{1-i\alpha}^2 \frac{\sigma^2}{n}$$

zodat

$$(15.10) \quad n = \frac{4 T_{1-i\alpha}^2 \sigma^2}{i^2}$$

[vergelijk met formule (7.5)].

In de praktijk hangt de keuze van de intervalbreedte i af van de vraag hoe nauwkeurig men μ wil localiseren. De keuze van de betrouwbaarheidscoëfficiënt $1 - \alpha$ (die bepalend is voor de waarde van $T_{1-i\alpha}^2$) hangt af van het belang dat men erbij heeft, dat het te vinden betrouwbaarheidsinterval de werkelijke waarde van μ bevat. De variantie σ^2 van de populatie zal gewoonlijk niet precies bekend zijn, maar men zal deze wel ongeveer kunnen aangeven (desnoods enigszins aan de hoge kant, waardoor men een wat te grote n vindt).

Voorbeeld

15.4. De spreiding van het hemoglobinegehalte van jongens van 13 jaar is ongeveer 1 g%. Hoe groot moet een aselechte steekproef zijn, als men een 99% schattingsinterval van het gemiddelde hemoglobinegehalte van 13-jarige jongens wil verkrijgen met een intervalbreedte van 0,2 g%?

De hemoglobinegehalten blijken pseudo-normaal verdeeld te zijn, zodat formule (15.10) bij benadering geldt. In dit geval is: $\sigma = 1$, $i = 0,2$, $1 - \alpha = 0,99$, dus $\alpha = 0,01$ en $T_{i\alpha} = -2,58$, zodat

$$n = \frac{4 (2,58)^2 (1)^2}{(0,2)^2} = \frac{4 \times 6,6564 \times 1}{0,04} = 666.$$

Men zal dus een kleine 700 waarnemingen moeten verrichten. Wil men

met een 95% betrouwbaarheidsinterval genoeg nemen, dan vindt men

$$n = \frac{4 (1,96)^2 (1)^2}{(0,2)^2} = \frac{4 \times 3,8416 \times 1}{0,04} = 385 .$$

Is tevens een grotere intervalbreedte, zoals $i = 0,04$, acceptabel, dan is

$$n = \frac{4 (1,96)^2 (1)^2}{(0,4)^2} = \frac{4 \times 3,8416 \times 1}{0,16} = 96 .$$

Uit dit voorbeeld blijkt wederom zeer duidelijk, dat het bepalen van de noodzakelijke steekproefomvang (op grond van de eisen die men aan de uitkomsten wil stellen) een essentieel onderdeel van de planning van een onderzoek is. Kiest men de steekproefomvang arbitrair, dan kan dit later tot onaangename verrassingen leiden. Het gevonden betrouwbaarheidsinterval kan smaller zijn dan voor praktisch gebruik nodig is (in welk geval men een te kostbaar onderzoek heeft verricht), of — en dit komt meer voor — het is te breed om praktische bruikbaarheid te bezitten (in welk geval men het onderzoek beter achterwege had kunnen laten).

15.2.4. OPGAVEN

15.8. Bepaal het $100(1-\alpha)\%$ betrouwbaarheidsinterval van μ voor de onderstaande steekproeven:

St. pr.	\bar{x}_0	n	σ	s_0	α	St. pr.	\bar{x}_0	n	σ	s_0	α
<i>a</i>	15	4	1	—	0,05	<i>g</i>	40	60	—	6	0,05
<i>b</i>	18	9	4	—	0,01	<i>h</i>	40	60	6	—	0,05
<i>c</i>	18	9	—	3	0,01	<i>i</i>	40	60	6	8	0,01
<i>d</i>	40	16	10	12	0,001	<i>j</i>	40	120	—	12	0,05
<i>e</i>	40	36	—	12	0,001	<i>k</i>	40	120	—	18	0,05
<i>f</i>	40	36	—	6	0,01	<i>l</i>	40	120	—	18	0,01

15.9. Bepaal het 95% en het 99% betrouwbaarheidsinterval voor:

- a.* De gemiddelde gewichtstoename in opgave 15.5.
b. Het gemiddelde uterusgewicht in opgave 15.6.

15.10. Men wil het gemiddelde glucosegehalte (in mg/100 ml) vaststellen in door venapunctie verkregen bloed bij vrouwen van 20 t/m 24 jaar. Men wenst een schattingsinterval met $1-\alpha = 0,99$ en een breedte van ongeveer 2 mg/100 ml. Op grond van literatuurgegevens verwacht men dat $\sigma \cong 4$. Welke omvang dient de te trekken steekproef (ongeveer) te hebben? Welke onderstelling dient men te maken, om deze omvang te kunnen vaststellen?

X 15.3. Toepassing op het gemiddelde van verschillen tussen gepaarde waarnemingen

15.3.1. DE TOETS VAN STUDENT OP VERSCHILLEN ¹

In deel I hebben wij twee technieken behandeld die kunnen worden

¹ Zie ook 17.3.5.

toegepast als men beschikt over n gepaarde waarnemingen en *verdelingsvrij* de hypothese wil toetsen, dat de waargenomen verschillen op toeval berusten, nl. de tekentoets, waarbij uitsluitend gebruik wordt gemaakt van de tekens van de verschillen tussen de paren en de rang-tekentoets, waarbij tevens de rangorde naar grootte van deze verschillen in aanmerking wordt genomen.

Indien de waargenomen n verschillen beschouwd kunnen worden als een steekproef van n elementen uit een normaal verdeelde populatie (van verschillen) met gemiddelde μ_v en spreiding σ_v kan echter een toets met een groter onderscheidingsvermogen worden toegepast. Noemt men

$$(15.11) \quad \bar{v} = \frac{\sum v}{n}$$

en

$$(15.12) \quad s_v^2 = \frac{\sum v^2 - \frac{(\sum v)^2}{n}}{n - 1}$$

dan volgt volgens (13.11) de groottheid

$$(15.13) \quad t = \frac{(\bar{v} - \mu_v) \sqrt{n}}{s_v}$$

een STUDENT-verdeling met $n - 1$ vrijheidsgraden. Hypothesen betreffende μ_v kunnen dus worden getoetst met de in 15.1.2 beschreven methode. De hypothese dat de waargenomen verschillen toevallig zijn komt neer op $H_0: \mu_v = 0$. Onder deze hypothese volgt de groottheid

$$(15.14) \quad t = \frac{(\bar{v} - 0) \sqrt{n}}{s_v} = \frac{\bar{v} \sqrt{n}}{s_v}$$

een STUDENT-verdeling met $\nu = n - 1$.

Voorbeeld

15.5. Beschouw de volgende waargenomen verdeling van verschillen tussen gepaarde waarnemingen (verschillen tussen de hemoglobinebepalingen in g% van twee onderzoekers met dezelfde sicca-hemometer; de bepalingen zijn onmiddellijk na elkaar, doch onderling volledig onafhankelijk verricht):

v	0,6	0,5	0,4	0,3	0,2	0,1	0	-0,1	-0,2	-0,3	Totaal
f	1	1	2	3	5	6	3	2	1	1	25
fv	0,6	0,5	0,8	0,9	1,0	0,6	0	-0,2	-0,2	-0,3	3,7
fv^2	0,36	0,25	0,32	0,27	0,20	0,06	0	0,02	0,04	0,09	1,61

$$\bar{v}_0 = \frac{3,7}{25} = 0,148 \quad s_v^2 = \frac{1,61 - (3,7)^2/25}{24} = 0,044267 \quad s_v = 0,2104.$$

De toetsing verloopt als volgt:

a. $H_0: \mu_v = 0$, $H_1: \mu_v \neq 0$.

b. Kies $\alpha = 0,05$.

c. De toetsingsgrootheid is $t = \bar{v} \sqrt{n}/s_v$.

d. De waargenomen verdeling van de verschillen is zodanig, dat ondersteld kan worden dat zij uit een pseudo-normaal verdeelde populatie stammen. Onder H_0 volgt dan de grootheid t bij benadering een STUDENT-verdeling met $n - 1 = 24$ vrijheidsgraden.

e. Bereken: $t_0(24) = \frac{0,148 \sqrt{25}}{0,2104} = 3,52$.

f. De tweezijdige overschrijdingskans van t_0 is blijkens tabel C iets kleiner dan 0,002 [immers $t_{0,999}(\nu = 24) = 3,467$]. H_0 wordt dus verworpen en de conclusie luidt, dat er een systematisch verschil tussen de uitkomsten van de onderzoekers bestaat.

15.3.2. BETROUWBAARHEIDSGRENZEN VAN HET GEMIDDELDE VERSCHIL

Volgens de methode in 15.2.2 verkrijgt men een $100(1-\alpha)\%$ betrouwbaarheidsinterval van μ_v met

$$(15.15) \quad \bar{v}_0 + t_{\frac{1}{2}\alpha} \frac{s_v}{\sqrt{n}} < \mu_v < \bar{v}_0 + t_{1-\frac{1}{2}\alpha} \frac{s_v}{\sqrt{n}}, \text{ of: } \bar{v}_0 \pm t_{\frac{1}{2}\alpha} \frac{s_v}{\sqrt{n}},$$

waarin $t_{\frac{1}{2}\alpha} = -t_{1-\frac{1}{2}\alpha}$ met $\nu = n - 1$ in tabel C kan worden opgezocht.

Zo vindt men bv. als 95% betrouwbaarheidsinterval van het gemiddelde verschil bij de gegevens in voorbeeld 15.5

$$(t_{0,975} = -t_{0,025} = 2,064 \text{ met } \nu = 24):$$

$$0,148 - (2,064) \frac{0,2104}{\sqrt{25}} < \mu_v < 0,148 + (2,064) \frac{0,2104}{\sqrt{25}},$$

$$0,06 < \mu_v < 0,24.$$

15.3.3. OPGAVEN

15.11. Men beschikt over de volgende gegevens betreffende de prothrombinetijd (in sec.) bij 16 konijnen, $V = \text{vóór}$, en $N = \text{na verblijf van 4 dagen op een hoogte van 6000 m (in een caisson)}$:

Konijn	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
V	24	21	23	18	20	21	25	18	25	18	19	22	22	23	24	21
N	15	16	12	12	13	11	12	9	16	10	11	12	15	14	16	10
$V-N$	9	5	11	6	7	10	13	9	9	8	8	10	7	9	8	11

Het is duidelijk, dat de prothrombinetijd 'na' zeer significant lager is dan 'voor'. Hoe groot is de overschrijdingskans bij toepassing van: a) De tekentoets, b) De rangtekentoets, c) De t -toets op de verschillen? Geef aan: d) Welke

onderstelling men bij het gebruik van laatstgenoemde toets dient te maken en of deze naar Uw mening gerechtvaardigd is.

Men wenst een schattingsinterval met betrouwbaarheidscoëfficiënt 0,99 op te stellen: e) Hoe luidt dit interval?

15.12. Uit H. J. SLUITER, Cor pulmonale, Diss., Groningen, 1955.

Bij 33 emphyseempatiënten zonder decompensatio cordis werden de volgende 1 secundewaarden (%) van de vitale capaciteit waargenomen, V = voor en N = na 0,3 mg adrenaline s.c.

Pat.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
V	23	37	40	23	47	24	20	39	47	44	43	35	29	38	43	34	27
N	30	43	40	23	50	30	20	47	46	48	46	36	38	33	52	38	32
$N-V$	7	6	0	0	3	6	0	8	-1	4	3	1	9	-5	9	4	5
Pat.	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	
V	36	30	24	34	40	22	30	20	43	45	41	52	45	32	42	29	
N	35	29	29	38	35	29	29	20	45	45	44	55	55	35	45	29	
$N-V$	-1	-1	5	4	-5	7	-1	0	2	0	3	3	10	3	3	0	

Geef bij benadering het 95% betrouwbaarheidsinterval van μ_v .

15.13. Pas de t -toets toe op de gegevens in tabel 10.3 (blz. 222). Welke tweezijdige overschrijdingskans vindt U bij benadering? Vergelijk deze uitkomst met die van de rangtekentoets (0,125).

15.4. De nauwkeurigheid van waarnemingen

Een meestal onvermijdelijke eigenschap van waarnemingsmethoden is, dat zij bij herhaling niet steeds dezelfde uitkomsten geven. Slechts bij zeer eenvoudige tellingen of bij grove metingen doet dit verschijnsel zich niet voor. Het is hierdoor noodzakelijk, dat wij enige aandacht wijden aan het begrip 'nauwkeurigheid van waarnemingen' en het gebruik dat men daarbij van de statistiek kan maken. Wij zullen eerst de uitkomsten van metingen en vervolgens die van tellingen beschouwen.

15.4.1. METINGEN

Onderstel, dat men beschikt over een flinke hoeveelheid serum, waarvan het calciumgehalte bekend is en ξ g% bedraagt. Laat men nu een analyst A volgens een bepaalde methode één bepaling van dit gehalte verrichten, dan kan de uitkomst daarvan worden beschouwd als een aselechte trekking uit de populatie, bestaande uit alle bepalingen die deze analyst door middel van deze methode zou kunnen verrichten. Verricht A een aantal onderling onafhankelijke bepalingen aan hetzelfde serum, steeds volgens dezelfde methode, dan vormt de frequentieverdeling van de waargenomen uitkomsten een schatting van deze

populatieverdeling. Het is gebleken, dat zo'n waargenomen verdeling bij een uitgebreide categorie van waarnemingsmethoden (zoals lengtemetingen, wegingen, titraties, enz., ieder vaak herhaald aan hetzelfde object) bij benadering normaal is. Van de (fictieve) populatieverdeling van calciumbepalingen kan ook worden ondersteld, dat zij normaal is, zodat zij volledig kan worden gekarakteriseerd door de parameters μ (g%) en σ (g%). Deze verdeling typeert echter de *waarnemingsmethode*, gehanteerd door analyst *A*, en de parameters μ en σ representeren zodoende twee aspecten van de nauwkeurigheid van deze methode, die wij achtereenvolgens bespreken:

1. De zuiverheid van de waarnemingsmethode

De waarnemingsmethode wordt *zuiver* genoemd, als $\mu = \xi$ is. Is $\mu \neq \xi$, dan spreekt men van een *onzuivere* methode of van een methode met een *systematische fout* van de grootte $\mu - \xi$.

Een systematische fout kan bv. ontstaan door een telkens herhaalde verkeerde handeling bij de waarneming, door het gebruik van een pipet waarvan de opgegeven verdunningsfactor onjuist is, door het gebruik van een reagens met een foutieve concentratie, e.d. Men kan nagaan of μ gelijk is aan ξ door middel van een *ijking* van de meettechniek, d.i. door toepassing ervan op één of meer objecten waarvan ξ bekend is. Bij herhaalde meting moeten de (onderling onafhankelijke!) waarnemingen dan ongeveer symmetrisch om ξ gespreid liggen en op basis van het gemiddelde van n waarnemingen kan men de hypothese $\mu = \xi$ toetsen met behulp van de toets van STUDENT (zie 15.1 en voorbeeld 15.1). In het — ideale — geval, dat het verschil $\mu - \xi$ vrijwel constant blijkt te zijn, kan men door een nauwkeurig onderzoek de grootte van dit verschil bepalen, zodat bij de volgende waarnemingen een correctie kan worden aangebracht waardoor de waarnemingsmethode van een onzuivere in een zuivere overgaat.

In de praktijk doen zich t.a.v. de inrichting van een ijkingsoefening verschillende mogelijkheden (en moeilijkheden) voor, die men van geval tot geval dient te beschouwen en waarover wij slechts enkele algemene opmerkingen willen maken.

In de eerste plaats geldt, dat men er in de regel goed aan zal doen de verschillende hulpmiddelen, die bij de waarnemingsmethode worden gebruikt (zoals bv. pipetten, sicca-hemometers bij hemoglobinebepalingen, weegschalen of gewichtjes, enz.) afzonderlijk te (doen) ijken. Verder kan het nuttig zijn de ijking in eerste instantie voor verschillende waarnemers afzonderlijk uit te voeren. Een derde punt is, dat het soms nodig kan zijn de ijking over een zeker tijdsverloop uit te strekken teneinde na te gaan, of de tijd een invloed op de waarnemingen uitoefent en zo ja, welke (het kan bv. voorkomen, dat het verschil tussen μ en ξ een systematisch verloop in de tijd vertoont). En zonder te preten-deren dat daarmee alle mogelijkheden zijn uitgeput noemen wij tenslotte een onderzoek naar de invloed van het tijdsverloop tussen het verkrijgen van het te onderzoeken object (bv. het afnemen van bloed) en de waarneming (als het kan voorkomen, dat deze enkele uren of zelfs dagen later plaatsvindt), of tussen het verkrijgen en het voor onderzoek gereed maken (zie opgave 16.12).

2. De betrouwbaarheid van de waarnemingsmethode

De parameter σ karakteriseert de betrouwbaarheid (reproduceerbaarheid) van de waarnemingsmethode. Een grote waarde van σ houdt in, dat (vrij) grote afwijkingen van μ (en dus bij een zuivere methode van ξ) vaak voorkomen, een kleine, dat dit tot de uitzonderingen behoort. Is σ (bij benadering) bekend, bv. uit een groot aantal voorafgaande waarnemingen, en verricht men één waarneming x_1 , dan kan men bv. een 95% betrouwbaarheidsinterval van μ verkrijgen door te nemen

$$x_1 - 1,96\sigma < \mu < x_1 + 1,96\sigma .$$

Verricht men n onderling onafhankelijke waarnemingen en bezitten deze het gemiddelde \bar{x} , dan geldt volgens (15.8) als 95% betrouwbaarheidsinterval

$$\bar{x} - 1,96\sigma/\sqrt{n} < \mu < \bar{x} + 1,96\sigma/\sqrt{n} .$$

In de praktijk is σ vaak onbekend, zodat men in plaats van deze parameter een schatting s met ν vrijheidsgraden dient te gebruiken. Voor $\nu > 100$ kan men dan zonder bezwaar $\sigma = s$ stellen en voorgaande formule gebruiken, maar voor een kleiner aantal vrijheidsgraden dient men gebruik te maken van het betrouwbaarheidsinterval, dat door (15.9) gegeven wordt.

Een veel voorkomende situatie is, dat men beschikt over k objecten en een analyst per object twee onderling onafhankelijke bepalingen (duplo's) laat verrichten. Elke duplo kan dan beschouwd worden als een aselechte steekproef van 2 elementen uit een (fictieve) populatie met gemiddelde μ_i en spreiding σ_i en men kan dan een zuivere schatting van de populatievariantie verkrijgen met

$$(15.16) \quad s_i^2 = \frac{d_i^2}{2} ,$$

waarin d_i het verschil van het tweetal bepalingen is. Onderstelt men, dat de k populaties ($i = 1, 2, \dots, k$) dezelfde spreiding bezitten, zodat $\sigma_1 = \sigma_2 = \dots = \sigma_k = \sigma$, dan kan men met (13.29) een zuivere schatting van σ^2 verkrijgen:

$$(15.17) \quad s^2 = \frac{\frac{(2-1)d_1^2}{2} + \frac{(2-1)d_2^2}{2} + \dots + \frac{(2-1)d_k^2}{2}}{k(2-1)} = \frac{\sum_{i=1}^k d_i^2}{2k} ,$$

d.i. dus door de som van de kwadraten van de verschillen tussen de duplo's te delen door $2k =$ het totale aantal waarnemingen. Deze schatting van σ^2 bezit k vrijheidsgraden.

Voorbeeld

15.6. Uit Chr. L. RÜMKE, syllabus 'Medische statistiek', 1959. De onderstaande tweetallen bepalingen betreffen het bloedsuikergehalte

(in mg per 100 ml) van twee kort na elkaar afgenomen bloedmonsters. Wij onderstellen, dat deze bepalingen alle door één analytist zijn verricht.

Duplo nr.	Waarnemingen		Verschil $d = a - b$	d^2
	a	b		
1	99	92	7	49
2	101	101	0	0
3	89	84	5	25
4	92	98	-6	36
5	66	75	-9	81
6	108	101	7	49
7	103	108	-5	25
8	98	94	4	16
9	85	87	-2	4
10	99	94	5	25
11	94	89	5	25
12	71	77	-6	36
13	80	73	7	49
14	77	77	0	0
15	87	78	9	81
16	96	96	0	0
17	89	85	4	16
18	84	84	0	0
19	85	84	1	1
20	80	80	0	0
21	94	91	3	9
22	96	91	5	25
23	103	106	-3	9
24	96	101	-5	25
25	103	104	-1	1

$\Sigma d^2 = 587$

Volgens (15.17) is: $s^2 = 587/50 = 11,74$, zodat $s = \sqrt{11,74} = 3,4$. Nu is bv. het gemiddelde van het 6-de tweetal waarnemingen 104,5, zodat men met (15.9) als 95% betrouwbaarheidsinterval van μ_6 vindt

$$104,5 - 2,06 \frac{3,4}{\sqrt{2}} < \mu_6 < 104,5 + 2,06 \frac{3,4}{\sqrt{2}},$$

$$99,5 < \mu_6 < 109,5 .$$

Indien de methode zuiver is, levert dit interval dus een schatting met een betrouwbaarheid van 0,95 van ξ_6 , het bloedsuikergehalte in het zesde bloedmonster. In deze schatting zijn alle foutenbronnen verwerkt, vanaf die welke kunnen optreden bij het afnemen van het bloed tot en met die welke ontstaan bij het aflezen op de buret of colorimeter. Als men mag aannemen of vastgesteld heeft, dat de

betrouwbaarheid van analyst A (ongeveer) constant is, kan men de voor s gevonden waarde ook voor het bepalen van de betrouwbaarheid van latere bepalingen gebruiken.

Onderstel vervolgens, dat de eerste 15 duplo's door analyst A en de volgende 10 door analyst B zijn verricht. Men vindt dan met (15.17)

voor analyst A :

$$s_A^2 = 501/30 = 16,7; \quad s_A = 4,1$$

voor analyst B :

$$s_B^2 = 86/20 = 4,3; \quad s_B = 2,1.$$

Deze varianties verschillen significant bij een 5% drempel, daar $F_0 = 16,7/4,3 = 3,88$ en $F_{0,975}(\nu_1 = 15, \nu_2 = 10) = 3,52$. De conclusie luidt dus, dat de twee analysten bepalingen met een verschillende betrouwbaarheid leveren en men kan dan beter bij elk stel waarnemingen met de betrouwbaarheid van de betrokken analyst werken. Voor het 6-de tweetal (analyst A) vindt men dan als 95% betrouwbaarheidsinterval ($s_A = 4,1$, $\nu = 15$, $t_{0,975} = 2,131$):

$$104,5 - 2,131 \frac{4,1}{\sqrt{2}} < \mu_6 < 104,5 + 2,131 \frac{4,1}{\sqrt{2}}, \\ 98,3 < \mu_6 < 110,7.$$

Voor het 23-ste tweetal (analyst B), waarvan het gemiddelde gehalte eveneens 104,5 is, vindt men als 95% betrouwbaarheidsinterval ($s_B = 2,1$, $\nu = 10$, $t_{0,975} = 2,228$):

$$104,5 - 2,228 \frac{2,1}{\sqrt{2}} < \mu_{23} < 104,5 + 2,228 \frac{2,1}{\sqrt{2}}, \\ 101,2 < \mu_{23} < 107,8.$$

Onderstelt men echter, dat de oneven duplo's door A en de even door B zijn verricht, dan vindt men dat

$s_A^2 = 375/26 = 14,42$, $s_B^2 = 212/24 = 8,83$, $F_0 = 14,42/8,83 = 1,63$, dus een duidelijk insignificant verschil. Hier kan men dan ook de betrouwbaarheid van A en B gelijk stellen.

15.4.2. TELLINGEN

Men pleegt het aantal leucocyten, erythrocyten, eosinophile granulocyten, e.d. in het bloed te bepalen door middel van een steekproef. Een bepaalde hoeveelheid bloed wordt in een pipet opgezogen en na verdunning wordt het aantal cellen met behulp van een telkamer onder een microscoop geteld.

RÜMKE (190) merkt omtrent deze steekproef het volgende op: 'De steekproef is wel zeer klein ten opzichte van hetgeen men wil kennen: de ongeveer 0,025 ml bloed, die in een leucocytenpipet wordt opgezogen, vertegenwoordigt bij benadering 1/200.000 van de hoeveelheid bloed in de circulatie. In slechts een deel van het tienvoudig verdunde monster wordt daarna het aantal cellen

geteld! Reeds op grond van deze overwegingen is het duidelijk, dat de uitkomsten van tellingen slechts een zeer oppervlakkige indruk kunnen geven van het aantal cellen, dat zich in het bloed bevindt. Ook kan men zich afvragen, in hoeverre de uitdrukking 'aantal cellen per mm^3 bloed' inderdaad een bestaande toestand weergeeft. Het is immers mogelijk dat de cellen niet homogeen in het circulerende bloed zijn verdeeld. Bovendien heeft men bij de beoordeling van de uitkomsten aan het ziekbed rekening te houden met de variaties in aantal bij de verschillende mensen en de physiologische schommeling van het aantal cellen in de loop van de dag, het dagrythme.'

Als wij aannemen, dat een bepaald bloedmonster een juist beeld geeft van de situatie in het organisme van de bloedgever, dan kunnen zich bij het *tellen* van de cellen weer verschillende foutenbronnen voordoen, die in vier soorten kunnen worden onderscheiden: 1. De fout van de pipet. De verdunningsfactoren van de in de handel verkrijgbare leucocyten- en erythrocytenpipetten wijken soms vrij aanzienlijk van 10, resp. 100 af. Deze fout kan worden voorkomen, door de te gebruiken pipetten nauwkeurig te (laten) ijken. 2. De fout van de telkamer door het optreden van afwijkingen van het onderstelde volume, bv. tengevolge van variaties in de diepte en het gebruik van dekglazen, die niet geheel vlak zijn. Wij nemen aan, dat men ook deze fout door een nauwkeurige ijking vrijwel geheel kan vermijden. 3. De fouten van de onderzoeker. Tot op zekere hoogte zal men deze kunnen ontlopen door zeer precies tewerk te gaan, bv. door de pipet nauwkeurig tot beide merktekens te vullen, het ontstaan van stolsels te voorkomen, zorgvuldig te tellen (onzorgvuldige telling geeft in het algemeen een te laag aantal!), enz. Onvermijdelijk zijn echter moeilijkheden bij de beslissing of een bepaalde cel binnen het te tellen gebied ligt (deze zijn des te erger, naarmate er minder cellen zijn), dan wel of een bepaald vormsel tot de te tellen celsoort behoort. Door middel van proeven zal men o.i. in ieder laboratorium een indruk kunnen krijgen omtrent de ernst en omvang van deze soort fouten.¹ 4. De zg. verdelingsfout, d.w.z. de variatie die ontstaat, doordat de cellen niet gelijkmatig over de vele vierkantjes van het telnet verspreid liggen. Deze onregelmatigheid is onvermijdelijk en vormt de voornaamste bron van onnauwkeurigheid van de uitkomsten van de telling. Bij deze variatiebron kan men echter weer de statistiek te hulp roepen, zodat wij er onze verdere bespreking aan wijden.

Onderstel, dat zich in een bloedmonster C leucocyten per mm^3 bevinden. Men maakt een tienvoudige verdunning en laat een persoon A met een telkamer van $0,1 \text{ mm}^3$ (= 25 van de gebruikelijke vierkantjes) een leucocytentelling verrichten. Deze uitkomst kan dan beschouwd worden als een aselechte trekking uit de populatie van alle mogelijke tellingen die A zou kunnen verrichten. Verricht A een aantal tellingen,

¹ Het valt ons echter op, dat deze proeven slechts zelden schijnen te worden gedaan — of dat hun uitkomsten alleen sporadisch worden gepubliceerd.

dan vormt de frequentieverdeling daarvan een schatting van deze populatieverdeling. Het blijkt (en is ook op theoretische gronden te verklaren), dat deze waargenomen verdeling ongeveer volgens Poisson verdeeld is (zie voor een voorbeeld tabel 15.2), zodat de populatieverdeling der tellingen als een Poisson-verdeling kan worden beschouwd.

Tabel 15.2. Het aantal waargenomen eosinophiele granulocyten (x) in 144 vakjes van een telkamer, f = waargenomen frequenties, e = frequenties volgens best-passende Poissonverdeling met $\mu = \bar{x} = 4,8472$

x	0	1	2	3	4	5	6	7	8	9	10	11	12	Totaal
f	2	7	9	24	24	29	16	15	8	6	3	1	0	144
e	1,1	5,5	13,3	21,5	26,0	25,2	20,4	14,1	8,5	4,6	2,2	1,0	0,6	144

Na groepering van de waarden 0 en 1, resp. 9 en hoger, vindt men χ^2_0 ($\nu = 7$) = 4,89, $0,5 < P < 0,7$, dus een zeer bevredigende aanpassing.

Deze kan volledig worden gekarakteriseerd door één parameter, μ , d.w.z. heeft als gemiddelde μ en als spreiding $\sqrt{\mu}$. Als het aantal cellen echter betrekkelijk groot is, d.w.z. als $\mu \geq 15$ is, kan deze Poisson-verdeling reeds goed worden benaderd door een normale verdeling met gemiddelde μ en spreiding $\sqrt{\mu}$.¹

De verdeling typeert de tellingsmethode, gehanteerd door A , en als men mag aannemen dat geen systematische fout optreedt, dan is dus $10 \times 10 \times \mu = C$. Heeft men nu bij een tienvoudige verdunning in een vakje van $0,1 \text{ mm}^3$ bv. 81 leucocyten geteld dan kan men, bij benadering van de Poissonverdeling door de normale verdeling, als 95% betrouwbaarheidsinterval van het werkelijke aantal leucocyten in het in de telkamer onderzochte volume nemen (i.p.v. 1,96 is met 2 vermenigvuldigd)

$$81 - 2 \sqrt{81} < \mu < 81 + 2 \sqrt{81},$$

$$63 < \mu < 99.$$

Het aantal leucocyten per mm^3 van het bloedmonster kan dan met een betrouwbaarheid van 0,95 worden geschat op een tussen 6300 en 9900 liggend aantal. Men bedenke hierbij echter, dat bij deze schatting slechts met de verdelingsfout rekening is gehouden en omdat andere bronnen van onnauwkeurigheid wel nooit geheel vermeden zullen kunnen worden, liggen de intervalgrenzen in feite verder uiteen.

Telt men in hetzelfde monster in de telkamer bv. het aantal cellen in $0,4 \text{ mm}^3$ (dus 4×25 van de gebruikelijke vierkantjes) en vindt men daarin 324 leucocyten (dus gemiddeld 81 leucocyten per $0,1 \text{ mm}^3$), dan

¹ Is dit niet het geval, dan dient men gebruik te maken van tabellen of nomogrammen van Poisson-verdelingen. Wij laten deze situatie buiten beschouwing.

vindt men voor de grenzen van het 95% betrouwbaarheidsinterval

$$324 - 2\sqrt{324} < \mu < 324 + 2\sqrt{324},$$

$$288 < \mu < 360.$$

Deze uitkomst dient door 4 te worden gedeeld om het overeenkomstige interval voor 0,1 mm³ van de telkamer te verkrijgen, zodat hiervoor als grenzen 72 en 90 worden gevonden. Het aantal leucocyten per mm³ van het bloed wordt dan dus (afgezien van andere foutenbronnen) met een betrouwbaarheid van 0,95 geschat op een aantal tussen 7200 en 9000 en de intervalbreedte blijkt dus te zijn gehalveerd door een viermaal zo groot volume te tellen.

Uit het voorgaande blijkt, dat de grootte van de verdelingsfout uitsluitend afhankelijk is van het totale aantal getelde cellen en dus slechts indirect van het aantal vierkantjes binnen de telkamer waarin men de cellen telt. De praktische betekenis van deze bevinding is, dat men van bloed met bv. 4000 leucocyten per mm³ het aantal cellen in tweemaal zoveel vierkantjes moet tellen, als men een even grote betrouwbaarheid wil bereiken als bij bloed waarin zich 8000 leucocyten per mm³ bevinden.

In het reeds eerder geciteerde artikel van RÜMKE (190), waarin hij dieper op deze materie ingaat en dat wij een ieder ter lezing aanbevelen, wordt dan ook o.m. het volgende geconcludeerd: 'De uitkomsten van tellingen van cellen kunnen waardevolle gegevens aan het ziekbed zijn. Toch is het wel duidelijk, dat het niet juist is grote waarde te hechten aan het vinden van een betrekkelijk geringe verhoging of verlaging van het aantal bloedcellen bij een patiënt, wanneer de methode, waarmee deze uitkomsten zijn verkregen, zo onnauwkeurig is.... De onvermijdelijke verdelingsfout kan men zo klein mogelijk houden door in de telkamer de cellen in een zo groot mogelijk aantal vierkanten te tellen. Bovendien behoort dan voor de telling een gelijk aantal proximale en distale vierkanten ten opzichte van de plaats, waar men de telkamer uit de pipet vulde, te worden gebruikt. Aan deze voorwaarde is ook voldaan indien bij de gebruikelijke plaatsing van de telkamer onder de microscoop niet 'horizontaal' maar 'verticaal' wordt gewerkt.'

Een tweede probleem, waarop RÜMKE (193) in een ander artikel ingaat, en dat wij hier niet geheel onbesproken willen laten, betreft de differentiatie van cellen in een uitstrijkpreparaat met de bedoeling vast te stellen, welk percentage van het beschouwde type cellen bepaalde kenmerken bezit. Als bij differentiatie van een steekproef, bv. bestaande uit n witte bloedcellen, het aantal eosinophile granulocyten x en dus de fractie van deze cellen $p = x/n$ bedraagt, dan is deze fractie een schatting van de fractie P in de populatie (het bloed). In hoofdstuk 7 hebben wij reeds verschillende manieren behandeld, waarmee men op basis van de waargenomen fractie p een 100 (1 - α)% betrouwbaarheidsinterval van P kan verkrijgen. Voor niet te kleine waarden van p en n kan men bv. gebruik maken van figuur 7.3 op blz. 161. In de

bijlagen van dit deel hebben wij echter nog een tweetal tabellen opgenomen, waarmee dit eenvoudiger en voor zeer kleine waarden van p ook nauwkeuriger kan geschieden.

Tabel W bevat de 95% en 99% betrouwbaarheidsintervallen van het *percentage* elementen met een zeker kenmerk A in de populatie, die gelden indien x ($= 1, 2, \dots, 10$) elementen van de n ($= 100, 150, 200, 300, 500$ of 1000) elementen van de steekproef dit kenmerk blijken te bezitten. Tabel X geeft 95% betrouwbaarheidsintervallen van het percentage met kenmerk A in de populatie, indien a procent van de n ($= 100, 200, 500$ of 1000) elementen in de steekproef dit kenmerk dragen. Uiteraard kunnen deze tabellen worden gebruikt voor de uitkomsten van ieder onderzoek, waarbij is nagegaan of onder de gegeven aantallen elementen in een steekproef een aantal elementen een bepaald kenmerk bezit.

Voorbeelden

15.7 Bij differentiatie van 200 witte bloedcellen werden 8 (4 pct) basophiele granulocyten aangetroffen. Uit tabel W-2 blijkt, dat het percentage basophielen in het bloed behoudens een onbetrouwbaarheid van 1% tussen 1,3 en 9 pct ligt.

15.8. Onder 500 erythrocyten werden 4 reticulocyten gevonden. Het 95% betrouwbaarheidsinterval van het percentage reticulocyten in het bloed kan blijkens tabel W-1 op 0,22 — 2,0 pct worden gesteld.

15.9. Onder 100 witte bloedcellen treft men 29% lymfocyten aan. Uit tabel X leest men bij benadering als 95% betrouwbaarheidsinterval van het percentage lymfocyten in het bloed 21 tot 40 pct af. Voor dit geval leest men in figuur 7.3 ongeveer de grenswaarden 20,5 en 38 pct af.

Wij besluiten deze korte uiteenzetting weer met enkele opmerkingen van РѹМКЪ: 'De differentiatie van het witte bloedbeeld wordt in de praktijk nogal eens op deze wijze uitgevoerd, dat men de beslissing, of het noodzakelijk is (bijvoorbeeld) 200 cellen te differentiëren, afhankelijk stelt van de uitkomsten bij de eerste 100 cellen. Met nadruk moet erop worden gewezen, dat het onmogelijk is, enige uitspraak te doen over de betrouwbaarheidsintervallen van de uitkomsten, die men op deze manier door differentiatie van 200 cellen verkrijgt. Bij deze handelwijze worden na differentiatie van 100 cellen de uitkomsten verwaarloosd van de uitstrijkpreparaten die hetzij een bijzonder 'abnormale', hetzij een bijzonder 'normale' samenstelling schijnen te hebben. Door deze verwaarlozing wordt het onmogelijk, de regels der kansrekening toe te passen op de differentiaties van 200 cellen. . . . In vele leerboeken worden zg. normale waarden voor de procentuele samenstelling van het bloed opgegeven. Hierbij wordt nimmer vermeld, hoeveel cellen er bij elke onderzochte werden gedifferentieerd om tot vaststelling van de norm te geraken. Het is duidelijk, dat een dergelijke opgave onvolledig is. Het is denkbaar, dat de marges voor elke 'normale waarde' kleiner zullen blijken te zijn, als de percentages door differentiatie van een groter aantal cellen worden vastgesteld.'

15.5

GEMIDDELDE VAN EEN POPULATIE

15.4.3. OPGAVEN

15.14. De onderstaande waarnemingen betreffen duplo's, die door één analytist werden verkregen bij het uitvoeren van een analyse volgens twee verschillende methoden op 10 monsters.

- a. Geef schattingen van de betrouwbaarheid van deze methoden en vergelijk deze.
 b. Wijzen de uitkomsten op een systematisch verschil in betrouwbaarheid tussen de twee methoden? Kies $\alpha = 0,05$.

Monster	Methode			
	I		II	
1	13,13	13,12	13,06	13,07
2	13,38	13,48	13,44	13,42
3	19,97	19,89	19,98	19,98
4	14,77	14,79	14,72	14,73
5	13,11	13,19	13,14	13,17
6	17,88	17,92	17,93	17,92
7	17,69	17,63	17,65	17,65
8	18,17	18,15	18,17	18,14
9	16,89	16,92	16,89	16,88
10	15,70	15,70	15,70	15,71

15.15. Geef (bij benadering) van de onderstaande uitkomsten van tellingen van leucocyten (geteld bij tienvoudige verdunning met een telkamer van $0,2 \text{ mm}^3$) het 95% betrouwbaarheidsinterval van het aantal leucocyten in het bloed:
 a. 64 b. 100 c. 144 d. 196.

15.16. Geef 95% betrouwbaarheidsgrenzen op grond van de volgende uitkomsten:

- a. 3 eosinophile granulocyten onder 300 witte bloedcellen.
 b. 10 eosinophile granulocyten onder 1000 witte bloedcellen.
 c. 26% lymfocyten onder 500 witte bloedcellen.
 d. 15% cellen met een bepaald kenmerk onder 200 cellen in een beenmergpunctuaat.

15.5. Het signaleren van uitbijters

15.5.1. INLEIDING

Bij de analyse van waarnemingsuitkomsten kan zich de vraag voordoen, of één bepaalde waarneming als uitbijter dient te worden beschouwd. Er zijn verschillende toetsen voorgesteld, waarmee kan worden onderzocht of zo'n verdachte waarneming al dan niet door toeval kan zijn ontstaan en in het volgende punt van deze paragraaf zullen wij er één van bespreken. Eerst willen wij nog enkele opmerkingen maken omtrent de gedragslijn die men moet volgen, als zo'n toets uitwijst dat de verdenking gegrond is.

Wij menen, dat men dan in de eerste plaats moet onderzoeken, of de uitbijter door een grove waarnemings- of registratiefout kan zijn ver-

oorzaakt. Blijkt dit het geval te zijn en kan men een nieuwe waarneming verrichten, dan is de moeilijkheid opgelost; de gesignaleerde uitbijter kan er dan wellicht toe leiden, dat de meettechniek of wijze van registreren wordt gewijzigd of verbeterd, zodat dergelijke fouten in de toekomst worden vermeden. Kan de fout niet worden hersteld, dan zal het gewoonlijk verantwoord zijn de uitbijter te verwijderen.

De zaak ligt echter anders, als er geen fout aanwijsbaar blijkt te zijn. Ons inziens zal men dan uitsluitend tot het verwijderen van de uitbijter mogen overgaan, als de waarneming bv. op theoretische gronden onaanvaardbaar moet worden geacht en anders zal men de uitbijter nolens volens dienen te accepteren.

15.5.2. DE UITBIJTERTOETS VAN DIXON

Onderstel, dat men beschikt over de in opklimmende volgorde gerangschikte waarnemingen x_1, x_2, \dots, x_n en de hypothese H_0 wil toetsen, dat deze waarnemingen afkomstig zijn uit dezelfde normaal verdeelde populatie met als alternatieve hypothese dat de grootste waarneming x_n stamt uit een populatie met een hoger gemiddelde dan de overige $n - 1$ waarnemingen (of met als alternatieve hypothese, dat de kleinste waarneming x_1 afkomstig is uit een populatie met een lager gemiddelde dan de resterende $n - 1$ waarnemingen). Men kan dan voor $n \leq 10$ gebruik maken van de door DIXON (144, 145) aangegeven toetsingsgrootheid¹

$$(15.18.1) \quad r_{10} = \frac{x_n - x_{n-1}}{x_n - x_1},$$

resp.

$$(15.18.2) \quad r_{10} = \frac{x_2 - x_1}{x_n - x_1}.$$

Voor $11 \leq n \leq 30$ kan men beter gebruik maken van de grootheid

$$(15.19.1) \quad r_{22} = \frac{x_n - x_{n-2}}{x_n - x_3},$$

resp.

$$(15.19.2) \quad r_{22} = \frac{x_3 - x_1}{x_n - x_3}.$$

In tabel Y zijn voor $3 \leq n \leq 30$ de rechter kritieke waarden r_R opgenomen voor $\alpha = 0,10, 0,05$ en $0,01$. Bij toetsing van de hypothese H_0 met onbetrouwbaarheid α kan men dus tot verwerpen overgaan, als $r_0 \geq r_R(\alpha, n)$ is.

¹ De letter r is van oudsher het symbool voor de correlatiecoëfficiënt. Het gebruik van één symbool in meerdere betekenissen is echter niet meer te vermijden. De indices dienen om de verschillende door DIXON gegeven grootheden (in totaal 6) van elkaar te onderscheiden.

Voorbeelden

15.10. Gegeven is de volgende (reeds naar grootte gerangschikte) reeks waarnemingen: 1,4, 1,7, 1,7, 2,2, 2,3, 2,5, 2,6, 7,8. Men wenst de hypothese te toetsen, dat de waarde 7,8 afkomstig is uit een populatie met een hoger gemiddelde. Uit (15.18.1) volgt, dat $r_0 = \frac{7,8 - 2,6}{7,8 - 1,4} = 0,81$. Deze waarde is veel groter dan de kritieke waarde bij de 1% drempel (0,590), zodat de waarde als uitbijter wordt aangemerkt.

15.11. Kan de waarde 17,4 in de 24 serum-calcium gehalten van opgave 10.23 (blz. 254) als uitbijter worden beschouwd? Men vindt met formule (15.19.1), dat $r_0 = \frac{17,4 - 13,4}{17,4 - 9,1} = 0,482$. Tabel Y toont, dat de kritieke waarde van r met $\alpha = 0,05$ gelijk is aan 0,413, zodat de waarde 17,4 significant afwijkt bij een 5% drempel.

15.5.3. ANDERE UITBIJTERTOETSEN

Verschillende andere uitbijtertoetsen worden door DIXON (144) besproken en onderling vergeleken. Hierbij blijkt, dat zij voor $n < 30$ meer rekenwerk vergen en in het algemeen geen groter onderscheidingsvermogen bezitten dan de toets, welke DIXON zelf geeft. Voor $n > 30$ kan men o.i. het best gebruik maken van een door DOORNBOS (148) gegeven toets, maar men dient dan te beschikken over een uitvoerige tabel van STUDENT-verdelingen, die o.m. in de BIOMETRIKATABLES (74) te vinden is.

**TECHNIEKEN BETREFFENDE
DE GEMIDDELDEN VAN TWEE POPULATIES**

Beschouw twee normaal verdeelde populaties, t.w. populatie I met gemiddelde μ_1 en spreiding σ_1 en populatie II met gemiddelde μ_2 en spreiding σ_2 . Onderstel, dat men uit populatie I een aselechte steekproef van n_1 elementen heeft getrokken met gemiddelde \bar{x}_1 en spreiding s_1 en dat een aselechte steekproef van n_2 elementen uit populatie II als gemiddelde \bar{x}_2 en als spreiding s_2 heeft. In dit hoofdstuk bespreken wij nu een aantal technieken, die kunnen worden gebruikt als men op basis van deze twee aselechte steekproeven hypothesen omtrent het verschil $\mu_1 - \mu_2$ wil toetsen of betrouwbaarheidsgrenzen voor dit verschil wil geven.

16.1. Toets voor de hypothese $\mu_1 = \mu_2$ als σ_1 en σ_2 bekend zijn

16.1.1. TWEEZIJDIGE TOETSING

De toetsing verloopt op de volgende wijze:

- a. $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$ (σ_1 en σ_2 zijn bekend).
- b. Kies α .
- c. De toetsingsgrootheid is, als $\sigma_1 \neq \sigma_2$

$$(16.1) \quad \underline{T} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

en als $\sigma_1 = \sigma_2 = \sigma$

$$(16.2) \quad \underline{T} = \frac{\bar{x}_1 - \bar{x}_2}{\sigma} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

- d. Onder H_0 volgt de grootheid \underline{T} volgens (13.23) de standaardnormale verdeling.
- e. Bereken T_0 , door n_1 , n_2 , σ_1^2 en σ_2^2 (resp. σ) en de uit de steekproeven berekende gemiddelden \bar{x}_1 en \bar{x}_2 in (16.1), resp. (16.2) te substitueren.
- f. Verwerp H_0 , als $P_D(T_0) \leq \alpha$ is, d.i. voor $T_0 \leq T_{1-\alpha}$ of $T_0 > T_{1-\alpha}$.

Voorbeeld

16.1. Men wil met $\alpha = 0,05$ de hypothese H_0 toetsen, dat twee pseudo-normaal verdeelde populaties hetzelfde gemiddelde bezitten.

Bekend is, dat $\sigma_1^2 = 4$ en $\sigma_2^2 = 5$. Men trekt uit elke populatie aselect een steekproef van 4 elementen. De uitkomsten zijn: $-2,4, 1,7, -0,9, -1,6$ resp. $6,0, 1,0, 3,2, -0,4$.

Er wordt tweezijdig getoetst en daar $\bar{x}_1 = -0,80$ en $\bar{x}_2 = 2,45$, vindt men met (16.1):

$$T_0 = \frac{-0,80 - 2,45}{\sqrt{4/4 + 5/4}} = \frac{-3,25}{1,50} = -2,17.$$

Uit tabel A blijkt, dat $P_D(-2,17) = 2(0,0150) = 0,03$, zodat H_0 wordt verworpen.

16.1.2. EENZIJDIGE TOETSING

$H_0: \mu_1 = \mu_2$ (of: $\mu_1 \geq \mu_2$), $H_1: \mu_1 < \mu_2$ (σ_1 en σ_2 bekend).

In dit geval vindt links éézijdige toetsing plaats en wordt H_0 bij een drempelwaarde α verworpen ten gunste van H_1 , als $P_L(T_0) < \alpha$ is, d.i. voor $T_0 < T_\alpha$.

$H_0: \mu_1 = \mu_2$ (of: $\mu_1 \leq \mu_2$), $H_1: \mu_1 > \mu_2$ (σ_1 en σ_2 bekend).

Er wordt rechts éézijdig getoetst en H_0 wordt bij een drempelwaarde α verworpen ten gunste van H_1 , als $P_R(T_0) \leq \alpha$ is, d.i. voor $T_0 \geq T_{1-\alpha}$.

16.1.3. DE KEUZE VAN DE OMVANG DER STEEKPROEVEN

Als men kan aangeven, hoe groot het te verwachten verschil tussen de populatiegemiddelden v is, en vaststelt welke kansen α en β op het ten onrechte verwerpen van H_0 , resp. H_1 , men wil aanvaarden, kan ook bij deze toets de noodzakelijke steekproefomvang worden vastgesteld. Als σ_1 en σ_2 verschillen, is de beste procedure n_1 en n_2 evenredig aan hun respectieve populatiespreidingen te maken. Bij *éézijdige toetsing* is dan

$$(16.3.1) \quad n_1 = \frac{\sigma_1^2 + \sigma_1\sigma_2}{v^2} (T_{1-\alpha} + T_{1-\beta})^2,$$

$$(16.3.2) \quad n_2 = \frac{\sigma_2^2 + \sigma_1\sigma_2}{v^2} (T_{1-\alpha} + T_{1-\beta})^2.$$

Is echter $\sigma_1 = \sigma_2 = \sigma$, dan maakt men beide steekproeven even groot:

$$(16.3.3) \quad n = n_1 = n_2 = \frac{2\sigma^2}{v^2} (T_{1-\alpha} + T_{1-\beta})^2.$$

Bij *tweezijdige toetsing* komt in deze formules $T_{1-\alpha}$ in plaats van $T_{1-\alpha}$.

Voorbeeld

16.2. Onderstel dat men de hypothese $H_0: \mu_1 - \mu_2 = 0$ wil toetsen

met als alternatief $H_1: \mu_1 - \mu_2 \geq 1,5$. De populatiespreidingen zijn $\sigma_1 = 5$ en $\sigma_2 = 8$. Men kiest $\alpha = 0,05$ en $\beta = 0,20$. Uit (16.3.1) en (16.3.2) volgt dat

$$n_1 = \frac{25 + 5 \times 8}{(1,5)^2} (1,645 + 0,842)^2 = 179,$$

$$n_2 = \frac{64 + 5 \times 8}{(1,5)^2} (1,645 + 0,842)^2 = 286,$$

waarbij $n_1: n_2 = 179: 286 = \sigma_1: \sigma_2 = 5: 8$.

Blijken de gevonden n_1 en n_2 te groot te zijn, zodat zij bv. uit economisch oogpunt niet zijn te verwezenlijken, dan kan men overwegen de oorspronkelijke keuze van v en/of α en/of β te wijzigen. Bestudeer in dit verband de volgende uitkomsten, die alle gelden bij éézijdige toetsing en $\sigma_1 = 5$ en $\sigma_2 = 8$:

α	β	v	n_1	n_2	α	β	v	n_1	n_2
0,05	0,30	1,5	136	218	0,01	0,30	1,5	235	376
0,05	0,20	1,5	179	286	0,01	0,20	1,5	290	464
0,05	0,10	1,5	248	396	0,01	0,10	1,5	376	602
0,05	0,30	2,0	77	123	0,01	0,30	2,0	132	212
0,05	0,20	2,0	101	161	0,01	0,20	2,0	164	261
0,05	0,10	2,0	140	223	0,01	0,10	2,0	212	339
0,05	0,30	2,5	49	79	0,01	0,30	2,5	85	136
0,05	0,20	2,5	65	103	0,01	0,20	2,5	105	167
0,05	0,10	2,5	90	143	0,01	0,10	2,5	136	217

16.2. Toetsen voor de hypothese $\mu_1 = \mu_2$ als σ_1 en σ_2 onbekend, maar gelijk zijn

16.2.1. TWEEZIJDIGE TOETSING: DE TOETS VAN STUDENT¹

De toets wordt als volgt uitgevoerd:

a. $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$ (bekend is, of ondersteld wordt:

$$\sigma_1 = \sigma_2 = \sigma).$$

b. Kies α .

c. De toetsingsgrootheid is

$$(16.4) \quad t = \frac{\bar{x}_1 - \bar{x}_2}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

waarin:

$$(16.5) \quad s = \sqrt{\frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{\sum x^2 - \left[\frac{(\sum x_1)^2}{n_1} + \frac{(\sum x_2)^2}{n_2} \right]}{n_1 + n_2 - 2}}$$

¹ Zie ook 17.2.3.

Als $n_1 = n_2 = n$ kan men ook nemen

$$(16.6.1) \quad \bar{t} = \frac{(\bar{x}_1 - \bar{x}_2) \sqrt{n}}{\sqrt{s_1^2 + s_2^2}}.$$

- d. Onder H_0 volgt de grootte \bar{t} volgens (13.32) een STUDENT-verdeling met $n_1 + n_2 - 2$ vrijheidsgraden.
 e. Bereken s_0 met (16.5) en vervolgens t_0 met (16.4), of als $n_1 = n_2 = n$ met (16.6.1) of rechtstreeks met

$$(16.6.2) \quad t_0 = (a - b) \sqrt{\frac{n-1}{nc - (a^2 + b^2)}},$$

waarin $a = \Sigma x_1$, $b = \Sigma x_2$ en $c = \Sigma x_1^2 + \Sigma x_2^2 = \Sigma x^2$.

- f. Verwerp H_0 , als $P_D(t_0) < \alpha$ is, d.i. voor $t_0 < t_{1-\alpha}$ ($\nu = n_1 + n_2 - 2$) of voor $t_0 > t_{1-\alpha}$ ($\nu = n_1 + n_2 - 2$).

Voorbeelden

16.3. Ter vergelijking van de fosfaatuitscheiding in de urine (cg per dag) van gezonde en van aan een ziekte S lijdende mannen heeft men de volgende gegevens verzameld (de waarnemingen zijn in volgorde van opklimmende grootte gerangschikt):

Gezonde mannen: 77 84 92 94 98 101 103 105 106 108 111 114
 115 117 119 121 128 129 131 133 136 139 152 159

Mannen met ziekte S: 97 111 125 128 136 137 142 156 165

Men wenst (bij een 5% drempel) de hypothese te toetsen, dat de fosfaatuitscheiding in de urine niet beïnvloed wordt door de ziekte S.

Uit de waarnemingen volgt:

Gezonde mannen: $n_1=24$ $\Sigma x_1=2772$ $\bar{x}_1=115,5$ $\Sigma x_1^2=329734$ $s_1^2=416$
 Zieke mannen : $n_2=9$ $\Sigma x_2=1197$ $\bar{x}_2=133,0$ $\Sigma x_2^2=162729$ $s_2^2=441$

De meeste gehalten van stoffen in het serum of in de urine zijn (enigszins) scheef naar rechts verdeeld, maar afgezien van het feit dat de t -toets hierdoor weinig wordt beïnvloed (zie hoofdstuk 21) is hiervan bij de fosfaatuitscheiding weinig te bemerken; bij de gezonde mannen vindt men als mediaan 114,5 en als gemiddelde 115,5, bij de zieke mannen als mediaan 136 en als gemiddelde 133 (stel zelf een grove verdeling van de uitscheiding bij de gezonde mannen op en bestudeer de vorm daarvan). Men kan dus zonder bezwaar onderstellen, dat het hier steekproeven uit pseudo-normaal verdeelde populaties betreft. Ook de onderstelling, dat de populatievarianties gelijk zijn is gerechtvaardigd, daar tussen de steekproefvarianties een onbetekenend verschil bestaat ($F_0(8, 23) = 441/416 = 1,06$). Men mag dus verwachten, dat de t -toets een valide uitkomst geeft.

De te toetsen hypothese is $H_0: \mu_1 = \mu_2$ met als alternatieve hypo-

these dat de gemiddelde fosfaatuitscheidingen verschillen, dus $H_1: \mu_1 \neq \mu_2$, zodat tweezijdig wordt getoetst. Uit de steekproefgegevens volgt met (16.5) als schatting van σ^2 (σ)

$$s^2 = \frac{(24 - 1) 416 + (9 - 1) 441}{24 + 9 - 2} = \frac{13096}{31} = 422,4516, \quad s = 20,55,$$

waarna men door substitutie in (16.4) vindt dat

$$t_0(31) = \frac{115,5 - 133,0}{20,55} \sqrt{\frac{24 \times 9}{24 + 9}} = -2,18.$$

Uit tabel C blijkt dat $t_{0,025}(31) \cong -2,04$, zodat H_0 wordt verworpen. De conclusie luidt, dat de lijders aan ziekte S een systematisch hogere fosfaatuitscheiding in de urine vertonen dan gezonde mannen.

16.4. Bij een onderzoek van kinderen van 7, 8 en 9 jaar in Drente werd o.m. de alkalische fosfatase in het bloeds serum (in Bodansky-eenheden) bepaald. Twee aselekt getrokken steekproeven, ieder bestaande uit 15 jongens en 15 meisjes, leverden de volgende gegevens op: Kinderen van arbeiders, veengedeelte: $n_1=30$, $\bar{x}_1=9,65$, $s_1^2=4,025$, Kinderen van arbeiders, zandgedeelte: $n_2=30$, $\bar{x}_2=8,13$, $s_2^2=4,890$. De te toetsen hypothese luidt, dat deze steekproeven uit populaties met hetzelfde gemiddelde afkomstig zijn. Men kiest $\alpha = 0,05$.

De F -toets op de varianties levert op: $F_0(29, 29) = 4,890/4,025 = 1,21$, een uitkomst die ver van significant is. Als men op grond hiervan onderstelt dat de populatievarianties gelijk zijn en aanneemt dat de populatieverdelingen pseudo-normaal zijn, kan de hypothese $H_0: \mu_1 = \mu_2$ met de toets van STUDENT worden getoetst. Ook hier vindt tweezijdige toetsing plaats en daar $n_1 = n_2 = n = 30$ kan t_0 met (16.6.1) worden berekend:

$$t_0(58) = \frac{(9,65 - 8,13) \sqrt{30}}{\sqrt{4,025 + 4,890}} = \frac{1,52 \times 5,4772}{2,99} = 2,78.$$

De tweezijdige overschrijdingskans van deze uitkomst blijkt aanzienlijk kleiner te zijn dan α , daar men in tabel C vindt dat $t_{0,975}(58) \cong 2$. De uitspraak luidt dus, dat de gemiddelde alkalische fosfatase bij de arbeiderskinderen in het veengedeelte significant hoger is dan bij de arbeiderskinderen in het zandgedeelte.

16.5a. Tenslotte volgt nog een voorbeeld, waarbij t_0 met (16.6.2) is berekend. Het betreft de volgende aselechte steekproeven uit pseudo-normaal verdeelde populaties:

Steekproef 1: 19 22 15 23 23, Steekproef 2: 13 19 7 16 12.
Bekend is, dat de populatievarianties gelijk zijn. Men wil de hypothese $\mu_1 = \mu_2$ toetsen met $\alpha = 0,05$.

De steekproefgegevens leveren op: $n_1 = n_2 = n = 5$, $\Sigma x_1 = 102$, $\Sigma x_2 = 67$, $\Sigma x^2 = \Sigma x_1^2 + \Sigma x_2^2 = 2128 + 979 = 3107$, zodat met (16.6.2) gevonden wordt:

$$t_0(8) = (102 - 67) \sqrt{\frac{5 - 1}{5(3107) - (102)^2 - (67)^2}} = 2,76.$$

Daar $t_{0,975}(8) = 2,306$ wordt H_0 verworpen.

16.2.2. TWEEZIJDIGE TOETSING: DE TOETS VAN LORD

Als men beschikt over twee aselechte steekproeven van dezelfde omvang n uit twee normaal verdeelde populaties met dezelfde spreiding kan men voor $n \leq 10$ de hypothese $\mu_1 = \mu_2$ (met als alternatief $\mu_1 \neq \mu_2$) bijzonder eenvoudig toetsen via de door LORD (175) aangegeven zg. 'substitute t ratio':

$$(16.7) \quad t^* = \frac{\bar{x}_1 - \bar{x}_2}{\frac{R_1 + R_2}{2}},$$

waarin R_1 en R_2 de spreidingsbreedten van de beide steekproeven zijn. Tabel 16.1 geeft voor tweezijdige toetsing de rechter kritieke waarden $t_{0,975}^*$ en $t_{0,995}^*$ van de verdeling van t^* onder H_0 , die dus gelden voor de drempelwaarden $\alpha = 0,05$, resp. $0,01$. Verder geldt, wegens de symmetrie van de verdeling van t^* , dat $t_{0,025}^* = -t_{0,975}^*$ en $t_{0,005}^* = -t_{0,995}^*$.

Tabel 16.1. Rechter kritieke waarden van t^* bij tweezijdige toetsing

n	$\alpha = 0,05$ $t_{0,975}^*$	$\alpha = 0,01$ $t_{0,995}^*$
2	1,714	3,958
3	0,636	1,046
4	0,406	0,618
5	0,306	0,448
6	0,250	0,357
7	0,213	0,300
8	0,186	0,260
9	0,167	0,232
10	0,152	0,210

Voorbeeld

16.5b. Past men deze toets toe op de steekproeven in 16.5a, dan is $R_1 = 23 - 15 = 8$, $R_2 = 19 - 7 = 12$, $\bar{x}_1 = 102/5 = 20,4$ en $\bar{x}_2 = 67/5 = 13,4$, zodat

$$t_0^* = \frac{20,4 - 13,4}{\frac{8 + 12}{2}} = 0,35.$$

Uit tabel 16.1 blijkt, dat deze uitkomst significant is bij een 5% drempel, daar $t_{0,975}^*(n = 5) = 0,306$ is.

16.2.3. EENZIJDIGE TOETSING: DE TOETS VAN STUDENT

$H_0: \mu_1 = \mu_2$ ($\mu_1 \geq \mu_2$), $H_1: \mu_1 < \mu_2$ (bekend is, of ondersteld wordt: $\sigma_1 = \sigma_2$).

De toetsing geschiedt links éézijdig; H_0 wordt bij een drempelwaarde α verworpen, als $P_L(t_0) \leq \alpha$ is, d.i. voor $t_0 \leq t_\alpha$ ($\nu = n_1 + n_2 - 2$).

$H_0: \mu_1 = \mu_2$ ($\mu_1 < \mu_2$), $H_1: \mu_1 > \mu_2$.

Er wordt rechts éézijdig getoetst. H_0 wordt verworpen bij een drempelwaarde α , als $P_R(t_0) < \alpha$ is, d.i. voor

$$t_0 > t_{1-\alpha}(v = n_1 + n_2 - 2).$$

16.3. Toets voor de hypothese $\mu_1 = \mu_2$ als σ_1 en σ_2 onbekend en ongelijk zijn

Als men weet of vermoedt (bv. op grond van een F -toets, toegepast op de steekproefvarianties) dat de populatievarianties σ_1 en σ_2 niet gelijk zijn, kunnen de in 16.2 beschreven toetsen voor de hypothese $\mu_1 = \mu_2$ niet worden toegepast. Men kan dan gebruik maken van de volgende benaderende methoden.¹

16.3.1. KLEINE STEEKPROEVEN ($n_1 < n_2 < 100$)

In dit geval kan een methode worden toegepast, die aangegeven is door SMITH (195) en uitgewerkt door SATTERTHWAITE (194), WELCH (206) en ASPIN (120). Hierbij is de toetsingsgrootheid

$$(16.8) \quad t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

De BIOMETRIKA TABLES (74) bevatten tabellen met kritieke waarden van t' bij éézijdige toetsing met $\alpha = 0,05$ en $0,01$ voor verschillende waarden van $v_1 = n_1 - 1$ en $v_2 = n_2 - 1$ en

$$c = \frac{s_1^2/n_1}{s_1^2/n_1 + s_2^2/n_2}$$

Men kan echter ook gebruik maken van het feit, dat onder $H_0: \mu_1 = \mu_2$ de grootheid t' bij benadering een STUDENT-verdeling volgt met als aantal vrijheidsgraden

$$(16.9.1) \quad f = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2 (n_1-1)(n_2-1)}{(n_2-1)(s_1^2/n_1)^2 + (n_1-1)(s_2^2/n_2)^2}$$

Voor $n_1 = n_2 = n$ gaat deze formule over in

$$(16.9.2) \quad f = \frac{(s_1^2 + s_2^2)^2 (n-1)}{s_1^4 + s_2^4}$$

Voorbeeld

16.6. Gegeven zijn de twee volgende steekproeven uit pseudo-normaal verdeelde populaties:

$$1) \quad n_1 = 12 \quad \bar{x}_1 = 35,3 \quad s_1^2 = 6 \quad 2) \quad n_2 = 7 \quad \bar{x}_2 = 30,8 \quad s_2^2 = 28.$$

¹ Als men eveneens weet of vermoedt, dat de populatieverdelingen niet normaal zijn, kan men het best de toets van WILCOXON toepassen.

Men wil de hypothese $H_0: \mu_1 = \mu_2$ toetsen met $\alpha = 0,05$.

De steekproefvarianties verschillen significant bij een 5% drempel: $F_0(6, 11) = 28/6 = 4,67$, $F_{0,975}(6, 11) = 3,88$ (tabel F-2). Er is dus reden om aan te nemen, dat σ_2 groter is dan σ_1 . Past men daarom de benadering van WELCH toe, dan is:

$$\frac{s_1^2}{n_1} = \frac{6}{12} = 0,5, \quad \frac{s_2^2}{n_2} = \frac{28}{7} = 4, \quad t'_0 = \frac{35,3 - 30,8}{\sqrt{0,5 + 4,0}} = \frac{4,5}{2,12} = 2,12,$$

$$f = \frac{(0,5 + 4)^2 (11) (6)}{6 (0,5)^2 + 11 (4)^2} = \frac{1336,5}{177,5} = 7,5.$$

Er dient tweezijdig te worden getoetst, d.w.z. H_0 kan bij de gekozen 5% drempel worden verworpen, indien $t'_0 \leq t_{0,025}(\nu = f)$ of $t'_0 \geq t_{0,975}(\nu = f)$. Zoals gewoonlijk het geval zal zijn, is f echter geen geheel getal. Uit tabel C blijkt, dat $t_{0,975}(\nu = 7) = 2,365$ en $t_{0,975}(\nu = 8) = 2,306$. Daar $t'_0 = 2,12$ reeds lager is dan de gunstigste van deze beide kritieke waarden (met $\nu = 8$) wordt H_0 dus niet verworpen. (Zou t'_0 groter geweest zijn dan de ongunstigste kritieke waarde (met $\nu = 7$), dan besluit men tot het verwerpen van H_0 . Ligt t'_0 tussen beide kritieke waarden in, dan heeft men met een twijfelgeval te doen; men weet dan echter, dat $P_D(t'_0)$ dicht bij α ligt. Bij éézijdige toetsing geldt een analoge redenering, die wij aan de lezer overlaten.)

Past men de t -toets volgens 16.2.1 toe dan vindt men:

$$s^2 = \frac{11 \times 6 + 6 \times 28}{11 + 6} = \frac{234}{17} = 13,7647, \quad s = 3,71,$$

$$t_0 = \frac{35,3 - 30,8}{3,71} \sqrt{\frac{12 \times 7}{12 + 7}} = \frac{4,5 \times 2,1}{3,71} = 2,55.$$

Tabel C levert op: $t_{0,975}(\nu = 17) = 2,11$, zodat men ten onrechte tot het verwerpen van de hypothese $\mu_1 = \mu_2$ zou zijn overgegaan.

16.3.2. GROTE STEEKPROEVEN

Als de omvang der steekproeven relatief groot is (n_1 en n_2 groter dan 100) volgt de grootheid

$$(16.10) \quad \underline{T} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

bij benadering een standaardnormale verdeling. De toets verloopt dan als beschreven in 16.1, met s_1^2 in plaats van σ_1^2 en s_2^2 in plaats van σ_2^2 . De benadering is altijd beter voor $n_1 = n_2$ dan voor $n_1 \neq n_2$ en wordt nauwkeuriger naarmate $n = n_1 + n_2$ toeneemt.

16.3.3. OORZAKEN VAN ONGELIJKHEID VAN VARIANTIES

Als men in de praktijk ongelijke varianties waarneemt, verdient het aanbeveling deze niet zonder meer te aanvaarden en naar de zojuist beschreven technieken te grijpen. Men zal er goed aan doen, zich eerst af te vragen wat de oorzaak van deze ongelijkheid kan zijn. Soms zal dan blijken, dat bv. één der beschouwde verdelingen een veel grotere spreiding bezit, omdat deze uitgesproken scheef is of uitbijters vertoont. In zo'n geval kan men dan beter een verdelingsvrije methode kiezen.

Een andere mogelijkheid is, dat de verdelingen van het Poisson-type zijn, zodat de verdeling met een hoger gemiddelde ook een grotere spreiding bezit (dan is immers $s^2 \cong \bar{x}$). Men kan dan de ongelijkheid der spreidingen gewoonlijk wel opheffen door op de waarnemingen de vierkantswortel-transformatie $y = \sqrt{x}$ of $y = \sqrt{x + \frac{1}{2}}$ toe te passen. Ook kan men de door FREEMAN & TUKEY (157) voorgestelde transformatie $y = \sqrt{x} + \sqrt{x+1}$ nemen. Als x volgens Poisson verdeeld is, wordt door deze 'variance-stabilizing transformation' de variantie van y ongeveer gelijk aan 1. In tabel 16.2 is een voorbeeld van beide genoemde transformaties gegeven.

Tabel 16.2. Toepassing van de vierkantswortel-transformatie en de transformatie van FREEMAN & TUKEY op verdelingen van het Poisson-type (aantal fouten in een concentratietest)

	Met storing			Zonder storing		
	x_1	$y_1 = \sqrt{x_1 + \frac{1}{2}}$	$y'_1 = \sqrt{x_1} + \sqrt{x_1 + 1}$	x_2	$y_2 = \sqrt{x_2 + \frac{1}{2}}$	$y'_2 = \sqrt{x_2} + \sqrt{x_2 + 1}$
	6	2,55	5,10	2	1,58	3,15
	5	2,35	4,69	0	0,71	1,00
	10	3,24	6,48	4	2,12	4,24
	6	2,55	5,10	3	1,87	3,73
	12	3,54	7,07	2	1,58	3,15
	9	3,08	6,16	3	1,87	3,73
	4	2,12	4,24	5	2,35	4,69
	11	3,39	6,78	2	1,58	3,15
	9	3,08	6,16	1	1,22	2,41
Gem.	8	2,88	5,75	2,4	1,65	3,25
Var.	8	0,25105	0,9901	2,3	0,23636	1,1641

16.4. Toets voor de hypothese $\mu_1 = \mu_2$ én $\sigma_1 = \sigma_2$

Op basis van twee aselechte steekproeven uit normaal verdeelde populaties kan men ook de hypothese $H_0: \mu_1 = \mu_2$ én $\sigma_1 = \sigma_2$ toetsen. Deze hypothese houdt dus in, dat de beide steekproeven uit identieke normale verdelingen of uit dezelfde normale verdeling afkomstig zijn. De gelijkheid van de varianties is nu een onderdeel van de te toetsen hypothese en onder H_0 volgt dus de grootheid t (16.4) een STUDENT-verdeling met $n_1 + n_2 - 2$ vrijheidsgraden, zodat de toetsing volgens 16.2 kan worden uitgevoerd.

Tegen deze handelwijze zijn echter twee bedenkingen aan te voeren. In de eerste plaats geldt, dat bij het verwerpen van H_0 de conclusie is,

dat (a) $\mu_1 \neq \mu_2$, of (b) $\sigma_1 \neq \sigma_2$, of (c) $\mu_1 \neq \mu_2$ en $\sigma_1 \neq \sigma_2$, d.w.z. men kan met deze toets niet uitmaken, tot welke van deze drie mogelijkheden men zou moeten besluiten. Dit bezwaar vervalt, als men uitsluitend wil onderzoeken of de twee steekproeven uit dezelfde populatie stammen. Een tweede en ernstiger bezwaar is, dat de toets in sommige gevallen een gering onderscheidingsvermogen blijkt te bezitten, nl. als de steekproef van de kleinste omvang ook de kleinste variantie bezit.

Voorbeeld

16.7. Beschouw het geval, dat $\bar{x}_1 - \bar{x}_2 = 2,6$, met $n_1 = 6$, $s_1^2 = 2$, $n_2 = 21$ en $s_2^2 = 14$ (zodat $F_0 = 14/2 = 7$ met $\nu_1 = 20$ en $\nu_2 = 5$, een uitkomst die significant is bij een 5% drempel en tweezijdige toetsing). Toetst men nu de voorgaande 'dubbelhypothese', dan is de uitkomst:

$$s^2 = \frac{5 \times 2 + 20 \times 14}{5 + 20} = 11,6, s = 3,4059, t_0(25) = \frac{2,6}{3,4059} \sqrt{\frac{6 \times 21}{6 + 21}} = 1,65,$$

met een tweezijdige overschrijdingskans die tussen 0,10 en 0,20 ligt (tabel C). Past men echter — op grond van de significant verschillende varianties — de benadering van WELCH toe, dus toetst men de hypothese $\mu_1 = \mu_2$, dan vindt men:

$$\frac{s_1^2}{n_1} = \frac{2}{6} = \frac{1}{3}, \quad \frac{s_2^2}{n_2} = \frac{14}{21} = \frac{2}{3}, \quad t'_0 = \frac{2,6}{\sqrt{1/3 + 2/3}} = 2,6 \quad \text{en}$$

$$f = \frac{(1/3 + 2/3)^2 (5) (20)}{20 (1/3)^2 + 5 (2/3)^2} = \frac{100}{40/9} = 22,5,$$

zodat de tweezijdige overschrijdingskans van t'_0 iets kleiner is dan 0,02. Bij het gebruik van een 5% drempel kan men dus de hypothese $\mu_1 = \mu_2$ en $\sigma_1 = \sigma_2$ niet verwerpen, terwijl men met de F -toets tot het verwerpen van de hypothese $\sigma_1 = \sigma_2$ en met de toets van WELCH tot het verwerpen van de hypothese $\mu_1 = \mu_2$ besluit.

Wij menen uit het voorgaande te moeten concluderen, dat men in de praktijk de toetsing van de hypothese $\mu_1 = \mu_2$ en $\sigma_1 = \sigma_2$ beter achterwege kan laten.

16.5. Betrouwbaarheidsgrenzen van $\mu_1 - \mu_2$

16.5.1. POPULATIEVARIANTIES BEKEND

Als de populatievarianties σ_1 en σ_2 bekend zijn, kan men op grond van twee aselechte steekproeven met gemiddelde \bar{x}_1 , resp. \bar{x}_2 , een $100(1 - \alpha)\%$ betrouwbaarheidsinterval voor $\mu_1 - \mu_2$ verkrijgen met

$$(16.11) \quad (\bar{x}_1 - \bar{x}_2) + T_{1-\alpha} \sigma_{\bar{x}_1 - \bar{x}_2} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + T_{1-\alpha} \sigma_{\bar{x}_1 - \bar{x}_2}$$

waarin

$$(16.12) \quad \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}.$$

Als $\sigma_1 = \sigma_2 = \sigma$ kan men voor (16.11) schrijven:

$$(16.13) \quad (\bar{x}_1 - \bar{x}_2) + T_{1-\alpha} \sigma \sqrt{\frac{n_1 + n_2}{n_1 n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + T_{1-\alpha} \sigma \sqrt{\frac{n_1 + n_2}{n_1 n_2}}.$$

16.5.2. POPULATIEVARIANTIES ONBEKEND

Zijn de populatievarianties onbekend, maar is $\sigma_1 = \sigma_2 = \sigma$, dan kan men op basis van \bar{x}_1 en \bar{x}_2 , en s (berekend met formule (16.5) uit s_1^2 en s_2^2), een $100(1 - \alpha)\%$ betrouwbaarheidsinterval voor het verschil tussen de populatiegemiddelden verkrijgen door te nemen

$$(16.14) \quad (\bar{x}_1 - \bar{x}_2) + t_{1-\alpha} s \sqrt{\frac{n_1 + n_2}{n_1 n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{1-\alpha} s \sqrt{\frac{n_1 + n_2}{n_1 n_2}},$$

waarin $t_{1-\alpha} = -t_{1-1-\alpha}$ met $n_1 + n_2 - 2$ vrijheidsgraden in tabel C kan worden opgezocht.

Als de populatievarianties onbekend en ongelijk zijn, volgt voor kleine steekproeven het $100(1 - \alpha)\%$ betrouwbaarheidsinterval voor $\mu_1 - \mu_2$ bij benadering uit

$$(16.15) \quad (\bar{x}_1 - \bar{x}_2) + t_{1-\alpha} s_{\bar{x}_1 - \bar{x}_2} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{1-\alpha} s_{\bar{x}_1 - \bar{x}_2},$$

waarin

$$(16.16) \quad s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

en $t_{1-\alpha} = -t_{1-1-\alpha}$ volgens formule (16.9) f vrijheidsgraden bezit. Voor relatief grote steekproeven kan men in (16.15) echter t vervangen door T .

Voorbeelden

16.8. Uit de gegevens betreffende de steekproeven in voorbeeld 16.3 volgt met (16.14) het volgende 95% betrouwbaarheidsinterval voor $\mu_2 - \mu_1$:

$$17,5 - 2,04 \times 20,55 \times \sqrt{\frac{24+9}{24 \times 9}} < \mu_2 - \mu_1 < 17,5 + 2,04 \times 20,55 \sqrt{\frac{24+9}{24 \times 9}},$$

$$1,1 < \mu_2 - \mu_1 < 33,9.$$

16.9. Uit de gegevens in voorbeeld 16.6 volgt met (16.15) bij benadering het volgende 95% betrouwbaarheidsinterval voor $\mu_1 - \mu_2$

(dit is iets aan de veilige kant, omdat $t_{0,025} = -t_{0,975}$ met 7 vrijheidsgraden is genomen):

$$4,5 - 2,365 \times 2,12 < \mu_1 - \mu_2 < 4,5 + 2,365 \times 2,12, \\ -0,5 < \mu_1 - \mu_2 < 9,5.$$

16.5.3. DE KEUZE VAN DE OMVANG DER STEEKPROEVEN

Als de populatievarianties ongelijk, doch bekend zijn, kan men een $100(1 - \alpha)\%$ betrouwbaarheidsinterval met breedte i verkrijgen, door voor de omvang van de steekproef uit de populatie met spreiding σ_1 te nemen:

$$(16.17.1) \quad n_1 = \frac{4 T_{1-\frac{1-\alpha}{2}}^2 (\sigma_1^2 + \sigma_1 \sigma_2)}{i^2}$$

en voor de omvang van de steekproef uit de populatie met σ_2 :

$$(16.17.2) \quad n_2 = \frac{4 T_{1-\frac{1-\alpha}{2}}^2 (\sigma_2^2 + \sigma_1 \sigma_2)}{i^2}.$$

Voor $\sigma_1 = \sigma_2 = \sigma$ gaan deze formules over in

$$(16.18) \quad n_1 = n_2 = \frac{8 T_{1-\frac{1-\alpha}{2}}^2 \sigma^2}{i^2}.$$

Ook hier geldt weer, dat men σ_1 en σ_2 weliswaar zelden zal kennen, maar dat men in vele gevallen wel over een voldoende nauwkeurige schatting van deze populatiespreidingen beschikt, om door middel van deze formules een indruk te krijgen van de steekproefomvang die nodig is voor het verkrijgen van een betrouwbaarheidsinterval van (ongeveer) de gewenste breedte.

Voorbeeld

16.10. Uit enkele onderzoeken van grote groepen schoolkinderen tussen 6 en 10 jaar is gebleken, dat de spreiding van het hemoglobinegehalte ongeveer 1 g% bedraagt. Men wil nu een onderzoek instellen naar het verschil tussen de gemiddelde hemoglobinegehalten van twee populaties van 7, 8 en 9 jarige kinderen en wenst daarbij de omvang der steekproeven zodanig te kiezen, dat men voor $\mu_1 - \mu_2$ een 95% betrouwbaarheidsinterval met een breedte van ca. 0,5 g% verkrijgt. Uit (16.18) volgt:

$$n_1 = n_2 = \frac{8 \times (1,96)^2 \times 1}{(0,5)^2} = 32 \times 3,8416 = 122,9.$$

Men zal dus uit iedere populatie een steekproef van ongeveer 123 kinderen moeten trekken.

16.6. Opgaven

16.1. Uit de populaties $N(\mu_x, 4)$ en $N(\mu_y, 6)$ ¹ worden aselechte steekproeven getrokken, resp. van 9 en 25 elementen. Deze leveren op: $\bar{x} = 10$, $\bar{y} = 15$, $s_x^2 = 25$, $s_y^2 = 25$.

- Toets bij een 5% drempel de hypothese $\mu_x = \mu_y$.
- Bepaal het 95% betrouwbaarheidsinterval van $\mu_x - \mu_y$.

16.2. Uit de populaties $N(\mu_x, \sigma_x)$ en $N(\mu_y, \sigma_y)$ worden aselechte steekproeven getrokken, resp. van 21 en 21 elementen. Men vindt: $\bar{x} = 67,3$, $\bar{y} = 63,7$, $s_x^2 = 85$, $s_y^2 = 87$.

- Toets bij een 5% drempel de hypothese $\mu_x = \mu_y$.
- Bepaal het 95% betrouwbaarheidsinterval van $\mu_x - \mu_y$.

16.3. Van twee aselechte steekproeven uit normaal verdeelde populaties is gegeven: $n_1 = 6$, $\bar{x}_1 = 31,9$, $s_1^2 = 25$; $n_2 = 5$, $\bar{x}_2 = 25,1$, $s_2^2 = 49$.

Toets bij een 5% drempel de hypothese $H_0: \mu_1 < \mu_2$ (met $H_1: \mu_1 > \mu_2$)

- Als ondersteld wordt, dat $\sigma_1 = \sigma_2$.
- Zonder deze onderstelling te maken.

16.4. Voor twee aselechte steekproeven uit normaal verdeelde populaties geldt: $n_1 = 390$, $\bar{x}_1 = 63,4$, $s_1^2 = 15,3$ en $n_2 = 406$, $\bar{x}_2 = 61,8$, $s_2^2 = 28,8$.

Toets de hypothese $\mu_1 = \mu_2$ met $\alpha = 0,05$.

16.5. Gegeven zijn de volgende aselechte steekproeven:

steekproef 1: 3,28 3,29 3,26 3,28 3,31,

steekproef 2: 3,25 3,25 3,23 3,27 3,22.

Onderstel dat de populaties waaruit deze steekproeven stammen normaal verdeeld zijn en bepaal het 95% en het 99% betrouwbaarheidsinterval van $\mu_1 - \mu_2$. Tot welke uitspraak omtrent de hypothese $\mu_1 = \mu_2$ kunt U op grond van deze betrouwbaarheidsintervallen komen?

16.6. Bepaal op grond van de volgende gegevens de noodzakelijke steekproefgrootten:

a. $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 - \mu_2 = 2$, $\sigma_1 = 12$, $\sigma_2 = 16$, $\alpha = 0,05$, $\beta = 0,20$.

b. $H_0: \mu_1 - \mu_2 = 0$, $H_1: \mu_1 - \mu_2 = -3$, $\sigma_1 = \sigma_2 = 5$, $\alpha = 0,01$, $\beta = 0,10$.

16.7. Hoe groot moet men in onderstaande gevallen de steekproeven kiezen, om een 95% betrouwbaarheidsinterval van $\mu_1 - \mu_2$ met een breedte i te verkrijgen?

a. $\sigma_1 = 5$, $\sigma_2 = 10$, $i = 3$

b. $\sigma_1 = \sigma_2 = 10$, $i = 5$.

16.8. De uitkomsten van de concentratietest in tabel 16.2 betreffen een groep van 18 recruten, die in opleiding waren voor radiotelegrafist. Deze groep werd aselekt in twee subgroepen van 9 man verdeeld. Eén subgroep kreeg de test met, de andere zonder storing. Men wil de hypothese toetsen, dat de aantallen fouten in deze groepen slechts toevallig verschillen. Voer de toetsing op elk van de volgende manieren uit, bepaal voor elke toetsing (bij benadering) de tweezijdige overschrijdingskans en vergelijk de gevonden kansen.

a. Toets van WILCOXON.

b. t -toets op de gemiddelden van de x -waarden, zonder rekening te houden met het verschil tussen de spreidingen.

c. Als (b), maar rekening houdend met het feit dat de spreidingen significant verschillen bij een 5% drempel.

¹ Wij memoreren, dat $N(\mu, \sigma)$ betekent: de normale verdeling met gemiddelde μ en spreiding σ .

- d. t -toets op de gemiddelden van de y -waarden.
 e. t -toets op de gemiddelden van de y' -waarden.
 f. Toets van LORD op de gemiddelden van de y -waarden.
 g. Toets van LORD op de gemiddelden van de y' -waarden.

16.9. Een groep van 16 albino-ratten van 180 dagen wordt aselekt in twee subgroepen van gelijke omvang gesplitst. Men voert vervolgens een experiment uit, waarbij subgroep 1 een zekere behandeling ondergaat en subgroep 2 als contróle dient. Daarna bepaalt men van alle ratten het stikstofgehalte in het plasma (gr per 100 cc). De uitkomsten zijn:

Groep 1	Rat nr	1	5	8	10	11	13	15	16
	Gehalte	1,19	1,21	1,13	1,18	1,33	1,27	1,20	1,07
Groep 2	Rat nr	2	3	4	6	7	9	12	14
	Gehalte	1,05	1,09	1,02	0,99	1,16	1,03	0,91	1,19

Bij voorgaande experimenten is reeds gebleken, dat het stikstofgehalte in het plasma pseudo-normaal verdeeld is. Toets bij een 1% drempel de hypothese, dat het stikstofgehalte door de behandeling niet is beïnvloed:

- a. Met de t -toets van STUDENT. Geef tevens het 99% betrouwbaarheids-interval voor $\mu_1 - \mu_2$.
 b. Met de toets van LORD.
 Welke onderstelling maakt U bij deze toetsen?

16.10. Uit H. ELSBACH, Klinische onderzoekingen met de laag-frekwente ballistocardiograaf volgens BURGER, Diss., Leiden, 1954. Frequentieverdelingen van de cd-tijd (sec. \times 100) bij mannen, 1 = beneden 40 jaar, 2 = 40 jaar en ouder

x	7	8	9	10	11	12	13	14	15	16	Totaal
1				2	3	10	5	7	2	1	30
2	1	2	2	2	1	3	7	1	1		20

Toets met $\alpha = 0,05$ de hypothese $\mu_1 = \mu_2$. Gegeven is, dat $\bar{x}_1 = 12,7$, $s_1^2 = 2,1333$, $\bar{x}_2 = 11,4$, $s_2^2 = 5,0947$.

16.11. Uit F. VAN FASSEN, Hypothyreoidie en aangeboren misvormingen, Diss., Utrecht, 1957.

De volgende gegevens betreffen albino ratten. Gemeten is de lengte van de binnenste plexiforme laag van de retina bij 80-voudige vergroting, ogen van foetus van normale moeders, A = direct gefixeerd, B = na 24-48 uur gefixeerd. Ter vereenvoudiging van de berekeningen zijn de gemeten lengten met 20 verminderd.

- A. 8,3 7,3 6,5 8,3 7,7 6,3 6,8 10,7 5,4 5,3 6,0 9,6 6,9 10,8 5,3.
 B. 4,0 5,2 2,1 7,1 5,9 6,0.

Gegeven is: $\bar{x}_A = 7,41$, $s_A^2 = 3,3155$, $\bar{x}_B = 5,05$, $s_B^2 = 3,1310$.

Onderstel, dat de groepen A en B aselekt zijn samengesteld en dat de lengteverdelingen pseudo-normaal zijn. Bepaal het 95% betrouwbaarheidsinterval van $\mu_A - \mu_B$.

TECHNIEKEN BETREFFENDE DE GEMIDDELDEN VAN k POPULATIES: VARIANTIE-ANALYSE

In hoofdstuk 16 zijn de methoden besproken, waarmee men op basis van aselechte steekproeven hypothesen kan toetsen of schattingsintervallen kan verkrijgen betreffende de gemiddelden van twee pseudo-normaal verdeelde populaties. Dit hoofdstuk handelt over de variantie-analyse, een techniek die kan worden gebruikt voor het toetsen van de hypothese, dat de gemiddelden van *meer dan twee* populaties gelijk zijn.

17.1. Enkelvoudige analyse: k aselechte steekproeven van dezelfde omvang

17.1.1. ONDERSTELLINGEN EN NULHYPOTHESE

Onderstel, dat men beschikt over k aselechte steekproeven van dezelfde omvang n uit k normaal verdeelde populaties met dezelfde spreiding σ . Door middel van de variantie-analyse kan men dan op basis van deze steekproeven de hypothese

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k = \mu$$

toetsen met als alternatieve hypothese, dat minstens één van de populatiegemiddelden van de andere afwijkt.

17.1.2. DE TOETSINGSGROOTHEID EN HAAR VERDELING ONDER H_0

Onder H_0 volgen de steekproefgemiddelden \bar{x}_j ($j = 1, 2, \dots, k$) een normale kansverdeling met gemiddelde $\mu_{\bar{x}} = \mu$ en met variantie $\sigma_{\bar{x}}^2 = \sigma^2/n$. Men kan uit de k waargenomen gemiddelden een zuivere schatting $s_{\bar{x}}^2$ van $\sigma_{\bar{x}}^2$ verkrijgen door te nemen

$$(17.1) \quad s_{\bar{x}}^2 = \frac{(\bar{x}_1 - \bar{x})^2 + (\bar{x}_2 - \bar{x})^2 + \dots + (\bar{x}_k - \bar{x})^2}{k - 1} = \frac{\sum_{j=1}^k (\bar{x}_j - \bar{x})^2}{k - 1},$$

waarin \bar{x} = het gemiddelde van de k steekproefgemiddelden
= het gemiddelde van alle kn waarnemingen.

Hieruit volgt, dat men een zuivere schatting van σ^2 kan verkrijgen met

$$(17.2) \quad s_f^2 = ns_{\bar{x}}^2 = \frac{n \sum_{j=1}^k (\bar{x}_j - \bar{x})^2}{k - 1}.$$

17.1 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

De index t is toegevoegd om aan te geven, dat deze schatting van σ^2 verkregen is uit de variantie tussen de steekproefgemiddelden. Men noemt

$$(17.3) \quad n \sum_{j=1}^k (\bar{x}_j - \bar{x})^2 = \frac{(\sum x_1)^2 + (\sum x_2)^2 + \dots + (\sum x_k)^2}{n} - \frac{(\sum x)^2}{kn}$$

de *kwadratensom tussen de steekproeven*. In deze formule is $\sum x = \sum x_1 + \sum x_2 + \dots + \sum x_k$ (wij laten het aan de lezer over, de herleiding van het linkerlid tot het rechterlid uit te voeren; bedenk hierbij dat $\bar{x}_j = \sum x_j/n$ en dat $\bar{x} = \sum x/kn$).

Ondersteld is, dat alle populatieverdelingen dezelfde spreiding σ bezitten. De steekproefvarianties s_j^2 zijn dus alle schattingen van dezelfde parameter σ^2 , zodat met (14.7) een tweede zuivere schatting van σ^2 kan worden verkregen:

$$(17.4) \quad s^2 = \frac{\sum_{j=1}^k (n_j - 1) s_j^2}{\sum_{j=1}^k (n_j - 1)} = \frac{\sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{k(n-1)} = \frac{\sum_{j=1}^k s_j^2}{k}$$

Men noemt

$$(17.5) \quad \sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 = (n-1) \sum s_j^2 = \sum x^2 - \frac{(\sum x_1)^2 + (\sum x_2)^2 + \dots + (\sum x_k)^2}{n}$$

de *kwadratensom binnen de steekproeven* (zie ook 13.7.2)¹.

Uit alle kn waarnemingen kan men nog een derde kwadratensom, de z.g. *totale kwadratensom*, verkrijgen door te nemen

$$(17.6) \quad \sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x})^2 = \sum x^2 - \frac{(\sum x)^2}{kn}$$

Wij gebruiken voor een kwadratensom het symbool S , met een index die aangeeft, welke kwadratensom bedoeld wordt; dus bv. $S_{\text{totaal}} =$ totale kwadratensom. De lezer kan nu eenvoudig verifiëren, dat tussen de drie behandelde kwadratensommen de volgende betrekking bestaat:

$$(17.7) \quad \sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x})^2 = n \sum_{j=1}^k (\bar{x}_j - \bar{x})^2 + \sum_{j=1}^k \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$$

$$(S_{\text{totaal}}) = (S_{\text{tussen}}) + (S_{\text{binnen}})$$

Bij de berekening van de kwadratensommen bepaalt men gewoonlijk eerst S_{totaal} , vervolgens S_{tussen} en tenslotte S_{binnen} als het verschil van

¹ Het dubbele somteken geeft aan, dat men eerst voor elke steekproef afzonderlijk $\sum (x_i - \bar{x})^2$ dient te berekenen en de verkregen kwadratensommen vervolgens over alle k steekproeven moet sommeren. Verder geldt, dat $\sum (n_j - 1) = \sum n_j - k = kn - k = k(n - 1)$.

de eerstgenoemde sommen. Daarna verkrijgt men dus de beide schattingen van σ^2 met

$$(17.8) \quad s_i^2 = \frac{S_{tussen}}{k-1}, \quad \text{en} \quad (17.9) \quad s_b^2 = \frac{S_{binnen}}{k(n-1)}.$$

Volgens 13.3.3 zijn deze twee schattingen van σ^2 stochastisch onafhankelijk, zodat onder H_0 volgens (13.13) de grootheid

$$(17.10) \quad \underline{F} = s_i^2/s_b^2$$

een F -verdeling volgt met $\nu_1 = k - 1$ en $\nu_2 = k(n - 1)$.

17.1.3. UITVOERING VAN DE TOETSING

De k populatieverdelingen bezitten dezelfde variantie, dus ook als H_0 niet juist is, blijft s_b^2 een zuivere schatting van σ^2 . De populatiegemiddelden zijn dan echter niet gelijk: er is minstens één gemiddelde, dat van de andere verschilt. De steekproefgemiddelden bezitten dan echter kansverdelingen, die *niet* alle hetzelfde gemiddelde hebben; diensengevolge zal s_i^2 systematisch hogere waarden aannemen dan s_b^2 . Er dient dus *rechts éénzijdig* te worden getoetst.

De toetsing kan worden uitgevoerd, door uit de waarnemingen s_i^2 en s_b^2 en vervolgens $F_0 = s_i^2/s_b^2$ te berekenen. Toetst men bij een drempelwaarde α , dan wordt H_0 verworpen, wanneer $P_R(F_0) \leq \alpha$ is, d.i. dus voor $F_0 \leq F_{1-\alpha}[\nu_1 = k-1, \nu_2 = k(n-1)]$. Hieruit volgt, dat men direct tot het niet verwerpen van H_0 kan besluiten, als $F_0 < 1$, d.i. als $s_i^2 < s_b^2$ is.

De hypothese, dat de steekproeven uit dezelfde populatie afkomstig zijn kan verdelingsvrij worden getoetst met de toets van KRUSKAL & WALLIS (H -toets, zie 10.6). Deze hypothese is niet identiek aan de nulhypothese bij de variantie-analyse en uit een significante uitkomst van de H -toets kan men dan ook strikt genomen niet concluderen, dat alleen de gemiddelden van de populaties een verschil of verschillen vertonen. Er zijn echter redenen om aan te nemen, dat de H -toets vrij ongevoelig is voor verschillen in de spreidingen van de populatieverdelingen. Vindt men dus een significante uitkomst, dan kan men deze zonder veel bezwaar als een indicatie voor een verschil in gemiddelden beschouwen.

Als H_0 verworpen wordt, zal men gewoonlijk de analyse van de verschillen tussen de gemiddelden nog verder willen voortzetten. In 17.1.5 bespreken wij een procedure welke daarbij gevolgd kan worden.

Voorbeelden

17.1. In tabel 17.1 is de enkelvoudige variantie-analyse toegepast op 4 aselect getrokken steekproeven van dezelfde omvang uit pseudo-normaal verdeelde populaties met dezelfde spreiding. De in deze tabel uitgevoerde berekeningen bestaan in feite uit drie delen:

17.1 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

- a. De berekening van Σx_j , en Σx_j^2 en van de steekproefgemiddelden \bar{x}_j . Hierbij kan men soms het rekenwerk sterk vereenvoudigen, door van alle waarnemingen een constante (liefst in de nabijheid van \bar{x}) af te trekken.
- b. De berekening van de kwadratsommen.
- c. De berekening van s_i^2 , s_b^2 en van F_0 . Men pleegt deze samen te vatten in een afzonderlijke tabel, welke eveneens in tabel 17.1 is opgenomen. Merk op, dat s_b^2 ook volgt uit $\Sigma s_j^2/k = 56,1/4 = 14,025$. Als men de steekproefvarianties heeft berekend, kan men deze formule dus als controle gebruiken.

De analyse levert op $F_0 = 4,04$. Volgens tabel F is $F_{0,95}(3, 16) = 3,24$ en $F_{0,975}(3, 16) = 4,08$. De hypothese, dat de populatiegemiddelden gelijk zijn kan dus verworpen worden met een onbetrouwbaarheid, die bijna gelijk is aan 0,025.

Tabel 17.1. Enkelvoudige variantie-analyse: toetsing van de hypothese $\mu_1 = \mu_2 = \mu_3 = \mu_4$ op basis van aselechte steekproeven met $n = 5$ uit pseudo-normaal verdeelde populaties met dezelfde spreiding

Steekproeven								
1		2		3		4		
x_1	x_1^2	x_2	x_2^2	x_3	x_3^2	x_4	x_4^2	
16	256	17	289	25	625	16	256	
18	324	22	484	28	784	16	256	
21	441	18	324	21	441	24	576	
12	144	23	529	19	361	19	361	
15	225	22	484	30	900	24	576	
Σx_j	82	102		123		99		$\Sigma x = 406$
Σx_j^2	1390	2110		3111		2025		$\Sigma x^2 = 8636$
\bar{x}_j	16,4	20,4		24,6		19,8		
s_j^2	11,3	7,3		21,3		16,2		$\Sigma s_j^2 = 56,1$

Berekening van de kwadratensommen

$$S_{\text{totaal}} = \Sigma x^2 - (\Sigma x)^2/kn = 8636 - (406)^2/4 \times 5 = 8636 - 8241,8 = 394,2$$

$$S_{\text{tussen}} = \frac{(\Sigma x_1)^2 + \dots + (\Sigma x_4)^2}{n} - \frac{(\Sigma x)^2}{kn} = \frac{(82)^2 + \dots + (99)^2}{5} - 8241,8 = 169,8$$

$$S_{\text{binnen}} = 224,4$$

Uitkomsten van de variantie-analyse

	S	Vrijheidsgraden	Variantieschatting
Tussen de steekproeven	169,8	$k-1 = 3$	$169,8/3 = 56,600 = s_i^2$
Binnen de steekproeven	224,4	$k(n-1) = 16$	$224,4/16 = 14,025 = s_b^2$
Totaal	394,2	$kn-1 = 19$	$F_0(3, 16) = \frac{56,600}{14,025} = 4,04$

17.2. In tabel 17.2 zijn de uitkomsten vermeld van een proef, waarbij ieder van 10 analysten (op identieke, nauwkeurig voorgeschreven wijze) 5 onderling onafhankelijke bepalingen verrichtte van de *Fe*-concentratie van een standaardoplossing (met 2,95% *Fe*). Men wil onderzoeken of, en zo ja welke, systematische verschillen tussen de uitkomsten van deze analysten optreden.

De verrichte waarnemingen kunnen worden beschouwd als aselecte steekproeven van 5 elementen uit 10 verschillende populaties. De denkbeeldige populatie *i* bestaat dan uit alle bepalingen, die de *i*-de analist zou kunnen verrichten. Wil men nu via de variantie-analyse de hypothese H_0 toetsen, dat deze populaties hetzelfde gemiddelde bezitten (dat de verschillen tussen de uitkomsten van de analysten toevallig zijn), dan dient te worden ondersteld, dat deze populaties pseudo-normaal verdeeld zijn en dat zij dezelfde spreiding σ hebben.

Tabel 17.2. Uitkomsten van *Fe*-bepalingen van 10 analysten (5 onderling onafhankelijke bepalingen per analist) betreffende een standaardoplossing met 2,95% *Fe*

Analyst	Bepaling (x)					Gemiddelde
	1	2	3	4	5	
1	2,963	2,996	2,979	2,970	2,979	2,9774
2	2,958	2,964	2,955	2,932	2,941	2,9500
3	2,956	2,945	2,963	2,950	2,975	2,9578
4	2,948	2,960	2,953	2,944	2,950	2,9510
5	2,953	2,961	2,961	2,953	2,949	2,9554
6	2,941	2,940	2,931	2,942	2,930	2,9368
7	2,963	2,928	2,925	2,940	2,934	2,9380
8	2,987	2,989	2,988	2,983	2,974	2,9842
9	2,946	2,950	2,955	2,969	2,954	2,9548
10	2,950	2,947	2,947	2,960	2,954	2,9516
						2,9557

Uit BENNETT & FRANKLIN (4), tabel 7.7.

De eerstgenoemde onderstelling onttrekt zich bij dit kleine aantal bepalingen per analist aan verificatie, maar uit proeven van grotere omvang is gebleken, dat dit soort bepalingen inderdaad ongeveer normaal verdeeld is. Voor de tweede onderstelling beschouwen wij de verkregen uitkomsten. Teneinde de berekeningen te vereenvoudigen, zijn deze in tabel 17.3 omgezet in de hulpgrootheid $u = 1000(x - 2,955)$, waarna per analist de variantie van de vijf *u*-waarden is bepaald. Deze blijken niet significant te verschillen (toets van HARTLEY, $k = 10$, $n - 1 = 4$, $F_{max} = 228,5/28,8 = 7,93$, $F_{0,95} = 44,6$: tabel P). De onderstelling, dat de populatievarianties gelijk zijn, kan daarom o.i. zonder bezwaar worden gemaakt.

17.1 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

Tabel 17.3. Uitkomsten van de F_e -bepalingen in tabel 17.2, omgezet in de hulpgrootheid $u = 1000 (x - 2,955)$

Analyst	Bepaling (u)					Σu_i	Σu_i^2	\bar{u}_i	S_u	s_i^2
	1	2	3	4	5					
1	8	41	24	15	24	112	3122	22,4	613,2	153,3
2	3	9	0	-23	-14	-25	815	-5,0	690,0	172,5
3	1	-10	8	-5	20	14	590	2,8	550,8	137,7
4	-7	5	-2	-11	-5	-20	224	-4,0	144,0	36,0
5	-2	6	6	-2	-6	2	116	0,4	115,2	28,8
6	-14	-15	-24	-13	-25	-91	1791	-18,2	134,8	33,7
7	8	-27	-30	-15	-21	-85	2359	-17,0	914,0	228,5
8	32	34	33	28	19	146	4414	29,2	150,8	37,7
9	-9	-5	0	14	-1	-1	303	-0,2	302,8	75,7
10	-5	-8	-8	5	-1	-17	179	-3,4	121,2	30,3
$k = 10 \quad n = 5$						35	13913	0,7	3736,8	934,2
						Σu	Σu^2	\bar{u}	ΣS_u	Σs_i^2

Tabel 17.3 bevat ook alle gegevens die nodig zijn, om de variantie-analyse (toetsing van de hypothese $H_0: \mu_1 = \mu_2 = \dots = \mu_{10}$) te kunnen uitvoeren. Volgens (17.6) is

$$S_{\text{totaal}} = \Sigma u^2 - \frac{(\Sigma u)^2}{kn} = 13913 - \frac{(35)^2}{10 \times 5} = 13913 - 24,5 = 13888,5$$

en uit (17.3) volgt

$$S_{\text{tussen}} = \frac{(112)^2 + (-25)^2 + \dots + (-17)^2}{5} - \frac{(35)^2}{10 \times 5} = 10176,2 - 24,5 = 10151,7.$$

Hieruit volgt

$$S_{\text{binnen}} = S_{\text{totaal}} - S_{\text{tussen}} = 13888,5 - 10151,7 = 3736,8.$$

Merk op, dat deze kwadratensom ook direct uit tabel 17.3 kan worden afgelezen door te nemen ΣS_u . De verkregen kwadratensommen en de hiermee corresponderende variantieschattingen zijn in tabel 17.4 samengevat.

Tabel 17.4. Uitkomsten van de variantie-analyse

	Kwadraten-som	Vrijheids-graden	Variantie-schatting
Tussen de analysten	10151,7	9	1127,97 = s_t^2
Binnen de analysten	3736,8	40	93,42 = s_b^2
Totaal	13888,5	49	

Ter controle kan dienen, dat s_b^2 ook volgt uit $\Sigma s_i^2/k = 934,2/10 = 93,42$. Uit de tabel blijkt, dat $F_0 = s_t^2/s_b^2 = 1127,97/93,42 = 12,07$. Daar $F_{0,99} (v_1 = 9, v_2 = 40) = 2,89$, wordt H_0 verworpen: de overschrij-

dingskans van F_0 is aanzienlijk kleiner dan 0,01.¹ De uitspraak luidt dus, dat er systematische verschillen tussen de uitkomsten van de analysten aanwezig zijn.

Met deze conclusie is echter slechts het eerste deel van de vraagstelling van de proef beantwoord. In de praktijk zal men immers veelal meer omtrent de geconstateerde verschillen tussen de analysten willen weten. Wij bespreken de wijze, waarop de statistische analyse dan kan worden voortgezet in 17.1.5.

17.1.4.* Toetsing via de spreidingsbreedte

De in het voorgaande beschreven toets voor de hypothese $\mu_1 = \mu_2 \dots = \mu_k$ kan — zij het met enig verlies aan onderscheidingsvermogen — ook worden gebaseerd op de spreidingsbreedte van de k steekproefgemiddelden, $\bar{R} = \bar{x}_{\max} - \bar{x}_{\min}$. Als de omvang van elk der steekproeven n is en als s_b^2 de variantie binnen de steekproeven is, volgt de grootheid

$$(17.11) \quad \underline{w} = \bar{R} \sqrt{\frac{n}{s_b^2}} = \frac{\bar{R}}{s_b} \sqrt{n}$$

een kansverdeling, waarvan voor $k = 2, 3, \dots, 20$ en voor verschillende waarden van het aantal bij s_b^2 behorende vrijheidsgraden, $\nu = k(n-1)$, de rechter kritieke waarden (fractielen) $w_{0,95}$ en $w_{0,99}$ in tabel Z gegeven zijn. Men kan dus bij een drempelwaarde α tot het verwerpen van H_0 besluiten, wanneer $w_0 > w_{1-\alpha} [k, k(n-1)]$ is.

Deze toets kan bv. worden gebruikt, als men beschikt over een publicatie, waarin alleen de gemiddelden en de varianties van k steekproeven van dezelfde omvang gegeven zijn en de juistheid van H_0 snel wil onderzoeken.

Voorbeelden

17.3. Voor de $k = 4$ steekproeven van de omvang $n = 5$ in tabel 17.1 geldt: $k(n-1) = 16$, $\bar{R}_0 = 24,6 - 16,4 = 8,2$, $s_b^2 = 14,025$, zodat

$$w_0 = 8,2 \sqrt{5/14,025} = 4,90.$$

Uit tabel Z blijkt, dat voor $k = 4$ en $\nu = 16$ de kritieke waarden van \underline{w} zijn: $w_{0,95} = 4,05$ en $w_{0,99} = 5,19$. Ook deze toets via de spreidingsbreedte der gemiddelden leidt dus tot de conclusie, dat H_0 bij een 5% drempel kan worden verworpen.

17.4. Voor de F_0 -bepalingen in voorbeeld 17.2 is (zie tabel 17.3): $\bar{R}_0 = 29,2 - (-18,2) = 47,4$ en $s_b^2 = 93,42$, zodat

$$w_0 = 47,4 \sqrt{5/93,42} = 10,97.$$

In tabel Z vindt men voor $k = 10$ en $\nu = 40$ als kritieke waarde $w_{0,99} = 5,60$.

17.1.5. VOORTZETTING VAN DE ANALYSE ALS H_0 VERWORPEN WORDT

Als bij toetsing van de hypothese $H_0: \mu_1 = \mu_2 = \dots = \mu_k$ tot het verwerpen van H_0 wordt besloten, kunnen zich omtrent de verschillen tussen de populatiegemiddelden allerlei vragen voordoen. Deze worden bv. door BENNETT & FRANKLIN (4) uitvoerig besproken. Wij beperken ons tot de vraag, die gewoonlijk praktisch van het meeste belang is:

¹ Uit uitvoeriger tabellen van de F -verdeling blijkt, dat $F_{0,99} = 4,02$, zodat deze kans zelfs veel kleiner is dan 0,001.

17.1 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

hoe kunnen de populaties gesplitst worden in groepen, waarbinnen de gemiddelden (ongeveer) hetzelfde zijn?

In de afgelopen jaren hebben velen zich met dit probleem bezig gehouden en zijn er verschillende, min of meer bevredigende technieken voor de oplossing ervan ontwikkeld. DUNCAN (150) bespreekt een aantal hiervan en geeft ook een uitgebreide literatuuropgave. Wij behandelen slechts één methode, afkomstig van TUKEY (205), die kan worden gebruikt als alle steekproeven dezelfde omvang hebben en waarvoor geen speciale tabellen nodig zijn. Deze methode bestaat uit drie fasen: (a) De steekproefgemiddelden worden naar grootte gerangschikt en op elk paar aangrenzende gemiddelden wordt, met een drempelwaarde α , de t -toets van STUDENT toegepast. Openvolgende gemiddelden, die niet significant blijken te verschillen, vormen dan een groep (zoals later zal blijken kan zo'n groep bestaan uit 1, 2, of meer dan 2 gemiddelden). (b) Elke groep met meer dan twee gemiddelden wordt vervolgens afzonderlijk onderzocht op het voorkomen van gemiddelden, die 'te veel' van het groepsgemiddelde afwijken (TUKEY noemt dergelijke gemiddelden 'stragglers'). Hiertoe beschouwt men successievelijk het verschil $v_{\max} = \bar{x}_a - \bar{x}_g$, waarbij \bar{x}_g het gemiddelde van de gehele groep is en \bar{x}_a het steekproefgemiddelde in deze groep, dat daarvan het meest afwijkt. Hierbij wordt gebruik gemaakt van een benadering van de verdeling van de grootheid $(\bar{x}_a - \bar{x}_g)/s$, die afkomstig is van NAIR (181). Indien het extreme gemiddelde \bar{x}_a bij een drempelwaarde α significant blijkt af te wijken van het groepsgemiddelde, dan wordt dit van de groep afgescheiden. Deze procedure wordt voor elke groep voortgezet, totdat geen verdere gemiddelden meer afvallen. (c) Tenslotte wordt elke groep met drie of meer gemiddelden, die bij de fasen (a) of (b) is verkregen, gecontroleerd op homogeniteit door middel van de variantie-analyse.

Een bezwaar van deze methode is, dat de verschillende achtereenvolgens toegepaste toetsen niet onderling onafhankelijk zijn, zodat de onbetrouwbaarheid van de uitspraak omtrent de verschillen tussen de (groepen) gemiddelden die men ermee verkrijgt, iets groter zal zijn dan de drempelwaarde α , die men bij elk der afzonderlijke toetsingen gebruikt. TUKEY zelf spreekt dan ook van 'a primitive and rough standard'. Onze ervaring is, dat de methode betrekkelijk snel tot een bruikbare en althans objectieve splitsing van de gemiddelden leidt.

De volledige procedure verloopt nu, uitgevoerd met een onbetrouwbaarheidsdrempel α , als volgt:

(a) Bereken met behulp van de waarde van s_y^2 , verkregen bij de voorafgaande variantie-analyse, het verschil tussen twee steekproefgemiddelden, dat bij toepassing van de t -toets van STUDENT juist significant zou zijn bij tweezijdige toetsing met de drempelwaarde α , d.i.

$$(17.12) \quad \bar{v} = t_{1-\frac{1}{2}\alpha} s_{\bar{x}_j} \sqrt{2},$$

waarin $s_{\bar{x}_j} = s_b/\sqrt{n}$. De waarde van $t_{1-\frac{1}{2}\alpha}$ met $k(n-1)$ vrijheidsgraden kan worden opgezocht in tabel C.

Rangschik de steekproefgemiddelden naar grootte en verdeel ze in groepen door elk verschil tussen twee opeenvolgende gemiddelden, dat groter is dan \bar{v} , als een scheiding tussen twee groepen te beschouwen. Als geen enkele groep meer dan twee gemiddelden bevat, wordt de analyse beëindigd en de ontstane groepen (die dan ook uit één gemiddelde kunnen bestaan) worden 'significant verschillend' verklaard. Twee gemiddelden in één groep worden geacht niet significant te verschillen.

(b) Bereken voor elke groep van drie of meer gemiddelden het groepsgemiddelde \bar{x}_g en vervolgens $v_{\max} = |\bar{x}_a - \bar{x}_g|$, waarbij \bar{x}_a het gemiddelde in de groep is, dat het meest van het groepsgemiddelde verschilt. Als de groep uit drie gemiddelden bestaat berekent men vervolgens

$$(17.13) \quad T_0 = \frac{\frac{v_{\max}}{s_{\bar{x}_j}} - \frac{1}{2}}{3 \left(\frac{1}{4} + \frac{1}{\nu} \right)}$$

en als er meer dan drie gemiddelden zijn

$$(17.14) \quad T_0 = \frac{\frac{v_{\max}}{s_{\bar{x}_j}} - \frac{6}{5} \log k'}{3 \left(\frac{1}{4} + \frac{1}{\nu} \right)}.$$

Hierin is weer $s_{\bar{x}_j} = s_b/\sqrt{n}$ en $\nu = k(n-1)$, terwijl in (17.14) k' het aantal gemiddelden in de groep is. Is $T_0 \geq T_{1-\frac{1}{2}\alpha}$, dan wordt het gemiddelde \bar{x}_a van de groep afgescheiden en in een afzonderlijke groep geplaatst.

De zojuist beschreven procedure wordt nu voor de resterende $k' - 1$ gemiddelden in de groep herhaald en als het meest extreme gemiddelde dan weer dient te worden afgescheiden, gaat men verder met de overblijvende $k' - 2$ gemiddelden, totdat een niet-significant verschil wordt gevonden. Worden op deze wijze twee opeenvolgende gemiddelden uit een groep verwijderd, dan vormen deze tezamen een nieuwe groep.

Tabel 17.5. Enkele waarden van $\frac{6}{5} \log k'$

k'	4	5	6	7	8	9
$\frac{6}{5} \log k'$	0,7225	0,8388	0,9338	1,0141	1,0837	1,1451
k'	10	11	12	13	14	15
$\frac{6}{5} \log k'$	1,2000	1,2497	1,2950	1,3367	1,3754	1,4113

17.1 GEMIDDELDEN VAN k POPULATIES: VARIANTIE-ANALYSE

Zijn het er echter drie of meer, dan wordt de procedure ook op deze nieuwe groep toegepast. Bij het gebruik van (17.14) kan men voor $k' = 4, 5, \dots, 15$ de waarden van $\frac{1}{2} \log k'$ in tabel 17.5 aflezen.

(c) Van elke groep van drie of meer gemiddelden, die bij de voorgaande stappen van de analyse tot stand is gekomen, wordt de homogeniteit onderzocht door te berekenen

$$(17.15) \quad F_0 [v_1 = m - 1, v_2 = k(n - 1)] = \frac{s_t^2}{s_b^2},$$

waarin s_t^2 de variantieschatting is, die met (17.2) verkregen kan worden uit de m gemiddelden in de groep en waarin s_b^2 ontleend wordt aan de volledige variantie-analyse op alle k steekproeven.

Voorbeeld

17.5. Wij passen de methode van TUKEY toe op de gemiddelden van de Fe -gehalten van de 10 analysten in tabel 17.2. Gemakshalve gebruiken wij bij deze analyse weer de hulpgrootte $u = 1000$ ($\alpha = 2,955$) en wij kiezen $\alpha = 0,05$. De opeenvolgende stappen verlopen als volgt:

(a) Uit tabel 17.4 blijkt, dat $s_b^2 = 93,42$, zodat $s^2_{\bar{x}_j} = 93,42/5 = 18,684$ en $s_{\bar{x}_j} = 4,32$. Uit tabel C volgt, dat $t_{1-\frac{1}{2}\alpha} = t_{0,975}$ met 40 vrijheidsgraden gelijk is aan 2,021, zodat volgens (17.12)

$$\bar{v} = (2,021)(4,32)\sqrt{2} = 12,3.$$

De gemiddelden per analyst zijn, naar grootte gerangschikt (zie tabel 17.3):

6	7	2	4	10	9	5	3	1	8
-18,2	-17,0	-5,0	-4,0	-3,4	-0,2	+0,4	+2,8	+22,4	+29,2

Aleen het verschil tussen de gemiddelden 3 en 1 is groter dan 12,3, zodat de eerste fase van de analyse een splitsing oplevert in twee groepen, respectievelijk bestaande uit de laagste acht en de hoogste twee gemiddelden. Daar de laatste groep slechts twee gemiddelden bevat, behoeft de tweede stap van de analyse slechts voor de eerste groep met $k' = 8$ gemiddelden te worden uitgevoerd.

(b) Wij beschouwen nu dus de groep van 8 gemiddelden met $\bar{x}_g = -5,6$. Gemiddelde 6 wijkt hiervan het meest af, zodat

$$v_{\max} = |-18,2 - (-5,6)| = 12,6.$$

Met (17.14) vinden wij

$$T_0 = \frac{12,6}{4,32} - \frac{6}{5} \log 8 = \frac{40}{33} \left(2,92 - 1,08 \right) = 2,23,$$

en daar $T_{1-\frac{1}{2}\alpha} = T_{0,975} = 1,96$ wordt gemiddelde 6 apart geplaatst.

Voor de overblijvende groep met 7 gemiddelden herhalen wij deze procedure. Voor deze groep is $\bar{x}_g = -3,8$ en $\bar{x}_a = -17,0$ (gemiddelde 7), zodat $v_{\max} = |-17,0 - (-3,8)| = 13,2$ en

$$T_0 = \frac{13,2 - \frac{6}{5} \log 7}{3 \left(\frac{1}{4} + \frac{1}{40} \right)} = \frac{40}{33} (3,06 - 1,01) = 2,48.$$

Ook gemiddelde 7 wordt dus apart geplaatst.

Voor de nu nog overblijvende 6 gemiddelden is $\bar{x}_g = -1,6$ en $\bar{x}_a = +2,8$ (gemiddelde 3), zodat $v_{\max} = |2,8 - (-1,6)| = 4,4$. Bij toepassing van (17.14) blijkt $T_0 = 0,11$ te zijn, zodat gemiddelde 3 niet voldoende afwijkt en fase twee wordt stopgezet. De twee gemiddelden (6 en 7), die hierbij zijn afgezonderd, volgen in grootte op elkaar, zodat zij een afzonderlijke groep vormen. Op dit punt zijn de tien analisten dus in de volgende groepen gesplitst:

Groep	Analyst	Gemiddelde (\bar{u}_j)
A	6, 7	-18,2 -17,0
B	2, 4, 10, 9, 5, 3	- 5,0 - 4,0 -3,4 -0,2 +0,4 +2,8
C	1, 8	+22,4 +29,2

(c) Daar de groepen A en C slechts twee gemiddelden bevatten, die reeds bij stap (a) niet significant verschillend zijn bevonden, behoeft hiermee verder niets te gebeuren. Van groep B onderzoeken wij nu nog de homogeniteit, door voor de 6 tot deze groep behorende gemiddelden s_i^2 te berekenen en te vergelijken met de waarde van s_b^2 , die bij de volledige analyse werd verkregen (tabel 17.4, $s_b^2 = 93,42$):

$$\begin{aligned} S_{\text{tussen}} &= \frac{(-25)^2 + (-20)^2 + \dots + (14)^2}{5} - \frac{(-25 - 20 + \dots + 14)^2}{6 \times 5} \\ &= \frac{1515}{5} - \frac{(-47)^2}{30} = 229,3667, \\ s_i^2 &= \frac{229,3667}{5} = 45,8733, \text{ zodat } F_0(5, 40) = \frac{45,87}{93,42} = 0,49. \end{aligned}$$

Deze F_0 is kleiner dan 1 en dus ver van significant. De gemiddelden in groep B vertonen een zeer bevredigende homogeniteit.

Hiermede is de analyse voltooid. Zij leidt tot de uitspraak, dat de tien analisten in de drie groepen A, B en C kunnen worden gesplitst, die onderling significant verschillen; de verschillen binnen elke groep kunnen echter insignificant worden geacht.

17.2. Enkelvoudige analyse: k aselechte steekproeven van verschillende omvang**17.2.1. UITVOERING VAN DE ANALYSE**

Als de steekproeven 1, 2, ..., k resp. de omvang $n_1, n_2, \dots, n_j, \dots, n_k$ bezitten gelden voor het uitvoeren van de variantie-analyse de volgende formules:

$$(17.16) \quad S_{\text{totaal}} = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{\Sigma n_j},$$

waarin $\Sigma n_j = n_1 + n_2 + \dots + n_k$,

$$(17.17) \quad S_{\text{tussen}} = \sum_{j=1}^k n_j (\bar{x}_j - \bar{x})^2 = \\ = \frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \dots + \frac{(\Sigma x_k)^2}{n_k} - \frac{(\Sigma x)^2}{\Sigma n_j}$$

en

$$(17.18) \quad S_{\text{binnen}} = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2 = \\ = \Sigma x^2 - \sum \frac{(\Sigma x_j)^2}{n_j} = S_{\text{totaal}} - S_{\text{tussen}}.$$

Verder is

$$(17.19) \quad s_i^2 = \frac{S_{\text{tussen}}}{k-1} \quad \text{en} \quad (17.20) \quad s_b^2 = \frac{S_{\text{binnen}}}{\Sigma n_j - k}.$$

Onder H_0 volgt dan de grootheid $F = s_i^2/s_b^2$ een F -verdeling met $\nu_1 = k - 1$ en $\nu_2 = \Sigma n_j - k$. Verder verloopt de toetsing als beschreven is in 17.1.3.

Voorbeeld

17.6. Men beschikt over de fosforgehalten (mg%) in het serum van drie aselekt gekozen groepen konijnen van verschillende omvang en wil de hypothese $H_0: \mu_1 = \mu_2 = \mu_3$ toetsen met $\alpha = 0,05$. In tabel 17.6 zijn deze gehalten (alle verminderd met 35 mg%) opgenomen.

In tabel 17.6 zijn eerst de steekproefvarianties s_j^2 berekend. Deze blijken zo weinig te verschillen, dat de onderstelling $\sigma_1 = \sigma_2 = \sigma_3$ gerechtvaardigd is. Verder is uit voorgaande onderzoeken gebleken, dat de fosforgehalten een verdeling bezitten, die enige scheefheid naar rechts vertoont. Zoals wij later zullen zien, vormt deze scheefheid echter geen belangrijk bezwaar bij het toepassen van de variantie-analyse. De berekeningen, die hiervoor nodig zijn staan eveneens in tabel 17.6. Daar wij reeds over de kwadratensommen van de steekproeven beschikken, die bij sommatie S_{binnen} opleveren, hebben wij genomen: $S_{\text{tussen}} = S_{\text{totaal}} - S_{\text{binnen}}$. Uit de verdere berekeningen

Tabel 17.6. Enkelvoudige variantie-analyse: toetsing van de hypothese $\mu_1 = \mu_2 = \mu_3$ op basis van drie aselecte steekproeven van verschillende omvang

	1		2		3		
	u_1	u_1^2	u_2	u_2^2	u_3	u_3^2	
	7,3	53,29	6,8	46,24	11,9	141,61	$n_1 = 7$
	1,3	1,69	3,6	12,96	3,0	9,00	$n_2 = 5$
	7,9	62,41	2,0	4,00	7,3	53,29	$n_3 = 6$
	2,3	5,29	-3,2	10,24	15,3	234,09	
	3,6	12,96	8,1	65,61	14,3	204,49	$\Sigma n_j = 18$
	9,6	92,16			11,3	127,69	
	-3,0	9,00					
Σu_j	29,0		17,3		63,1		109,4 = Σu
Σu_j^2	236,80		139,05		770,17		1146,02 = Σu^2
$(\Sigma u_j)^2/n_j$	120,1429		59,8580		663,6017		843,6026 = $\Sigma (\Sigma u_j)^2/n_j$
$\Sigma u_j^2 - (\Sigma u_j)^2/n_j$	116,6571		79,1920		106,5683		302,4174 = S_{binnen}
s_j^2	19,4429		19,7980		21,3137		

$S_{\text{totaal}} = \Sigma u^2 - (\Sigma u)^2/\Sigma n_j = 1146,02 - (109,4)^2/18 = 1146,02 - 664,9089 = 481,1111$

$S_{\text{tussen}} = S_{\text{totaal}} - S_{\text{binnen}} = 481,1111 - 302,4174 = 178,6937$

	S	v.g.	Var. schatting
Tussen de steekproeven	178,6937	2	89,35
Binnen de steekproeven	302,4174	15	20,16
Totaal	481,1111	17	

$F_0(2, 15) = \frac{89,35}{20,16} = 4,43$

$F_{0,05} = 3,68 \quad F_{0,075} = 4,77$

$0,05 > P_R > 0,025$

blijkt, dat de overschrijdingskans van de waargenomen waarde van \bar{F} tussen 0,025 en 0,05 ligt, zodat H_0 bij de gekozen 5% drempel wordt verworpen. Daar de steekproefgemiddelden resp. zijn $\bar{u}_1 = 4,1$, $\bar{u}_2 = 3,5$ en $\bar{u}_3 = 10,5$ (dus: $\bar{x}_1 = 39,1$, $\bar{x}_2 = 38,5$ en $\bar{x}_3 = 45,5$) is het zonder meer duidelijk, dat het hoge gemiddelde van steekproef 3 voor deze significantie verantwoordelijk is.

17.2.2. BETROUWBAARHEIDSGRENZEN VAN ÉÉN GEMIDDELDE

Ondersteld is, dat alle populaties normaal verdeeld zijn met spreiding σ . De grootheid

(17.21)
$$T = \frac{\bar{x}_j - \mu_j}{\sigma} \sqrt{n_j}$$

volgt dus volgens (13.4) de standaardnormale verdeling. Verder volgt volgens (13.6) de grootheid

(17.22)
$$\frac{(\Sigma n_j - k) s_b^2}{\sigma^2}$$

17.2 GEMIDDELDEN VAN k POPULATIES: VARIANTIE-ANALYSE

een χ^2 -verdeling met $\Sigma n_j - k$ vrijheidsgraden. Uit (17.21), (17.22) en (13.10) volgt dat de grootheid

$$(17.23) \quad t = \frac{\bar{x}_j - \mu_j}{s_b} \sqrt{n_j}$$

een STUDENT-verdeling volgt met $\Sigma n_j - k$ vrijheidsgraden. Bezitten alle steekproeven dezelfde omvang, zodat $n_1 = n_2 = \dots = n_k = n$, zoals in 17.1, dan is $\Sigma n_j = kn$ en $\nu = k(n - 1)$.

Op grond van een steekproefgemiddelde \bar{x}_j kan men nu het volgende $100(1 - \alpha)\%$ betrouwbaarheidsinterval van μ_j verkrijgen:

$$(17.24) \quad \bar{x}_j + t_{\frac{1}{2}\alpha} \frac{s_b}{\sqrt{n_j}} < \mu_j < \bar{x}_j + t_{1-\frac{1}{2}\alpha} \frac{s_b}{\sqrt{n_j}}$$

De waarde van $t_{1-\frac{1}{2}\alpha} = -t_{\frac{1}{2}\alpha}$ met $\Sigma n_j - k$ [of met $k(n - 1)$] vrijheidsgraden kan in tabel C worden opgezocht.

Voorbeeld

17.7. Bij voorbeeld 17.6 is van groep 1 het gemiddelde 39,1 en de uitgevoerde variantie-analyse heeft opgeleverd $s_b^2 = 20,16$ met 15 vrijheidsgraden. Uit tabel C volgt, dat $t_{0,975} = -t_{0,025}$ met $\nu = 15$ gelijk is aan 2,131, zodat het 95% betrouwbaarheidsinterval van het populatiegemiddelde van groep 1 bedraagt

$$39,1 - 2,131 \sqrt{\frac{20,16}{7}} < \mu_1 < 39,1 + 2,131 \sqrt{\frac{20,16}{7}}, \\ 35,5 < \mu_1 < 42,7.$$

17.2.3. DE ENKELVOUDIGE VARIANTIE-ANALYSE TOEGEPAST OP TWEE STEEKPROEVEN

Beschouwt men 2 aselechte steekproeven, dan kunnen de formules voor s_t^2 en voor s_b^2 als volgt worden herleid:

$$\begin{aligned} s_t^2 &= \frac{\sum_{j=1}^k n_j (\bar{x}_j - \bar{x})^2}{k - 1} = n_1 (\bar{x}_1 - \bar{x})^2 + n_2 (\bar{x}_2 - \bar{x})^2 \\ &= n_1 \left[\bar{x}_1 - \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2} \right]^2 + n_2 \left[\bar{x}_2 - \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2} \right]^2 \\ &= n_1 \left[\frac{n_1 \bar{x}_1 + n_2 \bar{x}_1 - n_1 \bar{x}_1 - n_2 \bar{x}_2}{n_1 + n_2} \right]^2 + n_2 \left[\frac{n_1 \bar{x}_2 + n_2 \bar{x}_2 - n_1 \bar{x}_1 - n_2 \bar{x}_2}{n_1 + n_2} \right]^2 \\ &= n_1 \left[\frac{n_2 (\bar{x}_1 - \bar{x}_2)}{n_1 + n_2} \right]^2 + n_2 \left[\frac{n_1 (\bar{x}_2 - \bar{x}_1)}{n_1 + n_2} \right]^2 \\ &= \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_1 - \bar{x}_2)^2, \end{aligned}$$

$$s_b^2 = \frac{\sum \sum (x_{ij} - \bar{x}_j)^2}{\sum n_j - k} = \frac{\sum (x_1 - \bar{x}_1)^2 + \sum (x_2 - \bar{x}_2)^2}{n_1 - 1 + n_2 - 1} = s^2,$$

zodat

$$F = \frac{s_t^2}{s_b^2} = \frac{(\bar{x}_1 - \bar{x}_2)^2}{s^2} \cdot \frac{n_1 n_2}{n_1 + n_2}.$$

Vergelijken wij deze uitkomst met die van de *t*-toets, formule (16.4), dan zien wij dat $F = t^2$ [vergelijk met 13.6, waarin reeds bleek dat $F_{1-\alpha}(v_1 = 1, v_2 = n_1 + n_2 - 2) = t_{1-\alpha}^2(v = n_1 + n_2 - 2)$]. De *t*-toets van STUDENT kan dus worden beschouwd als een speciaal geval van de enkelvoudige variantie-analyse, met $k = 2$.

17.3. Tweevoudige variantie-analyse op *k* verwante steekproeven

17.3.1. INLEIDING

In tabel 17.7 zijn de uitkomsten opgenomen van een proef, waarbij men zes geoefende rekenaars een bepaald type berekening op drie verschillende rekenmachines, A, B en C, liet uitvoeren (wij laten de gedetailleerde opzet van deze proef buiten beschouwing; hierbij is er echter voor gezorgd, dat de opeenvolgende berekeningen gelijkwaardig zijn en dat het oefeningselement is uitgeschakeld).

Tabel 17.7. Benodigde tijd (aantal seconden verminderd met 90) voor het uitvoeren van een berekening op de machines A, B en C door zes rekenaars

Rekenaars	Rekenmachine			Totaal	Rijge- middelde
	A	B	C		
1	32	36	13	81	27
2	22	30	17	69	23
3	22	32	9	63	21
4	25	22	10	57	19
5	33	22	5	60	20
6	34	38	12	84	28
Totaal	168	180	66	414	
Kolomgemiddelde	28	30	11		23

Wij hebben hier te doen met drie reeksen verwante waarnemingen en zoals in 10.4 is gebleken, kan de hypothese H_0 , dat de verschillen tussen de machines toevallig zijn, verdelingsvrij worden getoetst met de methode van FRIEDMAN (*m* rangschikkingen). Indien bepaalde onderstellingen worden gemaakt, kan men echter voor de analyse van deze gegevens, bestaande uit k ($= 3$) verwante steekproeven van de omvang n ($= 6$), ook gebruik maken van de variantie-analyse.

Bij het bespreken van deze zg. tweevoudige variantie-analyse maken

17.3 GEMIDDELDEN VAN k POPULATIES: VARIANTIE-ANALYSE

Tabel 17.8. Notatie bij de tweevoudige variantie-analyse

Elementen	Verwante steekproeven					Som	Gemiddelde	
	1	2	j			k
1	x_{11}	x_{12}	x_{1j}	x_{1k}	$\Sigma x_{1j} = T_{1.}$	$\bar{x}_{1.}$
2	x_{21}	x_{22}	x_{2j}	x_{2k}	$\Sigma x_{2j} = T_{2.}$	$\bar{x}_{2.}$
.
.
.
.
i	x_{i1}	x_{i2}	x_{ij}	x_{ik}	$\Sigma x_{ij} = T_{i.}$	$\bar{x}_{i.}$
.
.
.
n	x_{n1}	x_{n2}	x_{nj}	x_{nk}	$\Sigma x_{nj} = T_{n.}$	$\bar{x}_{n.}$
Som	Σx_{i1} $T_{.1}$	Σx_{i2} $T_{.2}$	Σx_{ij} $T_{.j}$	Σx_{ik} $T_{.k}$	T	
Gemiddelde	$\bar{x}_{.1}$	$\bar{x}_{.2}$	$\bar{x}_{.j}$	$\bar{x}_{.k}$		$\bar{x}_{..} = \bar{x}$

wij gebruik van de notatie, die in tabel 17.8 is opgenomen. Hierbij is elke waarneming van twee indices voorzien. De index i ($i = 1, 2, \dots, n$) geeft aan op welk element de waarneming betrekking heeft. De index j ($j = 1, 2, \dots, k$) duidt op de 'steekproef', waarin de waarneming zich bevindt. Voor tabel 17.7 geldt dus bv.: $x_{23} = 17$, $x_{61} = 34$, $x_{42} = 22$. Voor de rij- en kolomtotalen is ter vereenvoudiging het symbool T ingevoerd. Door een verschillende plaatsing van de punt in de aan T toegevoegde index kunnen zij van elkaar worden onderscheiden. Zo ziet men bv., dat $T_{2.}$ de som van de waarnemingen op de tweede rij is (= de som van de waarnemingen betreffende element 2) en dat $T_{.2}$ de som is van de waarnemingen in de tweede kolom (= de som van de waarnemingen in 'steekproef' 2). De som van alle kn waarnemingen kan dan worden aangeduid met $T_{..}$ of kortweg met T en het kwadraat van deze som met $(T)^2$. Voor de indices van de gemiddelden geldt dezelfde notatie. De som van de kwadraten van alle kn waarnemingen is $\sum_{j=1}^k \sum_{i=1}^n x_{ij}^2$, maar als geen misverstand mogelijk is schrijven wij hiervoor weer kortweg Σx^2 . Tenslotte kan men de formules, waarin de analyse kan worden samengevat, nog iets vereenvoudigen door te schrijven $(T)^2/kn = C$.

17.3.2. ONDERSTELLINGEN EN NULHYPOTHESEN

Wij onderstellen dat elke waarneming x_{ij} afkomstig is uit een

normaal verdeelde populatie met gemiddelde μ_{ij} en dat alle kn populaties dezelfde spreiding σ bezitten. Verder wordt aangenomen, dat een gemiddelde μ_{ij} als volgt is samengesteld ¹

$$(17.25) \quad \mu_{ij} = \mu + \xi_i + \eta_j, \text{ met}$$

$$(17.26) \quad \Sigma \xi_i = 0 \text{ en } \Sigma \eta_j = 0,$$

zodat μ het gemiddelde van alle populatiegemiddelden is.

Volgens dit model bestaat elk populatiegemiddelde dus uit (1) de component μ , die alle kn populaties gemeen hebben, (2) de component ξ_i , die de k populaties op de i -de rij gemeen hebben, en (3) de component η_j , die de n populaties in de j -de kolom gemeen hebben. Verder houdt dit model in, dat de componenten ξ_i en η_j optelbaar (additief) zijn, dus dat er tussen deze componenten geen wisselwerking (interactie) optreedt. Onder interactie verstaat men een systematische component, die zowel van de rijen als van de kolommen afhangt.

Voor het voorbeeld in tabel 17.7 met $k = 3$ en $n = 6$ zijn de volgens dit model opgebouwde populatiegemiddelden in tabel 17.9 opgenomen.

Tabel 17.9. Populatiegemiddelden van de waarnemingen in tabel 17.7

Rekenaar	Rekenmachines			Rijgemiddelden
	A = 1	B = 2	C = 3	
1	$\mu_{11} = \mu + \xi_1 + \eta_1$	$\mu_{12} = \mu + \xi_1 + \eta_2$	$\mu_{13} = \mu + \xi_1 + \eta_3$	$\mu_{1.}$
2	$\mu_{21} = \mu + \xi_2 + \eta_1$	$\mu_{22} = \mu + \xi_2 + \eta_2$	$\mu_{23} = \mu + \xi_2 + \eta_3$	$\mu_{2.}$
3	$\mu_{31} = \mu + \xi_3 + \eta_1$	$\mu_{32} = \mu + \xi_3 + \eta_2$	$\mu_{33} = \mu + \xi_3 + \eta_3$	$\mu_{3.}$
4	$\mu_{41} = \mu + \xi_4 + \eta_1$	$\mu_{42} = \mu + \xi_4 + \eta_2$	$\mu_{43} = \mu + \xi_4 + \eta_3$	$\mu_{4.}$
5	$\mu_{51} = \mu + \xi_5 + \eta_1$	$\mu_{52} = \mu + \xi_5 + \eta_2$	$\mu_{53} = \mu + \xi_5 + \eta_3$	$\mu_{5.}$
6	$\mu_{61} = \mu + \xi_6 + \eta_1$	$\mu_{62} = \mu + \xi_6 + \eta_2$	$\mu_{63} = \mu + \xi_6 + \eta_3$	$\mu_{6.}$
Kolomgemiddelden	$\mu_{.1}$	$\mu_{.2}$	$\mu_{.3}$	$\mu_{..} = \mu$

Dit model houdt in, dat het gemiddelde tijdsverschil tussen twee rekenaars (bv. tussen 1 en 2) bij elke machine hetzelfde is (nl. $\xi_1 - \xi_2$). Als dus rekenaar 1 met machine A per berekening gemiddeld 10 seconden langer werkt dan rekenaar 2, dan is hij ook bij de machines B en C gemiddeld 10 seconden langzamer dan rekenaar 2. En als machine C gemiddeld bv. 20 seconden sneller is dan machine B, dan geldt dit verschil voor elke rekenaar, dus zowel voor de man die langzaam als voor de man die snel werkt.

Het is duidelijk dat het gekozen model *geen* rekening houdt met de mogelijkheid, dat interactie optreedt, d.w.z. dat bv. een rekenaar die in het algemeen een laag tempo heeft, met een bepaalde machine

¹ ξ en η : de Griekse letters xi, resp. éta.

17.3 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

een hoog tempo weet te bereiken. Wij komen hierop echter in 17.4 terug.

Uitgaande van het nu gekozen model kan men in principe twee hypothesen toetsen:

$H_0(1)$: de rij-componenten zijn nul: $\xi_1 = \xi_2 = \dots = \xi_n = 0$,

$H_0(2)$: de kolom-componenten zijn nul: $\eta_1 = \eta_2 = \dots = \eta_k = 0$.

Voor het voorbeeld in tabel 17.7 toetst men met $H_0(2)$ de praktische hypothese, dat er geen verschillen tussen de machines zijn, of anders gezegd: dat de waargenomen verschillen tussen de met de machines bereikte tijden toevallig zijn. De hypothese $H_0(1)$ houdt in, dat alle rekenaars hetzelfde tempo bezitten. Daar men met verschillende rekenaars werkt kan men a priori verwachten, dat deze qua tempo zullen verschillen, zodat de toetsing van deze hypothese achterwege kan blijven. Het grote voordeel van deze opzet is nu juist, dat men bij de toetsing van $H_0(2)$ deze verschillen tussen de rekenaars heeft uitgeschakeld.

Opmerking: Voordat men met de variantie-analyse aanvangt dient onderzocht te worden of de varianties homogeen zijn. Dit kan geschieden, door de toets van BARTLETT of die van HARTLEY op de waargenomen varianties van de rijen en de kolommen toe te passen. Bij het voorbeeld in tabel 17.7 zijn de kolomvarianties resp. 31,6, 46,4 en 16,4 met $F_{\max} = 2,83$. Voor de rijvarianties geldt resp. 151, 43, 133, 63, 199 en 196, met $F_{\max} = 4,63$. Beide waarden van F_{\max} zijn ver van significant.

17.3.3. UITVOERING VAN DE TOETSINGEN

Onder $H_0(1)$ is $\xi_1 = \xi_2 = \dots = \xi_n = 0$, zodat dan voor de rij-gemiddelden geldt (ook als er verschillende kolom-componenten aanwezig zijn, zie tabel 17.9):

$$\mu_1 = \mu_2 = \dots = \mu_n = \mu + \sum_{j=1}^k \eta_j/k.$$

Als deze hypothese juist is zijn alle rij-gemiddelden \bar{x}_i , dus schattingen van deze parameter, zodat men een zuivere schatting van σ^2 kan verkrijgen door volgens (17.1) te nemen

$$(17.27) \quad s_{\text{rijen}}^2 = s_R^2 = \frac{k \sum_{i=1}^n (\bar{x}_i - \bar{x})^2}{n-1} = \frac{\frac{1}{k} \sum T_i^2 - C}{n-1}$$

(waarin $C = (\sum x)^2/kn$). Men noemt $k \sum_{i=1}^n (\bar{x}_i - \bar{x})^2$ de kwadraten-som tussen de rijen en wij schrijven hiervoor voortaan S_R .

Voor het voorbeeld vindt men met deze formule (zie tabel 17.7):

$$S_R = \frac{(81)^2 + (69)^2 + \dots + (84)^2}{3} - \frac{(414)^2}{(3)(6)} = 9732 - 9522 = 210,$$

zodat
$$s_R^2 = \frac{210}{5} = 42.$$

Onder H_0 (2) is $\eta_1 = \eta_2 = \dots = \eta_k = 0$, zodat dan voor de kolom-gemiddelden geldt (ook als er verschillende rij-componenten aanwezig zijn):

$$\mu_{.1} = \mu_{.2} = \dots = \mu_{.k} = \mu + \sum_{i=1}^n \xi_i/n.$$

Als deze hypothese juist is, zijn alle kolom-gemiddelden dus schattingen van deze parameter, zodat men een zuivere schatting van σ^2 kan verkrijgen door te nemen

$$(17.28) \quad s^2_{\text{kolommen}} = s_K^2 = \frac{n \sum_{j=1}^k (\bar{x}_{.j} - \bar{x})^2}{k-1} = \frac{\frac{1}{n} \sum T^2_{.j} - C}{k-1}.$$

Men noemt $n \sum_{j=1}^k (\bar{x}_{.j} - \bar{x})^2$ de kwadratensom tussen de kolommen en wij gebruiken hiervoor het symbool S_K .

Voor het voorbeeld geldt:

$$S_K = \frac{(168)^2 + (180)^2 + (66)^2}{6} - \frac{(414)^2}{(3)(6)} = 10830 - 9522 = 1308$$

en
$$s_K^2 = \frac{1308}{2} = 654.$$

Men kan nog een derde schatting van σ^2 verkrijgen en wel door uit te gaan van de variabiliteit die resteert, als men de invloed van de verschillen tussen de rij-gemiddelden en tussen de kolomgemiddelden heeft uitgeschakeld. Deze zg. *residuele variantie* of *restvariantie* volgt uit:

$$(17.29) \quad s_E^2 = \frac{\sum_{i=1}^n \sum_{j=1}^k (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2}{(kn-1) - (n-1) - (k-1)} \\ = \frac{\sum x^2 - \frac{1}{k} \sum T^2_{i.} - \frac{1}{n} \sum T^2_{.j} + C}{(k-1)(n-1)}$$

en wordt in 17.3.4 nader besproken. De kwadratensom in de teller van (17.29) noemt men de residuele kwadratensom, S_E .

De totale kwadratensom volgt uit

$$(17.30) \quad S_T = \sum x^2 - \frac{(\sum x)^2}{kn} = \sum x^2 - C,$$

17.3 GEMIDDELDEN VAN k POPULATIES: VARIANTIE-ANALYSE

en er bestaat, zoals men met de gegeven formules zelf kan nagaan, de volgende betrekking tussen de vier kwadratensommen en de bijbehorende aantallen vrijheidsgraden:

$$S_T = S_R + S_K + S_E$$

$$(kn-1) = (n-1) + (k-1) + (k-1)(n-1).$$

Men kan zodoende S_E verkrijgen, door S_T te berekenen en daarvan de som van S_R en S_K af te trekken. Voor het voorbeeld is $\sum x^2 = (32)^2 + (36)^2 + (13)^2 + (22)^2 + \dots + (12)^2 = 11302$, zodat $S_T = 11302 - 9522 = 1780$

en
$$S_E = 1780 - (210 + 1308) = 262.$$

Hieruit volgt:
$$s_E^2 = \frac{262}{10} = 26,2.$$

De formules waarmee de berekeningen bij deze tweevoudige analyse het eenvoudigst kunnen worden uitgevoerd zijn in tabel 17.10 samengevat.

Tabel 17.10. Formules voor de tweevoudige variantie-analyse op k verwante steekproeven van de omvang n

	Kwadratensom	Vrijheidsgraden
Tussen de rijen	$S_R = \frac{1}{k} \sum T^2_{.i} - C$	$n - 1$
Tussen de kolommen	$S_K = \frac{1}{n} \sum T^2_{.j} - C$	$k - 1$
Residueel	$S_E = S_T - (S_R + S_K)$	$(k - 1)(n - 1)$
Totaal	$S_T = \sum x^2 - C$	$kn - 1$

Voor het voorbeeld der rekenmachines heeft de variantie-analyse de in tabel 17.11 vermelde gegevens opgeleverd.

Tabel 17.11. Tweevoudige variantie-analyse, toegepast op de gegevens in tabel 17.7

	S	Vr. gr.	Variantieschatting
Tussen de rijen (rekenaars)	210	5	$s_R^2 = \frac{210}{5} = 42$
Tussen de kolommen (rekenmethoden)	1308	2	$s_K^2 = \frac{1308}{2} = 654$
Residueel	262	10	$s_E^2 = \frac{262}{10} = 26,2$
Totaal	1780	17	

Onder de hypothese $H_0(1): \xi_1 = \xi_2 = \dots = \xi_n = 0$ volgt de grootheid

$$(17.31) \quad \underline{F} = sR^2/sE^2$$

een F -verdeling met $\nu_1 = n - 1$ en $\nu_2 = (k - 1)(n - 1)$. De toetsing van deze hypothese kan dus worden uitgevoerd door $F_0 = sR^2/sE^2$ te berekenen en $H_0(1)$ bij een drempelwaarde α te verwerpen, als $P_R(F_0) < \alpha$ is, d.i. dus voor $F_0 \geq F_{1-\alpha}[\nu_1 = n-1, \nu_2 = (k-1)(n-1)]$.

Bij het voorbeeld is, zoals reeds werd opgemerkt, deze hypothese niet van betekenis. Merk echter op, dat $F_0 = 42/26,2 = 1,60$ en dat $F_{0,95}(5, 10) = 3,33$. De hypothese $H_0(1)$ kan hier dus bij een 5% drempel niet worden verworpen, wat erop wijst dat de verschillen tussen de rekenaars van betrekkelijk geringe betekenis zijn, of met andere woorden: dat zij een vrij homogene groep vormen.

Onafhankelijk van deze eerste hypothese kan men ook de hypothese $H_0(2): \eta_1 = \eta_2 = \dots = \eta_k$ toetsen. Indien deze hypothese juist is, volgt de grootheid

$$(17.32) \quad \underline{F} = sK^2/sE^2$$

een F -verdeling met $\nu_1 = k-1$ en $\nu_2 = (k-1)(n-1)$. Men berekent dus F_0 en besluit bij een drempelwaarde α tot het verwerpen van $H_0(2)$, wanneer $P_R(F_0) < \alpha$ is, d.i. voor $F_0 \geq F_{1-\alpha}[\nu_1 = k-1, \nu_2 = (k-1)(n-1)]$.

Het voorbeeld levert op: $F_0 = 654/26,2 = 24,96$, terwijl $F_{0,99}(2, 10) = 7,56$. De rechtse kans van F_0 is dus aanzienlijk kleiner dan 1% en de hypothese, dat er geen verschillen tussen de machines optreden wordt

Tabel 17.12. Hemoglobinegehalten op drie tijdstippen

Persoon nr.	Tijdstip			Totaal	Rijge- middelde
	1	2	3		
1	12,4	13,4	12,6	38,4	12,8
2	11,8	12,9	11,9	36,6	12,2
3	9,9	11,8	10,7	32,4	10,8
4	12,0	12,8	12,4	37,2	12,4
5	10,4	10,6	10,4	31,4	10,5
6	12,6	12,8	12,4	37,8	12,6
7	9,7	11,1	10,2	31,0	10,3
8	12,2	12,4	12,4	37,0	12,3
9	11,5	12,0	12,1	35,6	11,9
10	12,5	13,2	12,8	38,5	12,8
Totaal	115,0	123,0	117,9	355,9	
Kolom- gemiddelde	11,5	12,3	11,8		11,9

17.3 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

Berekeningen

$$\Sigma x^2 = (12,4)^2 + (13,4)^2 + \dots + (12,8)^2 = 4252,41$$

$$C = (355,9)^2 / (3) (10) = 4222,1603$$

$$S_T = 4252,41 - 4222,1603 = \dots \dots \dots 30,2497$$

$$S_R = \frac{(38,4)^2 + (36,6)^2 + \dots + (38,5)^2}{3} - 4222,1603$$

$$= 4247,3767 - 4222,1603 = \dots \dots \dots 25,2164$$

$$S_K = \frac{(115)^2 + (123)^2 + (117,9)^2}{10} - 4222,1603$$

$$= 4225,4410 - 4222,1603 = \dots \dots \dots 3,2807$$

$$S_E = 30,2497 - (25,2164 + 3,2807) = \dots \dots \dots 1,7526$$

verworpen. De conclusie luidt, dat de machines systematisch verschillen en in dit geval is zonder meer te zeggen, dat machine C te prefereren is.

Om de lezer met deze techniek vertrouwd te maken geven wij nog een tweede voorbeeld, waarvan zowel de oorspronkelijke gegevens als alle berekeningen in tabel 17.12 zijn opgenomen. Het betreft een onderzoek waarbij voor een steekproef van 10 vrouwen uit een bepaalde populatie het hemoglobinegehalte (in g%) op drie verschillende tijdstippen werd bepaald, t.w. (1) vlak voor, (2) één week na, en (3) één maand na het toepassen van een zekere therapie. Men wenst met een onbetrouwbaarheid van 5% de hypothese te toetsen, dat de verschillen tussen de tijdstippen toevallig zijn, met als alternatieve hypothese dat systematische verschillen optreden.

In tabel 17.13 zijn de uitkomsten van de variantie-analyse in de gijkte tabelvorm geplaatst. Wij onderstellen, dat het door de formules (17.25) en (17.26) gerepresenteerde model (met normaal verdeelde populaties, die dezelfde spreiding bezitten) toepasbaar is. Ook in dit geval is alleen de hypothese H_0 (2), dat de effecten η_j gelijk zijn aan nul, van belang. Wij berekenen dus $F_0 = s_K^2 / s_E^2 = 1,6404 / 0,0974 = 16,8$ en daar uit tabel F blijkt, dat $F_{0,95}(2, 18) = 3,55$ kan deze hypothese worden verworpen.

Tabel 17.13

	S	Vr. gr.	Var. schatting
Tussen de rijen	25,2164	9	25,2164/9 = 2,8018
Tussen de kolommen	3,2807	2	3,2807/2 = 1,6404
Residueel	1,7526	18	1,7526/18 = 0,0974
Totaal	30,2497	29	

17.3.4. DE RESIDUELE VARIANTIE

Teneinde de lezer een inzicht te geven in de aard van de residuele variantie, die uit formule (17.29) volgt, beschouwen wij nogmaals de uitkomsten van de proef met de drie rekenmachines in tabel 17.7. Deze zijn in tabel 17.14 opgenomen en als volgt bewerkt:

Tabel 17.14. Berekening van de residuele variantie uit de waarnemingen, nadat deze voor de waargenomen rij- en kolomcomponenten zijn gecorrigeerd (gegevens uit tabel 17.7)

A. Uitkomsten en de verschillen van de rij- en kolom-gemiddelden met het totaal-gemiddelde

Rekenaars	Rekenmachine			$T_{i.}$	$\bar{x}_{i.}$	$\bar{x}_{i.} - \bar{x}$
	A	B	C			
1	32	36	13	81	27	+4
2	22	30	17	69	23	0
3	22	32	9	63	21	-2
4	25	22	10	57	19	-4
5	33	22	5	60	20	-3
6	34	38	12	84	28	+5
$T_{.j}$	168	180	66	414		
$\bar{x}_{.j}$	28	30	11		23	
$\bar{x}_{.j} - \bar{x}$	+5	+7	-12			

B. Alle waarnemingen verminderd met $\bar{x}_{i.} - \bar{x}$

Rekenaars	A	B	C	$T_{i.}$	$\bar{x}_{i.}$	$\bar{x}_{i.} - \bar{x}$
1	28	32	9	69	23	0
2	22	30	17	69	23	0
3	24	34	11	69	23	0
4	29	26	14	69	23	0
5	36	25	8	69	23	0
6	29	33	7	69	23	0
$T_{.j}$	168	180	66	414		
$\bar{x}_{.j}$	28	30	11		23	
$\bar{x}_{.j} - \bar{x}$	+5	+7	-12			

C. Alle waarnemingen tevens verminderd met $\bar{x}_{.j} - \bar{x}$

Rekenaars	A	B	C	$T_{i.}$	$\bar{x}_{i.}$	$\bar{x}_{i.} - \bar{x}$
1	23	25	21	69	23	0
2	17	23	29	69	23	0
3	19	27	23	69	23	0
4	24	19	26	69	23	0
5	31	18	20	69	23	0
6	24	26	19	69	23	0
$T_{.j}$	138	138	138	414		
$\bar{x}_{.j}$	23	23	23		23	
$\bar{x}_{.j} - \bar{x}$	0	0	0			

17.3 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

A. De oorspronkelijke uitkomsten, met de rij- en kolomgemiddelden, zijn in deze tabel nogmaals opgenomen; per rij is echter ook de schatting van de rij-component ξ_i , dat is de waargenomen component $r_i = \bar{x}_i - \bar{x}$ vermeld, terwijl per kolom de schatting van de kolom-component η_j , dat is de waargenomen component $k_j = \bar{x}_j - \bar{x}$, is gegeven.

B. In deze tabel zijn de rijgemiddelden gelijk gemaakt, door van alle op rij i staande getallen $r_i = \bar{x}_i - \bar{x}$ af te trekken. De kolomgemiddelden en het totaalgemiddelde veranderen hierdoor niet.

C. Tenslotte zijn in deze tabel de kolomgemiddelden gelijk gemaakt, door van alle in kolom j staande getallen $k_j = \bar{x}_j - \bar{x}$ af te trekken. De rijgemiddelden en het totaalgemiddelde veranderen hierdoor niet.

De getallen, die na deze 'correctie's' zijn verkregen, zijn alle schattingen van de gemeenschappelijke component μ (zie tabel 17.9), zodat uit deze getallen een (derde) schatting van σ kan worden verkregen, die onafhankelijk is van s_R^2 en s_K^2 . Dit is de restvariantie, die volgt uit formule (17.29). De getallen in 17.14C zijn immers verkregen door te nemen

$$x'_{ij} = x_{ij} - r_i - k_j = x_{ij} - (\bar{x}_i - \bar{x}) - (\bar{x}_j - \bar{x}) .$$

Hun kwadratensom is

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^k (x'_{ij} - \bar{x})^2 &= \sum \sum (x'_{ij} - \bar{x})^2 = \sum \sum \left[\{x_{ij} - (\bar{x}_i - \bar{x}) - (\bar{x}_j - \bar{x})\} - \bar{x} \right]^2 \\ &= \sum \sum (x_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2 \end{aligned}$$

met als aantal vrijheidsgraden

$$\begin{aligned} kn - (n-1) - (k-1) - 1 &= kn - n + 1 - k + 1 - 1 \\ &= k(n-1) - (n-1) = (k-1)(n-1) . \end{aligned}$$

In tabel 17.14 D zijn de waarden van $x'_{ij} - \bar{x}$ voor het voorbeeld berekend. Uit deze tabel volgt, dat

$$S_E = \sum \sum (x'_{ij} - \bar{x})^2 = (0)^2 + (2)^2 + \dots + (-4)^2 = 262 .$$

Tabel 17.14 D. De waarden van $x'_{ij} - \bar{x}$

Rekenaars	A	B	C
1	0	+2	-2
2	-6	0	+6
3	-4	+4	0
4	+1	-4	+3
5	+8	-5	-3
6	+1	+3	-4

Deze uitkomst vonden wij reeds eerder (en veel eenvoudiger) door te nemen $S_E = S_T - (S_K + S_R)$. Toch is het van belang om de residuele

deviaties $x'_{ij} - \bar{x}$ te berekenen, omdat zij een zekere contrôle op de juistheid van het model verschaffen. Bij een klein aantal waarnemingen bestaat deze hieruit, dat men kan nagaan of de positieve en negatieve tekens toevallig over de tabel verdeeld zijn en dat men eventuele storende uitbijters kan signaleren. Bij een groter aantal waarnemingen kan men de frequentieverdeling van de residuele waarden x'_{ij} of van de residuele deviaties opstellen en onderzoeken, of deze ongeveer normaal is.

17.3.5. DE TWEEVOUDIGE VARIANTIE-ANALYSE TOEGEPAST OP GE-
PAARDE WAARNEMINGEN

De t -toets op verschillen tussen gepaarde waarnemingen in 15.3 kan beschouwd worden als een speciaal geval (met $k = 2$) van de in deze paragraaf beschreven tweevoudige variantie-analyse. Wij volstaan hier, met dit aan de hand van een voorbeeld te demonstreren. Wij hebben hiertoe op tien gepaarde waarnemingen eerst de variantie-analyse en vervolgens de t -toets van STUDENT toegepast:

	I	2	3	4	5	6	7	8	9	10	$k = 2$	$n = 10$
I: x_1	24	26	12	11	24	13	25	25	16	13	$\Sigma x_1 = 189$	$\Sigma x_1^2 = 3937$
II: x_2	33	12	27	22	28	24	23	26	35	23	$\Sigma x_2 = 253$	$\Sigma x_2^2 = 6765$
$x_1 + x_2$	57	38	39	33	52	37	48	51	51	36	$\Sigma x = 442$	$\Sigma x^2 = 10702$
$v = x_2 - x_1$	9	-14	15	11	4	11	-2	1	19	10	$\Sigma v = 64$	$\Sigma v^2 = 1226$

Variantie-analyse

$$S_{\text{totaal}} = \Sigma x^2 - (\Sigma x)^2 / kn = 10702 - (442)^2 / 20 = 10702 - 9768,2 = 933,8$$

$$S_{\text{rijen}} = \frac{(57)^2 + (38)^2 + \dots + (36)^2}{2} - \frac{(442)^2}{20} = 320,8$$

$$S_{\text{kolommen}} = \frac{(189)^2 + (253)^2}{10} - \frac{(442)^2}{20} = 204,8$$

525,6

$$S_{\text{residueel}} = \dots \dots \dots 408,2$$

	S	v. g.	Var.
Tussen de rijen	320,8	9	35,6444
Tussen de kolommen	204,8	1	204,8000
Residueel	408,2	9	45,3556
Totaal	933,8	19	

$$F_0 = \frac{204,8}{45,3556} = 4,5154$$

$$F_{0,95}(1, 9) = 5,12$$

17.4 GEMIDDELDEN VAN k POPULATIES: VARIANTIE-ANALYSE

t-toets op de verschillen

$$\bar{v} = \frac{\Sigma v}{n} = \frac{64}{10} = 6,4 \quad s_v^2 = \frac{\Sigma v^2 - (\Sigma v)^2/n}{n-1} = \frac{1226 - (64)^2/10}{9} = 90,7111$$

$$s_v = 9,5242$$

$$t_0 = \frac{\bar{v} \sqrt{n}}{s_v} = \frac{6,4 \sqrt{10}}{9,5242} = 2,1250 \quad t_{0,975} (\nu = 9) = 2,26$$

$t_0^2 = (2,1250)^2 = 4,5156 = F_{0,9}^2$ (het verschil in de vierde decimaal is een gevolg van afrondingsfoutjes in t_0).

17.4. Verwante steekproeven met replicaties

17.4.1. ONDERSTELLINGEN EN NULHYPOTHESEN

Bij het in 17.3 beschreven model voor k verwante steekproeven is aangenomen, dat de rij- en kolomcomponenten optelbaar zijn. Dit model is echter onjuist als er interactie optreedt, d.w.z. als er een systematische component λ_{ij} is¹, die zowel van de rijen als van de kolommen afhangt, zodat de rij- en kolomcomponenten *niet additief* zijn. Er dient dan een ander model te worden gekozen. Hierbij onderstellen wij, evenals in 17.3, dat elke waarneming x_{ij} afkomstig is uit een normaal verdeelde populatie met spreiding σ en dat het gemiddelde van zo'n populatie als volgt is samengesteld:

$$(17.33) \quad \mu_{ij} = \mu + \xi_i + \eta_j + \lambda_{ij} ,$$

$$(17.34) \quad \Sigma \xi_i = 0, \quad \Sigma \eta_j = 0, \quad \sum_{i=1}^n \lambda_{ij} = 0 \quad \text{en} \quad \sum_{j=1}^k \lambda_{ij} = 0 .$$

De juistheid van de onderstelling, dat de kn populaties dezelfde variantie bezitten, kan op basis van de kn waargenomen varianties met de toets van BARTLETT (voor $k \leq 12$ eenvoudiger met de toets van HARTLEY) worden onderzocht.

Indien men weet of vermoedt, dat dit model geldt kan men niet volstaan met een proefopzet met k verwante steekproeven. De residuele variantie s_E^2 , die men bij de in 17.3 beschreven analyse verkrijgt, is dan geen schatting van σ^2 , omdat hierin ook de variatie tengevolge van de interactie tot uitdrukking komt (men zegt dan, dat vermenging — Eng. 'confounding' — heeft plaatsgevonden). Deze moeilijkheid kan worden opgelost, door voor elk der kn populaties m onafhankelijke waarnemingen (z.g. herhalingen of *replicaties*) te verrichten. Bij deze proefopzet kan men dan niet alleen de hypothesen

$$H_0 (1): \xi_1 = \xi_2 = \dots = \xi_n = 0$$

en

$$H_0 (2): \eta_1 = \eta_2 = \dots = \eta_k = 0$$

¹ λ = de Griekse letter lambda.

toetsen, maar *tevens* de hypothese, dat er geen interactie aanwezig is:

$$H_0(3): \lambda_{11} = \lambda_{12} = \dots = \lambda_{ij} = \dots = \lambda_{nk} = 0.$$

17.4.2. BEREKENINGEN

Wij beschouwen nogmaals het onderzoek betreffende de drie rekenmachines (A, B en C), dat in 17.3 als voorbeeld werd gebruikt. Wij onderstellen nu, dat dit zo is uitgevoerd, dat per rekenaar drie replicaties aanwezig zijn, d.w.z. dat men per rekenaar over drie onderling onafhankelijke waarnemingen per machine beschikt. De (fictieve) uitkomsten van het op deze wijze uitgevoerde onderzoek zijn in tabel 17.15 opgenomen. Wij beschrijven eerst de uitvoering van de berekeningen, vervolgens de betekenis van de interactie-variantie en tenslotte de toetsingen.

Tabel 17.15. Benodigde tijd (aantal seconden minus 90) voor het uitvoeren van een bepaald soort berekening met de machines A, B, en C door zes rekenaars. Drie replicaties per rekenaar

Rekenaars	Rekenmachine									Rij-som
	A			B			C			
1	19	25	30	27	20	28	16	26	22	213
2	26	35	28	32	36	28	27	24	34	270
3	22	16	27	36	28	33	12	2	7	183
4	23	21	35	27	17	25	25	14	20	207
5	35	24	32	34	39	42	11	15	17	249
6	28	25	35	24	27	31	22	20	16	228
Kolomsom	486			534			330			1350

Voordat men met de variantie-analyse aanvangt dient te worden onderzocht of de onderstelling, dat de $kn = 3 \times 6 = 18$ populaties (A, 1), (B, 1), . . . , (C, 6) dezelfde variantie bezitten, houdbaar is. Men kan daartoe op de varianties van de 18 steekproefjes van de omvang $n = 3$ de toets van BARTLETT toepassen. Hoewel de uitkomst van deze toets, $M_0(17) = 3,513$ wegens de kleine steekproefomvang met reserve dient te worden gezien is zij zo ver van significant, dat deze onderstelling verantwoord kan worden geacht.

De variantie-analyse begint nu eveneens met de 18 steekproefjes van 3 waarnemingen. Hierop kan de enkelvoudige analyse van 17.1 worden toegepast en men vindt dan:

$$\Sigma \Sigma x = T = 1350, \quad \Sigma \Sigma x^2 = (19)^2 + (25)^2 + \dots + (16)^2 = 37300,$$

$$C = (T)^2 / knm = (1350)^2 / (3)(6)(3) = 33750,$$

zodat

$$S_{\text{totaal}} = 37300 - 33750 = 3550.$$

17.4 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

Verder is

$$\begin{aligned} \text{Stussen} &= \frac{(19+25+30)^2 + (27+20+28)^2 + \dots + (22+20+16)^2}{3} - \frac{(1350)^2}{(3)(6)(3)} \\ &= 36434,6667 - 33750 = 2684,6667 . \end{aligned}$$

De uitkomsten van deze enkelvoudige analyse zijn dus:

	Kwadratensom	Vr. gr.
Tussen de steekproeven	2684,6667	17
Binnen de steekproeven	865,3333	36
Totaal	3550,0000	53

Men kan nu in de tweede plaats de totalen per rekenaar per machine beschouwen, zoals deze in tabel 17.16 zijn samengevat.

Tabel 17.16. Totalen per rekenaar per machine

Rekenaar	Rekenmachine			Rijksom
	A	B	C	
1	74	75	64	213
2	89	96	85	270
3	65	97	21	183
4	79	69	59	207
5	91	115	43	249
6	88	82	58	228
Kolomsom	486	534	330	1350

In deze vorm lenen zich de uitkomsten voor de tweevoudige analyse, die in 17.3 is beschreven. Men dient echter bij de toepassing hiervan rekening te houden met het feit, dat elk getal in deze tabel op $m = 3$ waarnemingen berust, zodat elk rij-totaal op $km = 3 \times 3 = 9$ en elk kolomtotaal op $nm = 6 \times 3 = 18$ waarnemingen betrekking heeft. Men vindt dan:

$$\begin{aligned} \text{Stotaal} &= \frac{(74)^2 + (75)^2 + \dots + (58)^2}{3} - \frac{(1350)^2}{(3)(6)(3)} \\ &= 36434,6667 - 33750 = 2684,6667 . \end{aligned}$$

De totale kwadratensom bij deze tweede stap, de tweevoudige analyse, is dus gelijk aan de kwadratensom tussen de steekproeven bij de eerste stap, de enkelvoudige analyse.

Verder is:

$$S_{\text{rijen}} = \frac{(213)^2 + (270)^2 + \dots + (228)^2}{(3)(3)} - \frac{(1350)^2}{(3)(6)(3)}$$

$$= 34288 - 33750 = 538$$

en

$$S_{\text{kolommen}} = \frac{(486)^2 + (534)^2 + (330)^2}{(6)(3)} - \frac{(1350)^2}{(3)(6)(3)}$$

$$= 35014 - 33750 = 1264,$$

zodat

$$S_{\text{residueel}} = S_{\text{interactie}} = 2684,6667 - (538 + 1264) = 882,6667.$$

De volledige uitkomsten van de twee stappen zijn tenslotte in tabel 17.17 bijeengebracht.

Tabel 17.17. Uitkomsten van de variantie-analyse op de gegevens in tabel 17.15

	Kwadratensom	Vrijheidsgraden	Variantie-schatting
Tussen de rijen (rekenaars)	538,0000	$n - 1 = 5$	$107,60 = s_R^2$
Tussen de kolommen (machines)	1264,0000	$k - 1 = 2$	$632,00 = s_K^2$
Interactie (rekenaars \times machines)	882,6667	$(k - 1)(n - 1) = 10$	$88,27 = s_I^2$
Tussen de steekproeven (subtotaal)	2684,6667	$kn - 1 = 17$	
Binnen de steekproeven	865,3333	$kn(m - 1) = 36$	$24,04 = s_B^2$
Totaal	3550,0000	$knm - 1 = 53$	

Tabel 17.18. Formules voor de tweevoudige variantie-analyse: k verwante steekproeven van de omvang n met m replicaties

	Kwadratensom	Vrijheidsgraden
Tussen de rijen	$S_R = mk \sum_i (\bar{x}_{i.} - \bar{x})^2$	$n - 1$
Tussen de kolommen	$S_K = mn \sum_j (\bar{x}_{.j} - \bar{x})^2$	$k - 1$
Interactie $R \times K$	$S_I = m \sum_{ij} (\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2$	$(k - 1)(n - 1)$
Binnen de steekproeven	$S_B = \sum_{ijr} (x_{ijr} - \bar{x}_{ij})^2$	$kn(m - 1)$
Totaal	$S_T = \sum_{ijr} (x_{ijr} - \bar{x})^2$	$knm - 1$

De verschillende kwadratensommen en de corresponderende aantallen vrijheidsgraden zijn in tabel 17.18 in formulevorm gegeven. Wij hebben hier de herleide formules, waarmee de berekeningen het

17.4 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

best kunnen worden uitgevoerd, achterwege gelaten; de lezer kan deze zelf aan de hand van deze tabel, tabel 17.10 en de bij het voorbeeld uitgevoerde berekeningen afleiden. Ter vereenvoudiging zijn de volgende somtekens gebruikt:

$$\sum_i \text{ voor } \sum_{i=1}^n, \quad \sum_j \text{ voor } \sum_{j=1}^k, \quad \sum_{ij} \text{ voor } \sum_{i=1}^n \sum_{j=1}^k \text{ en } \sum_{ijr} \text{ voor } \sum_{i=1}^n \sum_{j=1}^k \sum_{r=1}^m.$$

17.4.3. DE INTERACTIE-VARIANTIE

In tabel 17.19A zijn de gemiddelden \bar{x}_{ij} van de $kn = 18$ steekproeven van de omvang $m = 3$ opgenomen (deze zijn verkregen, door de sommen in tabel 17.16 door drie te delen). Analoog aan de in 17.3.4 beschreven procedure kan men van elk gemiddelde eerst \bar{x}_i , $-\bar{x}$ en vervolgens $\bar{x}_j - \bar{x}$ aftrekken. Men verkrijgt dan de 'gecorrigeerde' gemiddelden

$$\bar{x}'_{ij} = \bar{x}_{ij} - (\bar{x}_i - \bar{x}) - (\bar{x}_j - \bar{x}),$$

die in tabel 17.19B staan (de berekeningen zijn niet opgenomen: desgewenst kan de lezer deze zelf uitvoeren. Om de afrondingsfout zo klein mogelijk te houden, zijn de gemiddelden in twee decimalen berekend).

Tabel 17.19. De gemiddelden \bar{x}_{ij} en \bar{x}'_{ij} van de waarnemingen in tabel 17.15

A. De gemiddelden \bar{x}_{ij}					
Rekenaars	Rekenmachine			\bar{x}_i	$\bar{x}_i - \bar{x}$
	A	B	C		
1	24,67	25 00	21,33	23,67	-1,33
2	29,67	32,00	28,33	30,00	+5,00
3	21,67	32,33	7,00	20,33	-4,67
4	26,33	23,00	19,67	23,00	-2,00
5	30,33	38,33	14,33	27,67	+2,67
6	29,33	27,33	19,33	25,33	+0,33
\bar{x}_j	27,00	29,67	18,33	25,00	
$\bar{x}_j - \bar{x}$	2,00	4,67	-6,67		0
B. De gemiddelden $\bar{x}'_{ij} = \bar{x}_{ij} - (\bar{x}_i - \bar{x}) - (\bar{x}_j - \bar{x})$					
1	24,00	21,67	29,33	25,00	
2	22,67	22,33	30,00	25,00	
3	24,33	32,33	18,33	25,00	
4	26,33	20,33	28,33	25,00	
5	25,67	31,00	18,33	25,00	
6	27,00	22,33	25,67	25,00	
	25,00	25,00	25,00	25,00	

Onder de hypothese $H_0(3)$, dat geen interactie aanwezig is, vertonen de (voor de gemiddelde rij- en kolomcomponenten gecorrigeerde) gemiddelden \bar{x}'_{ij} louter een toevallige variatie. Men beschikt dan over $kn = 18$ gemiddelden van steekproeven van de omvang $m = 3$ uit normaal verdeelde populaties met hetzelfde gemiddelde μ en met dezelfde spreiding σ . Deze gemiddelden zijn echter niet onderling onafhankelijk: zij bezitten slechts

$$kn - 1 - (n - 1) - (k - 1) = (k - 1)(n - 1)$$

vrijheidsgraden. Men kan dus, indien $H_0(3)$ juist is, een zuivere schatting van σ^2 verkrijgen door te nemen

$$(17.35) \quad S_I^2 = \frac{m \sum_{ij} (\bar{x}'_{ij} - \bar{x})^2}{(k-1)(n-1)} = \frac{m \sum_{ij} (\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2}{(k-1)(n-1)}.$$

Zoals in 17.4.2 is gebleken kan de interactie-kwadratensom veel eenvoudiger door aftrekking worden verkregen. Past men echter (17.35) toe op de gecorrigeerde gemiddelden in tabel 17.19B, dan vindt men

$$s_I^2 = \frac{3}{10} \left[(24,00)^2 + (21,67)^2 + \dots + (25,67)^2 - (450)^2/18 \right] = 87,97.$$

Het verschil met de uitkomst in tabel 17.17 is een gevolg van afrondingsfouten bij de berekening van \bar{x}'_{ij} .

Wanneer $H_0(3)$ onjuist is, zodat er wèl interactie aanwezig is, zal het effect daarvan in de gecorrigeerde gemiddelden tot uitdrukking komen. Deze gemiddelden stammen dan niet meer uit populaties met hetzelfde gemiddelde μ , zodat de variatie tussen deze gemiddelden systematisch te groot zal zijn.

17.4.4. DE VARIANTIE BINNEN DE STEEKPROEVEN

Deze variantie is eenvoudig te berekenen via $S_{\text{binnen}} = S_{\text{totaal}} - S_{\text{tussen}}$. Men kan natuurlijk ook gebruik maken van de formule in tabel 17.18, $S_{\text{binnen}} = \sum_{ijr} (x_{ijr} - \bar{x}_{ij})^2$, maar deze vereist veel meer rekenwerk en geeft tengevolge van afrondingsfouten gewoonlijk een iets afwijkende uitkomst. Toch heeft het zin de afzonderlijke waarden $x_{ijr} - \bar{x}_{ij}$ te bepalen, omdat zij een contrôle op de normaliteit mogelijk maken; de frequentieverdeling van deze waarden dient nl. ongeveer normaal te zijn.

17.4.5. UITVOERING VAN DE TOETSINGEN

Volgens het door (17.33) en (17.34) weergegeven model is elk gemiddelde μ_{ij} een constante. In tabel 17.20 is aangegeven, van welke parameters bij dit model de waargenomen varianties schattingen zijn.

De waargenomen variantie $s_B^2 = S_B/kn(m-1)$ is te allen tijde een zuivere schatting van σ^2 . Deze schatting is immers verkregen door kn varianties te combineren [zie formule (17.4)] en hierbij zijn a.h.w. alle verschillen tussen de gemiddelden uitgeschakeld.

17.4 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

Tabel 17.20. Geschatte parameters bij het door (17.33) en (17.34) weergegeven model

	Schatting	Geschatte parameter
Rijen	s_R^2	$\sigma^2 + \frac{mk \sum (\mu_{i.} - \mu)^2}{n - 1}$
Kolommen	s_K^2	$\sigma^2 + \frac{mn \sum (\mu_{.j} - \mu)^2}{k - 1}$
Interactie	s_I^2	$\sigma^2 + \frac{m \sum \sum (\mu_{ij} - \mu_{i.} - \mu_{.j} + \mu)^2}{(k - 1)(n - 1)}$
Binnen de steekproeven	s_B^2	σ^2

Onder de hypothese $H_0(3)$: $\lambda_{11} = \lambda_{12} = \dots = \lambda_{ij} = \dots = \lambda_{nk} = 0$ (geen interactie) is $\sum \sum (\mu_{ij} - \mu_{i.} - \mu_{.j} + \mu)^2 = 0$, zodat dan ook s_I^2 een zuivere schatting is van σ^2 . Als deze hypothese juist is, volgt dus de grootheid

$$\underline{F} = s_I^2/s_B^2$$

een F -verdeling met $\nu_1 = (k-1)(n-1)$ en $\nu_2 = kn(m-1)$. Is $H_0(3)$ echter onjuist, dan zal s_I^2 systematisch te grote waarden aannemen (zie ook 17.4.3) en de toetsing kan dus bij een drempelwaarde α worden uitgevoerd, door $F_0 = s_I^2/s_B^2$ te berekenen en $H_0(3)$ te verwerpen, wanneer $P_R(F_0) \leq \alpha$ is, d.i. voor $F_0 \geq F_{1-\alpha}(\nu_1, \nu_2)$.

Onder de hypothese $H_0(1)$: $\xi_1 = \xi_2 = \dots = \xi_n = 0$ is elke $\mu_{i.}$ gelijk aan μ , zodat ook $\sum (\mu_{i.} - \mu)^2 = 0$. Als deze hypothese juist is, dan is ook s_R^2 een zuivere schatting van σ^2 en volgt de grootheid

$$\underline{F} = s_R^2/s_B^2$$

een F -verdeling met $\nu_1 = n-1$ en $\nu_2 = kn(m-1)$. De toetsing kan dus worden uitgevoerd door $F_0 = s_R^2/s_B^2$ te berekenen en $H_0(1)$ te verwerpen, indien $P_R(F_0) \leq \alpha$ is.

Onder de hypothese $H_0(2)$: $\eta_1 = \eta_2 = \dots = \eta_k = 0$ is $\sum (\mu_{.j} - \mu)^2 = 0$, zodat s_K^2 een zuivere schatting van σ^2 is. Dan volgt de grootheid

$$\underline{F} = s_K^2/s_B^2$$

een F -verdeling met $\nu_1 = k-1$ en $\nu_2 = kn(m-1)$. De toetsing van deze hypothese kan dus geschieden, door $F_0 = s_K^2/s_B^2$ te berekenen en $H_0(2)$ te verwerpen, indien $P_R(F_0) \leq \alpha$ is.

Het tot dusver behandelde model en de daarop gebaseerde toetsingen bezitten echter t.a.v. de realiteit (i.c.: het onderzoek naar de verschillen tussen de drie rekenmachines) het nadeel, dat zij zich beperken tot de $n = 6$ rekenaars die aan het onderzoek deelnemen. Deze rekenaars vormen a.h.w. de populatie, waarop de conclusies van het onderzoek betrekking hebben. Gewoonlijk zal men echter willen komen tot een algemeen geldende uitspraak omtrent de in het onderzoek

betrokken machines, d.i. tot een uitspraak die geldt voor een willekeurig gekozen rekenaar. Dit houdt in, dat men de aan het onderzoek deelnemende rekenars beschouwt als een *steekproef* uit de populatie van alle rekenars, die men het bepaalde type berekening met deze machines zou kunnen laten uitvoeren. In dit geval dient men echter het model aan te vullen met enkele onderstellingen, die wij hier slechts beknopt bespreken.

De eerste onderstelling is, dat de n rekenars, die aan het onderzoek deelnemen, een aselechte steekproef vormen uit deze populatie van N rekenars en dat $n \ll N$ is. Verder onderstellen wij, dat de componenten ξ , die aan deze N rekenars verbonden zijn, normaal zijn verdeeld met gemiddelde 0 en spreiding σ_a en dat hun interactie-componenten λ (met de $k = 3$ machines) onafhankelijk hiervan normaal verdeeld zijn met gemiddelde 0 en spreiding σ_b . Wij vermelden nu zonder bewijs, dat de waargenomen varianties bij dit uitgebreide model schattingen zijn van de parameters, die in tabel 17.21 zijn gegeven.

Tabel 17.21. Geschatte parameters bij het uitgebreide model, waarbij de rekenars worden beschouwd als een aselechte steekproef uit een populatie van rekenars

	Schatting	Geschatte parameter
Rijen	s_R^2	$\sigma^2 + m\sigma_b^2 + mk\sigma_a^2$
Kolommen	s_K^2	$\sigma^2 + m\sigma_b^2 + \frac{mn \sum (u_{.j} - \mu)^2}{k-1}$
Interactie	s_I^2	$\sigma^2 + m\sigma_b^2$
Binnen de steekproeven	s_B^2	σ^2

De hypothese $H_0(3)$, dat geen interactie aanwezig is, houdt nu in dat $\sigma_b^2 = 0$. Uit tabel 17.21 blijkt, dat onder deze hypothese s_I^2 , evenals s_B^2 , een schatting is van σ^2 , zodat de toetsingsgrootheid

$$\underline{F} = s_I^2/s_B^2$$

een F -verdeling volgt met $\nu_1 = (k-1)(n-1)$ en $\nu_2 = kn(m-1)$. De toetsing van deze hypothese verloopt dus precies hetzelfde als bij het voorgaande model. Het verdere verloop van de analyse hangt echter van de uitkomst van deze toetsing af en hierbij zijn twee gevallen te onderscheiden:

a. $H_0(3)$ wordt verworpen. De conclusie luidt dan, dat er interactie aanwezig is, dus dat $\sigma_b^2 \neq 0$. Hieruit volgt, dat s_I^2 een schatting is van $\sigma^2 + m\sigma_b^2$.

Is $H_0(2)$ juist, dan is $\sum (\mu_{.j} - \mu)^2 = 0$, zodat dan s_K^2 eveneens een schatting is van $\sigma^2 + m\sigma_b^2$ (zie tabel 17.21). De toetsing van deze

hypothese dient dan plaats te vinden via de grootheid

$$\underline{F} = s_K^2/s_I^2$$

die onder H_0 (2) dus een F -verdeling volgt met

$$\nu_1 = k-1 \text{ en } \nu_2 = (k-1)(n-1).$$

Men berekent dus $F_0 = s_K^2/s_I^2$ en verwerpt H_0 (2) bij een drempelwaarde α , indien $P_R(F_0) \leq \alpha$ is.

Is H_0 (1) juist, dan is $\sigma_a^2 = 0$, d.w.z. voor *alle* rekenaars in de populatie (dus ook in de steekproef) is $\xi = 0$. Dan is ook s_K^2 een schatting van $\sigma^2 + m\sigma_b^2$ en de toetsingsgrootheid

$$\underline{F} = s_R^2/s_I^2$$

volgt een F -verdeling met $\nu_1 = n-1$ en $\nu_2 = (k-1)(n-1)$. Wil men deze hypothese toetsen, dan berekent men $F_0 = s_R^2/s_I^2$ en besluit tot het verwerpen van H_0 (1), als $P_R(F_0) \leq \alpha$ is.

b. H_0 (3) wordt niet verworpen. De conclusie is dan, dat $\sigma_b^2 = 0$, zodat zowel s_I^2 als s_B^2 schattingen zijn van σ^2 . Men pleegt in dit geval, om het aantal vrijheidsgraden te vergroten, deze twee schattingen te combineren door te nemen

$$s_C^2 = \frac{S_I + S_B}{(k-1)(n-1) + kn(m-1)}.$$

De hypothese H_0 (2) wordt dan getoetst via de grootheid $\underline{F} = s_K^2/s_C^2$ [met $\nu_1 = k-1$ en $\nu_2 = (k-1)(n-1) + kn(m-1)$] en de hypothese H_0 (1) via de grootheid

$$\underline{F} = s_R^2/s_C^2 \text{ [met } \nu_1 = n-1 \text{ en } \nu_2 = (k-1)(n-1) + kn(m-1)].$$

Opmerking: Tegen deze combinatie is als bezwaar aan te voeren, dat het niet-verwerpen van H_0 (3) niet identiek is aan het juist zijn van deze hypothese. Theoretisch is het dus juister, om zowel s_K^2 als s_R^2 met s_B^2 te vergelijken.

Wij passen nu deze toetsingsvoorschriften toe op het voorbeeld, wat dus impliceert, dat wij de zes rekenaars beschouwen als een aselechte steekproef. Als onbetrouwbaarheidsdrempel kiezen wij $\alpha=0,05$ en de variantieschattingen kunnen worden ontleend aan tabel 17.17.

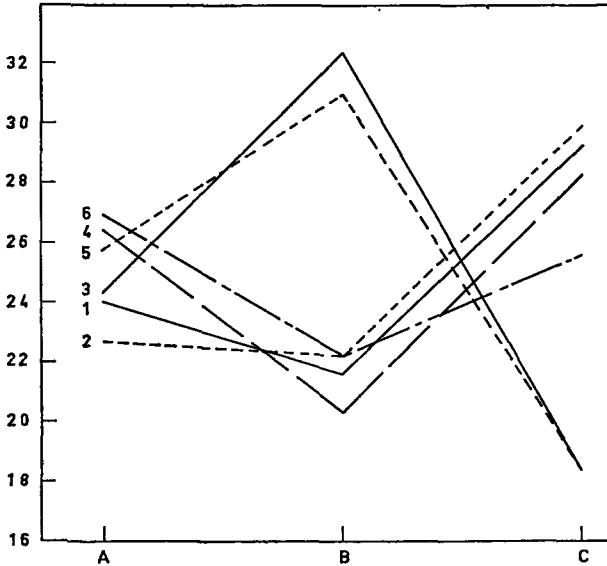
Bij toetsing van H_0 (3) — geen interactie — is:

$$F_0 = s_I^2/s_B^2 = 88,27/24,04 = 3,67.$$

Uit tabel F blijkt, dat $F_{0,05}(\nu_1 = 10, \nu_2 = 36)$ tussen 2,08 en 2,16 ligt, zodat H_0 (3) wordt verworpen. De verdere toetsing verloopt dus volgens (a) en hierbij is alleen de hypothese H_0 (2), dat de kolomcomponenten nul zijn, van betekenis. Wij berekenen dus

$$F_0 = s_K^2/s_I^2 = 632,00/88,27 = 7,16.$$

Daar $F_{0,95}(v_1 = 2, v_2 = 10) = 4,10$, wordt $H_0(2)$ eveneens verworpen. Merk op, dat men bij toetsing van de hypothese $H_0(1)$ vindt: $F_0 = s_R^2/s_I^2 = 107,60/88,27 = 1,22$, een ver van significante uitkomst. Dit wijst erop, dat geen belangrijke verschillen tussen de rekenaars aanwezig zijn, wat begrijpelijk is als men bedenkt dat de proef met goed geoefende personen werd uitgevoerd.



Figuur 17.1. De gemiddelden \bar{x}'_{ij} van de zes rekenaars. N.I.P.G. 878

In figuur 17.1 zijn de gecorrigeerde gemiddelden \bar{x}'_{ij} uit tabel 17.19B grafisch voorgesteld. Deze figuur laat zien, dat de geconstateerde interactie een gevolg is van het feit, dat de rekenaars 3 en 5 met machine B systematisch langere tijden produceren dan de overige rekenaars, en dat bij machine C het tegendeel het geval is. Het is duidelijk, dat een verschil tussen de gemiddelden van deze twee machines alleen dan *praktisch* van betekenis is, wanneer het relatief groot is t.o.v. de systematische verschillen, die bij elk van deze machines tussen de rekenaars optreden. Is dit niet het geval, dan loopt men het risico dat men later, bij het werkelijk uitvoeren van berekeningen, weliswaar over de gemiddeld het snelst te bedienen machine beschikt, maar dat deze wordt gebruikt door een rekenaar, die juist met deze machine minder goed overweg kan en daardoor langzamer werkt dan met een machine, die gemiddeld langzamer is, maar hem beter ligt. Zo is bv. bij het voorbeeld machine B gemiddeld langzamer dan machine A, maar rekenaar 4 werkt met B toch sneller dan met A.

17.5 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

17.5. Opgaven ¹

17.1. De onderstaande steekproeven stammen uit pseudo-normaal verdeelde populaties met dezelfde spreiding. *a.* Toets bij een 5% drempel de hypothese, dat de populatiegemiddelden gelijk zijn. *b.* Zet, als deze hypothese verworpen wordt, de analyse voort volgens de methode van TUKEY.

Steekproef	x					Σx	Steekproef	x					Σx
1	28	7	37	31	26	129	6	59	39	46	40	25	209
2	8	36	25	23	17	109	7	21	16	40	27	23	127
3	24	43	47	38	54	206	8	22	35	21	13	26	117
4	10	27	40	25	24	126	9	33	18	19	35	10	115
5	14	8	9	20	1	52	10	23	25	31	16	28	123

Gegeven: $\Sigma x = 1313$, $\Sigma x^2 = 42155$, $(\Sigma x)^2/50 = 34479,38$,

$$(129)^2 + (109)^2 + \dots + (123)^2 = 191391.$$

17.2. Toets met $\alpha = 0,05$ voor onderstaande steekproeven de hypothese $\mu_1 = \mu_2 = \dots = \mu_6$.

Steekproef	x							
1	16	13	9	10	5	8		
2	11	12	12	18	8			
3	15	14	16	9	10			
4	6	16	11	9	15	13	6	5
5	8	1	7	10	17	12	6	
6	16	4	5	4				

Gegeven: $\Sigma x = 357$, $\Sigma x^2 = 4295$, $(\Sigma x)^2/35 = 3641,40$.

17.3. Men beschikt over 7 aselechte steekproeven, elk bestaande uit 12 elementen, uit pseudo-normaal verdeelde populaties met dezelfde spreiding.

Gegeven is:

$$\bar{x}_1 = 22, s_1^2 = 15; \bar{x}_2 = 31, s_2^2 = 13; \bar{x}_3 = 18, s_3^2 = 16; \bar{x}_4 = 26, s_4^2 = 14; \bar{x}_5 = 23, s_5^2 = 17; \bar{x}_6 = 32, s_6^2 = 16; \bar{x}_7 = 22, s_7^2 = 12.$$

a. Toets de hypothese $\sigma_1 = \sigma_2 = \dots = \sigma_7$.

b. Toets de hypothese $\mu_1 = \mu_2 = \dots = \mu_7$ via de spreidingsbreedte.

Kies voor beide toetsingen $\alpha = 0,05$.

17.4. Pas de enkelvoudige variantie-analyse toe op de gegevens in opgave 10.20 (blz. 249).

¹ Overweeg bij elk van deze opgaven, welke onderstellingen bij de toepassing van de variantie-analyse dienen te worden gemaakt en of U deze verantwoord acht. Bedenk, dat soms de berekeningen aanzienlijk vereenvoudigd kunnen worden door van een hulpgrootheid gebruik te maken.

17.5. Een groep van 120 muizen wordt aselekt in drie subgroepen van 40 dieren gesplitst. Deze subgroepen worden geënt met drie typen typhus-bacteriën. De frequentieverdelingen van de levensduur der dieren zijn nevenstaand gegeven:

Levensduur in dagen	Type			Totaal
	9 D	11 C	DSC-1	
2	1	1	1	3
3	4	2	1	7
4	5	2	2	9
5	9	4	3	16
6	8	4	6	18
7	7	9	7	23
8	3	8	7	18
9	2	3	5	10
10	1	4	4	9
11		2	3	5
12		1	1	2
n	40	40	40	120
Σx	228	288	299	815
Σx^2	1430	2280	2439	6149
$\Sigma x^2 - (\Sigma x)^2/n$	130,4	206,4	203,975	540,775
\bar{x}	5,7	7,2	7,5	

Toets bij een 1% drempel de hypothese, dat de gemiddelden van de subgroepen toevallig verschillen.

17.6. Deze opgave is bedoeld om U vertrouwd te maken met de berekeningen, die bij de tweevoudige variantie-analyse moeten worden uitgevoerd. Pas deze analyse toe en bereken de residuele variantie ook volgens de methode in 17.3.4:

	Kolommen			Totaal	
	1	2	3		
Rijen	1	6	5	4	15
	2	15	10	8	33
	3	11	6	7	24
	4	16	14	12	42
	5	12	5	4	21
Totaal	60	40	35		135

17.7. Uit N.T.v.G., 101 (1957), 2278-2279.

De onderstaande tabel bevat de serum-ijzergehalten (in $\gamma/100$ ml) bij normale mannen op vier tijdstippen van dezelfde dag. Aan de hand van een grafiek van deze gegevens merkt men het volgende op: 'De dagcurve is bij onbelaste mannen onregelmatig door de sterke schommelingen en door de weinig constante vorm van de curve. De 'gemiddelde' lijn is representatief voor hetgeen bij de meerderheid der proefpersonen wordt gevonden, nl. de hoogste waarde om 12, en de laagste om 16 uur.' Geef Uw mening omtrent deze laatste uitspraak.

17.5 GEMIDDELDEN VAN K POPULATIES: VARIANTIE-ANALYSE

Nr.	9 uur	10 uur	12 uur	16 uur	Som
1	113	106	117	84	420
2	130	160	140	152	582
3	240	212	207	144	803
4	138	166	162	152	618
5	130	82	96	90	398
6	129	119	194	112	554
7	144	144	146	110	544
8	97	123	113	129	462
9	141	152	166	152	611
10	141	98	84	82	405
11	62	46	108	90	306
Som	1465	1408	1533	1297	5703

17.8. Een laboratorium ontvangt een zending van 50 muizen, die per tien stuks opgeborgen zijn in een kooi (de kooien zijn door de afzender genummerd van 1 t/m 5). Men wil deze dieren gebruiken voor een toxiciteitsonderzoek, waarbij van de te onderzoeken stof vijf verschillende doseringen zullen worden gebruikt en is van mening, dat men dit kan doen door per dosis één kooi met dieren te gebruiken (voor de laagste dosis kooi 1, voor de daaropvolgende kooi 2, enz.). Onderstel, dat men U vraagt deze proefopzet te beoordelen en U daarbij de beschikking geeft over de volgende, door de afzender bijgevoegde, gewichtsopgave (in grammen) van de muizen:

Kooi	Gewicht									
1	24	26	30	29	20	27	18	20	21	19
2	27	24	25	28	24	20	21	21	22	18
3	25	23	24	23	22	17	22	20	16	15
4	20	19	20	19	21	17	19	16	15	16
5	21	19	21	22	19	18	16	18	16	17

Geef zo nodig aan, hoe de proefopzet verbeterd kan worden.

Opmerking: Een werkelijk experiment op dit gebied is kortgeleden uitgevoerd door RÜMKE (191).

17.9. Bij een onderzoek betreffende bepalingen van het hemoglobinegehalte met de sicca-hemometer werden door drie geoefende analisten bij vier verschillende gehalten vijf onafhankelijke bepalingen verricht, waarvan de uitkomsten in onderstaande tabel zijn opgenomen. De analisten waren te voren niet ingelicht omtrent de gekozen gehalten en het aantal bepalingen per gehalte. De preparaten werden aangeboden in aselechte volgorde. De bepalingen werden verricht met hetzelfde apparaat, onmiddellijk na elkaar, zonder contact tussen de analisten onderling. De uitkomst van elke bepaling werd door de analisten op een briefje genoteerd, dat daarna onmiddellijk ter registratie werd weggenomen.

- Is er een systematisch verschil tussen de analisten te constateren? (Kies $\alpha = 0,01$).
- Welke schatting, die bij deze analyse wordt verkregen, is een maatstaf voor de betrouwbaarheid van de bepaling?

Gehalte	Analyst			Gehalte	Analyst		
	A	B	C		A	B	C
I	16,3	16,0	17,0	III	12,6	12,1	12,4
	16,5	16,7	17,0		11,9	11,9	12,4
	16,5	16,2	16,5		12,6	12,4	12,4
	16,5	16,5	16,7		12,4	11,9	12,4
	16,7	16,5	16,7		12,4	12,1	12,1
	<u>82,5</u>	<u>81,9</u>	<u>83,9</u>		<u>61,9</u>	<u>60,4</u>	<u>61,7</u>
II	14,4	14,1	14,2	IV	11,1	11,0	11,3
	14,1	13,7	14,1		11,3	10,8	11,6
	14,1	13,9	13,7		11,1	10,3	11,4
	14,1	13,6	14,1		10,8	10,5	11,1
	14,1	13,9	14,1		11,0	11,0	11,3
	<u>70,8</u>	<u>69,2</u>	<u>70,2</u>		<u>55,3</u>	<u>53,6</u>	<u>56,7</u>

REGRESSIE EN CORRELATIE

18.1. Tweedimensionale verdelingen

De statistische technieken in de hoofdstukken 14 t/m 17 betreffen waarnemingen van één (pseudo) normaal verdeelde grootheid, die vericht zijn aan de elementen in één of meer steekproeven. De populatieverdelingen, die hierbij worden bestudeerd, hebben dus betrekking op één grootheid x en zij worden daarom *ééndimensionale* verdelingen genoemd.

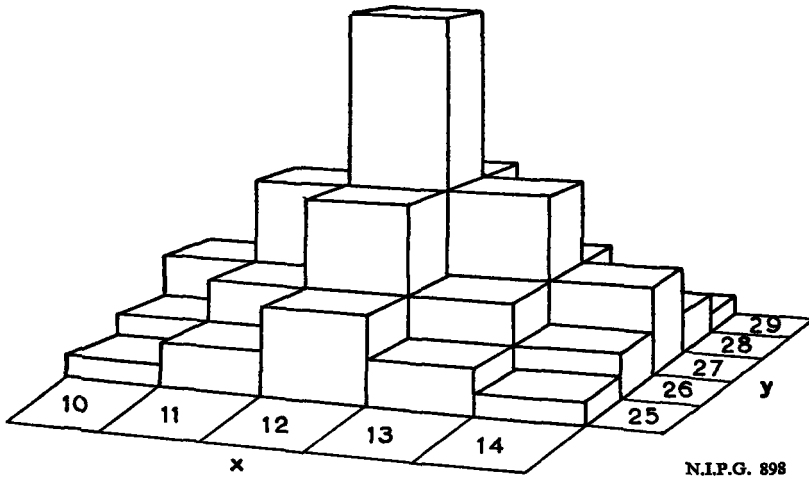
De hoofdstukken 18, 19 en 20 handelen over de klassieke methoden die kunnen worden toegepast, als men de simultane verdeling van twee grootheden x en y op een populatie wil bestuderen op basis van n waarnemingsparen (x, y) . Men noemt zo'n simultane verdeling ook wel een *tweedimensionale* verdeling.

Wij hebben gezien, dat een ééndimensionale verdeling grafisch kan worden voorgesteld in twee dimensies, bv. door middel van een kolommendiagram, dat in het platte vlak wordt getekend. Om een vergelijkbare grafiek te krijgen van een tweedimensionale populatieverdeling dient men een driedimensionaal diagram samen te stellen. Er zijn immers twee dimensies nodig om de variabelen x en y uit te zetten en een derde om de frequenties weer te geven. Beschouw bv. de tweedimensionale populatieverdeling in tabel 18.1.

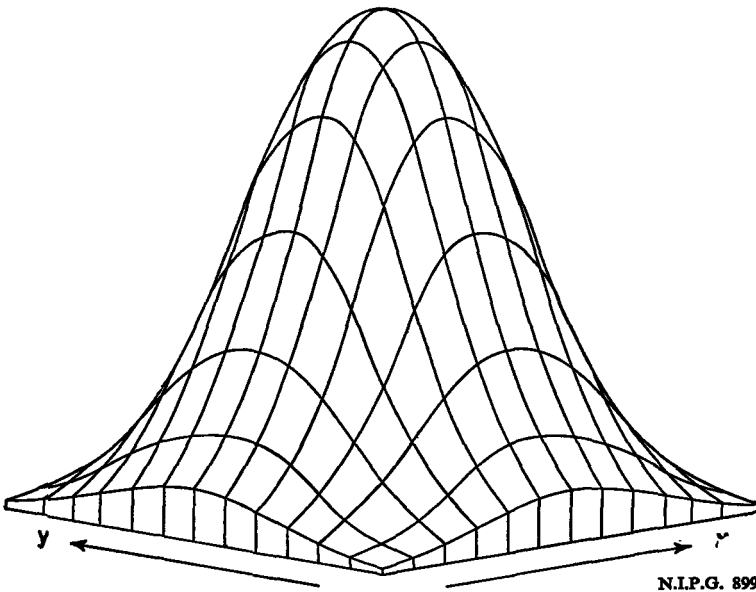
Tabel 18.1. Tweedimensionaal verdeelde populatie: de simultane verdeling van de grootheden x en y

$\begin{array}{c} x \\ \diagdown \\ y \end{array}$	10	11	12	13	14	Totaal
29	1	2	4	2	1	10
28	2	4	8	4	2	20
27	4	8	16	8	4	40
26	2	4	8	4	2	20
25	1	2	4	2	1	10
Totaal	10	20	40	20	10	100 = N

Van deze verdeling kan een kolommendiagram worden vervaardigd, door het platte vlak met behulp van lijnen evenwijdig aan de X - en Y -as in rechthoeken van dezelfde grootte te verdelen en op elke rechthoek een blok te plaatsen, waarvan de hoogte (en dus de inhoud) evenredig is aan de bij deze rechthoek behorende frequentie. Men



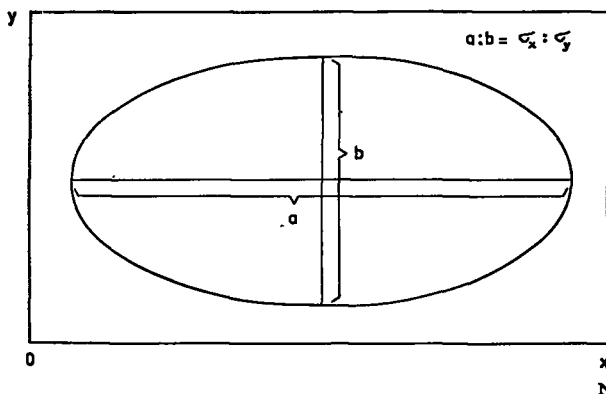
Figuur 18.1. Stereogram van de verdeling in tabel 18.1. Uit: Bok (41).



Figuur 18.2. Stereogram van een tweedimensionaal normale verdeling; de grootheden x en y zijn onafhankelijk en $\sigma_x = \sigma_y$. Uit: Bok (41).

noemt zo'n grafiek een *stereogram* en van de verdeling in tabel 18.1 is het stereogram in figuur 18.1 getekend.

Onderstel nu, dat men de simultane verdeling beschouwt van twee normaal verdeelde grootheden x en y . Tekent men van deze z.g. tweedimensionaal normale verdeling een grafische voorstelling, dan verkrijgt men het stereogram dat (afgeknot) in figuur 18.2 is weergegeven. In deze figuur is aangenomen, dat $\sigma_x = \sigma_y$, en tevens, dat de grootheden x en y onderling onafhankelijk zijn. In dit speciale geval is elke doorsnede evenwijdig aan het platte vlak een cirkel. Zijn de grootheden onafhankelijk, maar is $\sigma_x \neq \sigma_y$, dan is zo'n doorsnede een ellips, waarvan de assen evenwijdig aan de X - en Y -as lopen en zich verhouden als deze spreidingen: zie figuur 18.3.



N.I.P.G 901

Figuur 18.3. Horizontale doorsnede van het stereogram van een tweedimensionaal normale verdeling; de grootheden x en y zijn onafhankelijk, maar σ_x is groter dan σ_y .

Bij een onderzoek betreffende een tweedimensionale populatieverdeling (van de grootheden x en y) kan men onderscheid maken tussen regressieproblemen en correlatieproblemen. Wij leiden de regressieproblemen in 18.2 in en bespreken de *regressie-analyse* in hoofdstuk 19. De correlatieproblemen worden in 18.3 ingeleid en de *correlatierekening* wordt in hoofdstuk 20 besproken.

18.2. Regressie

18.2.1. REGRESSIEPROBLEMEN

Men spreekt van een regressieprobleem, wanneer men voor een tweedimensionale populatie wil onderzoeken, welk verloop de grootheid y vertoont bij verschillende waarden van de grootheid x . Men zal dan een steekproef, bestaande uit n waarnemingsparen (x, y) op de volgende wijze verkrijgen:

a. Men kiest een aantal waarden $x_1, x_2, \dots, x_j, \dots, x_k$ van de grootte x .

b. Men trekt uit de populatie aselekt een steekproef van in totaal n elementen, die bestaat uit: n_1 elementen met de waarde x_1 , n_2 elementen met de waarde x_2, \dots, n_k elementen met de waarde x_k . Zoals later zal blijken verdient het gewoonlijk aanbeveling $n_1 = n_2 = \dots = n_k$ te kiezen (soms is $n_1 = n_2 = \dots = n_k = 1$).

c. Van elk van deze elementen bepaalt men de waarde, die de grootte y heeft aangenomen.

Regressieproblemen kunnen bv. optreden bij een onderzoek naar:

1. De relatie tussen de uitkomsten van twee meettechnieken of analysemethoden, wanneer bekend is dat één daarvan zeer nauwkeurige uitkomsten oplevert.

2. Het verloop van de uitkomsten van een meettechniek of analysemethode, wanneer men de waarden x_j nauwkeurig kent (bv. indien men over een aantal preparaten met nauwkeurig bekende concentratie van een stof beschikt).

3. Het verloop van een grootte (y) in de tijd (x), wanneer men op bepaalde tijdstippen x_j over één of meer waarden van y beschikt.

4. Het verloop van de responsie (y) bij verschillende doses (x_j) van een stof, indien elke dosis aan een aselekt gekozen groep proefdiere is toegediend.

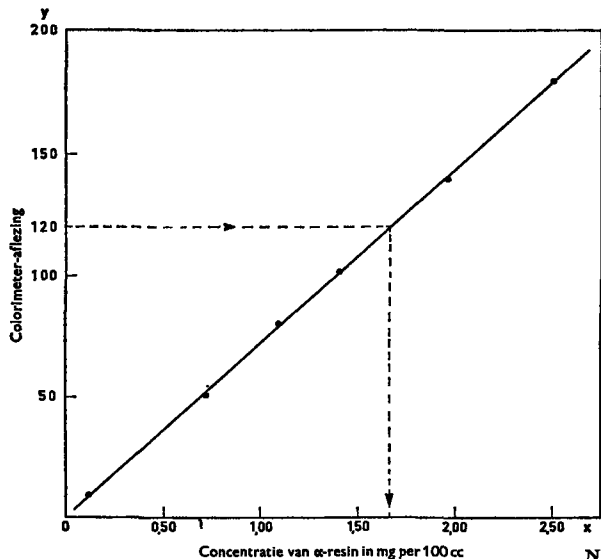
Voorbeelden

18.1. De directe kwantitatieve bepaling van de α -resin concentratie is zeer tijdrovend. BULLIS en ALDERTON verrichtten bij een zestal oplossingen van α -resin, waarvan de concentratie bekend was, colorimetrische bepalingen en hun onderzoek leverde de in tabel 18.2 opgenomen waarnemingsparen op, waarvan in figuur 18.4 het spreidingsdiagram is getekend.

Tabel 18.2

Concentratie in mg per 100 cc x	Colorimeter- aflezing y
0,12	8
0,71	50
1,09	81
1,38	102
1,95	140
2,50	181

Uit: WILKS (37).



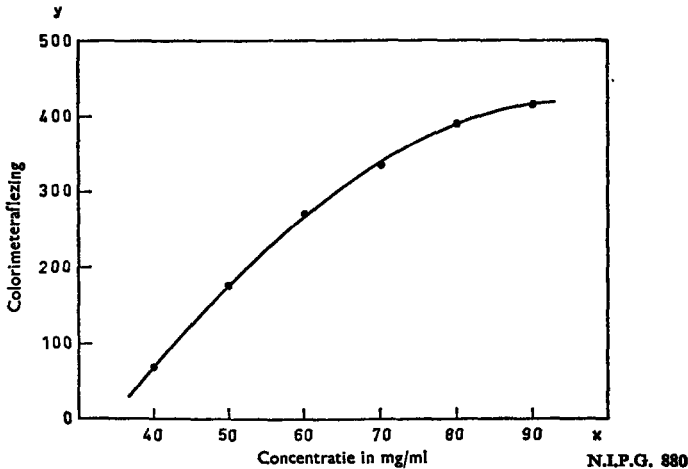
Figuur 18.4. Spreidingsdiagram van de waarnemingsparen in tabel 18.2 met de op het oog getrokken best-passende rechte.

Uit het spreidingsdiagram blijkt dat de punten (x, y) vrijwel op een rechte lijn door de oorsprong liggen en deze lijn kan op het oog reeds heel goed worden getekend. De vergelijking van deze lijn, $y = 72,5 x$, kan uit de figuur worden afgelezen. Met deze zg. *regressielijn* van y op x kan men dus voor een oplossing met bekende concentratie de colorimetrische waarde vrij nauwkeurig schatten. Omdat deze lijn zo goed bij de punten aansluit, kan men haar echter ook gebruiken om uit een colorimeteraflezing de concentratie van α -resin te schatten. De vergelijking wordt dan $x = y/72,5 = 0,0138 y$ en op deze wijze vindt men bv. bij een colorimeterwaarde van 120 een concentratie van (120) $(0,0138) = 1,66$ mg per 100 cc. De praktische betekenis hiervan is, dat men dus voor een oplossing met onbekende concentratie (in het bestudeerde gebied van 0 tot ongeveer 3 mg per 100 cc) de eenvoudige colorimetrische bepaling in de plaats kan stellen van de moeizame analyse.

Tabel 18.3

Concentratie (mg/ml)	Colorimeter-aflezing
x	y
40	69
50	175
60	272
70	335
80	390
90	415

18.2. De waarnemingsparen in tabel 18.3 zijn verkregen bij een onderzoek ter bestudering van de relatie tussen de hoeveelheid β -erythroidine in water en de troebelheid van de oplossing, bepaald met een colorimeter. Figuur 18.5 bevat het spreidingsdiagram van deze waarnemingsparen.



Figuur 18.5. Spreidingsdiagram van de waarnemingsparen in tabel 18.3 met de op het oog getekende best-passende kromme.

De regressie is hier weliswaar niet rechtlijnig, maar men kan toch op het oog een vloeiend verloopende kromme door de punten (x, y) tekenen, die zeer bevredigend past.

Bij de voorgaande voorbeelden kan met één waarneming van y (één colorimeteraflezing) per waarde van x (concentratie) worden volstaan, omdat er bij het verrichten van meerdere waarnemingen aan dezelfde nauwkeurig samengestelde concentratie een zo geringe spreiding tussen de uitkomsten blijkt te bestaan, dat deze zonder bezwaar kan worden verwaarloosd. Bij vele regressieproblemen is dit niet het geval. Er treedt dan, bv. tengevolge van de onnauwkeurigheid van de waarnemingen van y of door een natuurlijke variatie van de grootte y , een zodanige spreiding op, dat het niet meer mogelijk of verantwoord is de regressielijn op het oog te tekenen. Men dient dan te beschikken over een objectieve methode, waarmee de best-passende regressielijn of -kromme kan worden *berekend* en die tevens de mogelijkheid biedt om hypothesen betreffende de regressie te toetsen of betrouwbaarheids-grenzen van de verkregen lijn aan te geven. Men kan daartoe gebruik maken van de *methode der kleinste kwadraten*, die in de volgende punten worden besproken. Zoals reeds is opgemerkt komen de technieken der statistische regressie-analyse, die op deze methode berusten, in hoofd-

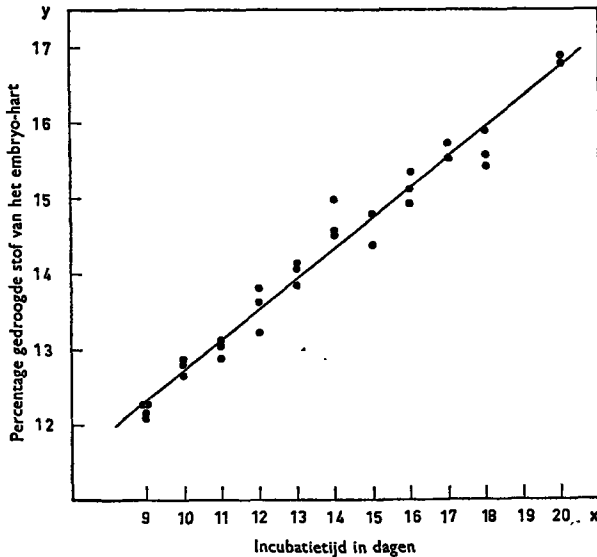
stuk 19 aan de orde en wij zullen daarbij onder meer gebruik maken van de volgende voorbeelden:

18.3. De gegevens in tabel 18.4 en figuur 18.6 betreffen de percentages droge stof (y) in het hart van kippeëmbryo's bij incubatietijden (x) van 9 t/m 20 dagen.

Tabel 18.4

x	y	\bar{y}	x	y	\bar{y}	x	y	\bar{y}
9	12,10	12,20	12	13,22	13,56	16	14,94	15,14
	12,16			13,64			15,12	
	12,26			13,81			15,37	
	12,26			40,67			45,43	
	48,78							
10	12,67	12,78	13	13,86	14,02	17	15,63	15,68
	12,80			14,05			15,72	
	12,87			14,15			31,35	
	38,34							
11	12,89	13,02	14	14,52	14,68	18	15,41	15,63
	13,05			14,54			15,57	
	13,12		14,99	46,88				
	39,06		44,05					
			15	14,37	14,59	20	16,79	16,83
		14,80		16,87				
				29,17			33,66	

Uit: J. W. DuvFF, Acta Neerl. Morph. A (1939), 165—169.

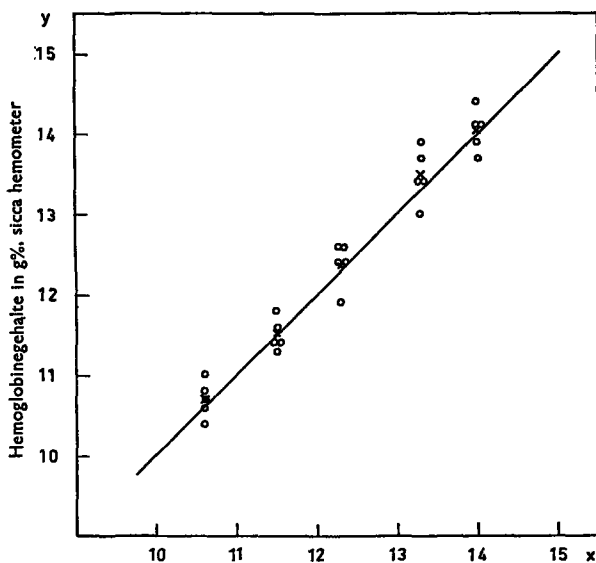


Figuur 18.6. Spreidingsdiagram van de gegevens in tabel 18.4 met de lineaire regressielijn volgens de vergelijking $y = 8,66 + 0,405x$ (zie ook voorbeeld 19.2).

18.4. Bij een onderzoek aangaande de bepaling van het hemoglobinegehalte met de sicca-hemometer werden door een geoefend analyst bij elk van vijf gehalten, die nauwkeurig bekend waren, vijf onderling onafhankelijke bepalingen met de sicca-hemometer verricht. De uitkomsten staan in tabel 18.5 en het spreidingsdiagram is in figuur 18.7 getekend.

Tabel 18.5

Gehalte x	sicca-bepalingen y					Σy_i	\bar{y}_i
	14,0	14,4	13,7	14,1	13,9		
13,3	13,4	13,7	13,9	13,4	13,0	67,4	13,48
12,3	12,6	11,9	12,6	12,4	12,4	61,9	12,38
11,5	11,8	11,6	11,4	11,4	11,3	57,5	11,50
10,6	10,7	11,0	10,8	10,4	10,6	53,5	10,70



Hemoglobinegehalte in g%, Beckman spectrofotometer

N.I.P.G. 885

Figuur 18.7. Spreidingsdiagram van de gegevens in tabel 18.5 met de 'ideale' regressielijn $y = x$. De waargenomen gemiddelden der y -waarden zijn met een kruisje aangegeven.

18.5. De gegevens in tabel 18.6 betreffen een proef met vitamine B_1 . Een twintigtal ratten met bradycardie werd aselekt in vier groepen van vijf ratten gesplitst. Elke groep kreeg eenmaal een bepaalde dosis vitamine B_1 toegediend, waarna de duur van de afwezigheid van de bradycardie (in halve dagen) werd geregistreerd.

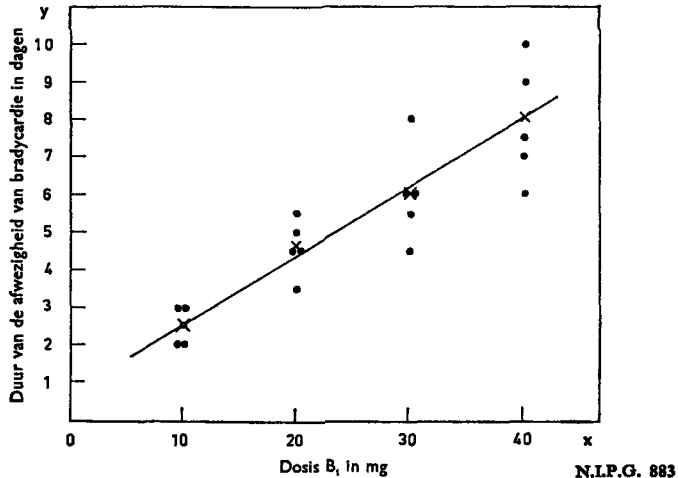
Het spreidingsdiagram in figuur 18.8 laat zien, dat de gemiddelden

\bar{y}_j ongeveer op een rechte liggen, maar dat de spreiding van de y -waarden sterk toeneemt, naarmate de dosis B_1 groter wordt.

Tabel 18.6

Dosis in mg x	Duur v. d. afwezigheid van bradycardie in dagen y					Σy_j	\bar{y}_j
10	2	2	2,5	3	3	12,5	2,5
20	3,5	4,5	4,5	5	5,5	23,0	4,6
30	4,5	5,5	6	6	8	30,0	6,0
40	6	7	7,5	9	10	39,5	7,9

Uit: P. C. LEONG, Bioch. J. 33 (1939), 1397.



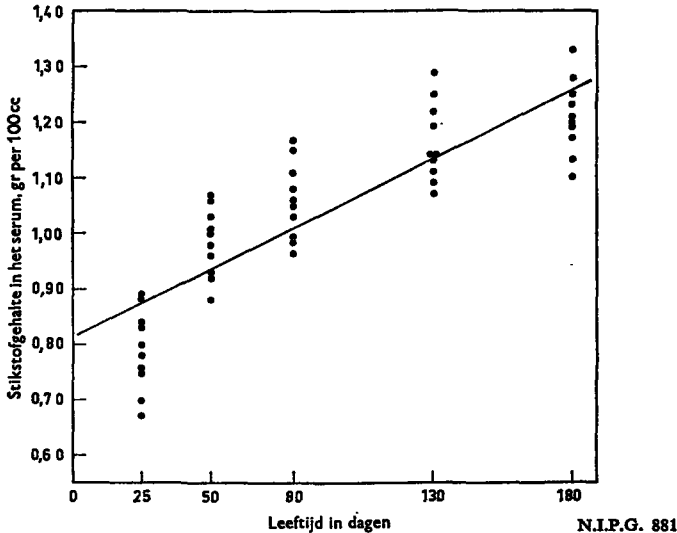
Figuur 18.8. Spreidingsdiagram van de gegevens in tabel 18.6 met de op het oog getrokken regressielijn, die zo goed mogelijk bij de gemiddelden \bar{y}_j (x) past.

18.6. De gegevens in tabel 18.7, waarvan het spreidingsdiagram in figuur 18.9 is getekend, betreffen de stikstofgehalten in het serum

Tabel 18.7

Leeftijd in dagen x	Stikstofgehalte in gr per 100 cc serum y											Σy_j	\bar{y}_j
25	0,67	0,70	0,75	0,76	0,78	0,80	0,83	0,84	0,88	0,89		7,90	0,790
50	0,88	0,92	0,93	0,96	0,98	1,00	1,01	1,03	1,06	1,07		9,84	0,984
80	0,96	0,98	0,99	1,03	1,05	1,06	1,08	1,11	1,15	1,17		10,58	1,058
130	1,07	1,09	1,11	1,13	1,14	1,14	1,19	1,22	1,25	1,29		11,63	1,163
180	1,10	1,13	1,17	1,19	1,20	1,21	1,23	1,25	1,28	1,33		12,09	1,209

(in gr per 100 cc serum) bij vijf aselect gekozen groepen van 10 ratten op vijf verschillende leeftijden (in dagen). Het is duidelijk, dat bij dit probleem de regressie van y op x kromlijinig is.



Figuur 18.9. Spreidingsdiagram van de gegevens in tabel 18.7: *kromlijnige* regressie. De berekende best-passende rechte (zie voorbeeld 19.5), die in deze figuur is getekend, blijkt dan ook bijzonder slecht te passen.

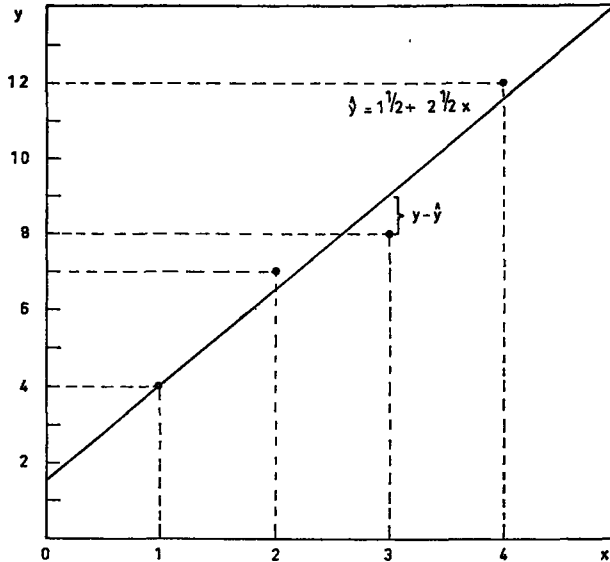
18.2.2. BEREKENING VAN EEN REGRESSIELIJN VOLGENS DE METHODE DER KLEINSTE KWADRATEN

Wij demonstreren deze methode aan de hand van het volgende eenvoudige voorbeeld:

x	1	2	3	4
y	4	7	8	12

De punten (x, y) zijn in figuur 18.10 getekend. Gevraagd wordt naar een rechte met de vergelijking $\hat{y} = a + bx$, die zo goed mogelijk bij deze punten past.¹ Bij de methode der kleinste kwadraten worden nu voor a en b zodanige waarden gezocht, dat de som van de gekwadrateerde verschillen tussen de gegeven en de uit deze vergelijking volgende waarden van y , $\sum (y - \hat{y})^2$, zo klein mogelijk wordt (slechts als alle punten (x, y) op een rechte liggen, is deze som gelijk aan 0).

¹ Zie bijlage II, par. 6. Het dakje boven de y dient om aan te geven, dat wij hier niet te maken hebben met een functionele (algebraïsche) relatie.



Figuur 18.10.

N.I.P.G. 888

Wij stellen nu eerst een tabel samen, waarin de waarden van x en y , de uit de vergelijking $\hat{y} = a + bx$ volgende waarden van \hat{y} en de verschillen $y - \hat{y}$ zijn opgenomen:

x	y	\hat{y}	$y - \hat{y}$
1	4	$a + b$	$4 - a - b$
2	7	$a + 2b$	$7 - a - 2b$
3	8	$a + 3b$	$8 - a - 3b$
4	12	$a + 4b$	$12 - a - 4b$

Door de kwadraten van de verschillen uit te werken en te sommeren, vindt men

$$\begin{aligned} (4 - a - b)^2 &= a^2 + b^2 + 16 + 2ab - 8a - 8b \\ (7 - a - 2b)^2 &= a^2 + 4b^2 + 49 + 4ab - 14a - 28b \\ (8 - a - 3b)^2 &= a^2 + 9b^2 + 64 + 6ab - 16a - 48b \\ (12 - a - 4b)^2 &= a^2 + 16b^2 + 144 + 8ab - 24a - 96b \end{aligned}$$

$$\Sigma (y - \hat{y})^2 = 4a^2 + 30b^2 + 273 + 20ab - 62a - 180b \quad +$$

$$(18.1) \quad = 4a^2 + a(20b - 62) + (273 - 180b + 30b^2)$$

$$(18.2) \quad = 30b^2 + b(20a - 180) + (273 - 62a + 4a^2)$$

Beschikt men echter over een vergelijking van de vorm

$$S = pm^2 + qm + c,$$

waarin p en q bekend zijn, dan is S minimaal voor $m = \frac{-q}{2p}$.

(Luidt bv. de vergelijking: $S = 9m^2 - 18m + 4$, dan is S minimaal voor $m = -(-18)/18 = 1$. Dit minimum is dan $S = 9 - 18 + 4 = -5$. Voor elke andere waarde van m neemt S een waarde aan, die groter is dan -5 .)

Uit (18.1) volgt dus, dat $\Sigma(y - \hat{y})^2$ minimaal is, indien

$$a = \frac{-(20b - 62)}{(2)(4)}, \quad \text{dat is als}$$

$$(18.3) \quad 4a + 10b - 31 = 0.$$

Uit (18.2) volgt, dat $\Sigma(y - \hat{y})^2$ minimaal is, indien

$$b = \frac{-(20a - 180)}{(2)(30)}, \quad \text{dat is als}$$

$$(18.4) \quad 10a + 30b - 90 = 0.$$

Wij beschikken nu over twee vergelijkingen met twee onbekenden, zodat hieruit a en b kunnen worden opgelost. Vermenigvuldigt men (18.3) met -3 en telt men bij de nieuwe vergelijking (18.4) op, dan vindt men:

$$\begin{array}{r} -12a - 30b + 93 = 0 \\ 10a + 30b - 90 = 0 \\ \hline -2a \qquad \qquad + 3 = 0, \end{array}$$

zodat $a = 1,5$ en $b = 2,5$. De best-passende rechte volgens de methode der kleinste kwadraten is dus

$$\hat{y} = 1,5 + 2,5x.$$

Deze regressielijn van y op x is in figuur 18.10 getekend. De richtingscoëfficiënt b ($b_{y \cdot x}$) van deze lijn noemt men de *regressiecoëfficiënt van y op x* . De regressielijn loopt door het punt (\bar{x}, \bar{y}) , daar voor

$$x = \bar{x} = 10/4 = 2,5 \text{ geldt, dat } \hat{y} = 1,5 + (2,5)(2,5) = 7,75 = \bar{y}.$$

Het voorgaande demonstreert de gedachtengang, die aan de methode van de kleinste kwadraten ten grondslag ligt, maar men kan b en a ook op eenvoudiger wijze berekenen. Beschouwen wij nevenstaande tabel en vergelijken wij de voor het voorbeeld verkregen waarden van Σx , Σx^2 , Σy , Σy^2 en Σxy met de vergelijkingen (18.3) en (18.4), dan is in te zien, dat voor vergelijking (18.3) kan worden geschreven

$$(18.5) \quad na + b\Sigma x - \Sigma y = 0,$$

$n = 4$			
x	y	x^2	xy
1	4	1	4
2	7	4	14
3	8	9	24
4	12	16	48
10	31	30	90
Σx	Σy	Σx^2	Σxy

en voor vergelijking (18.4):

$$(18.6) \quad a \Sigma x + b \Sigma x^2 - \Sigma xy = 0.$$

Vermenigvuldigt men nu (18.5) met $\Sigma x/n$ en trekt men de uitkomst van (18.6) af, dan vindt men:

$$\begin{aligned} a \Sigma x + b \Sigma x^2 - \Sigma xy &= 0 \\ a \Sigma x + \frac{b(\Sigma x)^2}{n} - \frac{\Sigma x \Sigma y}{n} &= 0 \\ \hline b \left[\Sigma x^2 - \frac{(\Sigma x)^2}{n} \right] &= \Sigma xy - \frac{\Sigma x \Sigma y}{n}, \end{aligned}$$

zodat

$$(18.7.1) \quad b = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}},$$

of

$$(18.7.2) \quad b = \frac{\Sigma xy - n \cdot \bar{x} \cdot \bar{y}}{\Sigma x^2 - n \cdot \bar{x}^2}.$$

$$\begin{aligned} \text{Maar: } \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) &= \Sigma (x_i y_i - x_i \cdot \bar{y} - y_i \cdot \bar{x} + \bar{x} \cdot \bar{y}) \\ &= \Sigma x_i y_i - \bar{y} \cdot \Sigma x_i - \bar{x} \cdot \Sigma y_i + n \cdot \bar{x} \cdot \bar{y} \\ &= \Sigma x_i y_i - n \cdot \bar{x} \cdot \bar{y} \end{aligned}$$

en

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \Sigma x_i^2 - n \cdot \bar{x}^2,$$

zodat men ook kan schrijven

$$(18.7.3) \quad b = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{\Sigma (x_i - \bar{x})^2} = \frac{\Sigma (x - \bar{x})(y - \bar{y})}{\Sigma (x - \bar{x})^2}.$$

Verder volgt uit (18.5):

$$(18.8) \quad a = \frac{\Sigma y - b \cdot \Sigma x}{n} = \bar{y} - b \cdot \bar{x}.$$

De regressievergelijking van y op x kan dus ook worden geschreven in de vorm:

$$\hat{y} = a + bx = \bar{y} - b \cdot \bar{x} + b \cdot x,$$

of

$$(18.9) \quad \hat{y} - \bar{y} = b(x - \bar{x}).$$

Het voorgaande dient om te laten zien, hoe men de best-passende regressielijn van y op x volgens de methode der kleinste kwadraten kan berekenen. Wij hebben hierbij gebruik gemaakt van een eenvoudig

voorbeeld met een viertal waarnemingsparen (x, y) en ons niet bezig gehouden met de vraag, of deze een populatie dan wel een steekproef betreffen.

Bij de behandeling van de regressie in hoofdstuk 19 zullen wij echter met de regressielijn van y op x in een populatie en met een schatting van deze regressielijn op basis van een steekproef uit deze populatie te maken hebben. Wij zullen daarbij de coëfficiënten van de populatie-regressielijn met A en B en hun schattingen uit de steekproef met a en b aanduiden. Daar het hier steeds om de regressie van y op x gaat, behoeft men deze symbolen niet van een index te voorzien.

Bij de correlatierekening, waarvan in 18.3.2 reeds enkele begrippen worden besproken, komen de regressie van y op x en de regressie van x op y aan de orde, zodat de respectieve coëfficiënten moeten worden onderscheiden. Bij de regressie van y op x zullen wij deze aanduiden met B_{yx} (b_{yx}) en A_{yx} (a_{yx}) en bij de regressie van x op y met B_{xy} (b_{xy}) en A_{xy} (a_{xy}).

18.2.3. BEREKENING VAN EEN BEST-PASSENDE KROMME VOLGENS DE METHODE DER KLEINSTE KWADRATEN

Als men over n waarnemingsparen (x, y) beschikt kan men ook de vergelijking van een best-passende kromme, zoals bv. een k -de graads vergelijking van de vorm

$$y = a + b_1x + b_2x^2 + b_3x^3 + \dots + b_kx^k$$

zodanig berekenen, dat $\Sigma(y - \hat{y})^2$ minimaal is. Hoewel wij de kromlijnige regressie verder niet behandelen, vermelden wij dat men voor het verkrijgen van zo'n vergelijking van de k -de graad a, b_1, b_2, \dots, b_k dient op te lossen uit de vergelijkingen:

$$\begin{aligned} \Sigma y &= an + b_1 \Sigma x + b_2 \Sigma x^2 + \dots + b_k \Sigma x^k \\ \Sigma xy &= a \Sigma x + b_1 \Sigma x^2 + b_2 \Sigma x^3 + \dots + b_k \Sigma x^{k+1} \\ \Sigma x^2 y &= a \Sigma x^2 + b_1 \Sigma x^3 + b_2 \Sigma x^4 + \dots + b_k \Sigma x^{k+2} \\ &\vdots \\ \Sigma x^k y &= a \Sigma x^k + b_1 \Sigma x^{k+1} + b_2 \Sigma x^{k+2} + \dots + b_k \Sigma x^{2k}. \end{aligned}$$

Voor het aanpassen van een kromme van de tweede graad luiden deze vergelijkingen dus:

$$\begin{aligned} \Sigma y &= an + b_1 \Sigma x + b_2 \Sigma x^2 \\ \Sigma xy &= a \Sigma x + b_1 \Sigma x^2 + b_2 \Sigma x^3 \\ \Sigma x^2 y &= a \Sigma x^2 + b_1 \Sigma x^3 + b_2 \Sigma x^4. \end{aligned}$$

Links in tabel 18.8 zijn zes waarnemingsparen (x, y) opgenomen. De sommen, die nodig zijn voor berekening van de best-passende kromme van de tweede graad (parabool) zijn in deze tabel berekend.

Tabel 18.8. Sommen voor de berekening van de best-passende kromme van de tweede graad volgens de methode der kleinste kwadraten

x	y	xy	x^2y	x^2	x^3	x^4
1	2	2	2	1	1	1
2	3	6	12	4	8	16
3	6	18	54	9	27	81
4	11	44	176	16	64	256
5	18	90	450	25	125	625
7	38	266	1862	49	343	2401
22	78	426	2556	104	568	3380
Σx	Σy	Σxy	Σx^2y	Σx^2	Σx^3	Σx^4

De vergelijkingen, waaruit a , b_1 en b_2 kunnen worden opgelost luiden dus:

$$\begin{aligned} \text{I.} \quad & 78 = 6a + 22b_1 + 104b_2 \\ \text{II.} \quad & 426 = 22a + 104b_1 + 568b_2 \\ \text{III.} \quad & 2556 = 104a + 568b_1 + 3380b_2. \end{aligned}$$

Vermenigvuldigt men vergelijking II met 3 en vergelijking I met 11, dan verkrijgt men:

$$\begin{array}{r} 1278 = 66a + 312b_1 + 1704b_2 \\ 858 = 66a + 242b_1 + 1144b_2 \\ \hline 420 = \quad \quad 70b_1 + 560b_2 \end{array}$$

of

$$b_1 + 8b_2 = 6, \quad b_1 = 6 - 8b_2.$$

Vermenigvuldigt men vergelijking III met 3 en vergelijking I met 52, dan is

$$\begin{array}{r} 7668 = 312a + 1704b_1 + 10140b_2 \\ 4056 = 312a + 1144b_1 + 5408b_2 \\ \hline 3612 = \quad \quad 560b_1 + 4732b_2 \end{array}$$

Substitueert men in deze vergelijking $b_1 = 6 - 8b_2$, dan vindt men dat $b_2 = 1$ en $b_1 = -2$. Door substitutie van deze waarden in één der oorspronkelijke vergelijkingen volgt tenslotte $a = 3$. De best-passende tweedegraadskromme van y op x luidt dus $\hat{y} = 3 - 2x + x^2$ en deze parabool blijkt precies bij de punten (x, y) te passen, zodat $\Sigma(y - \hat{y})^2 = 0$.

Het is duidelijk, dat het oplossen van een vergelijking van een hogere graad op deze wijze een moeizaam rekenwerk wordt en er zijn dan ook verschillende methoden ontwikkeld om dit min of meer te vereenvoudigen. Als de waarden van x equidistant zijn, kan reeds een aanzienlijke vereenvoudiging worden bereikt door op de in 19.1.6 beschreven wijze een hulpgrootheid u in te voeren (alle sommen van

oneven machten van x zijn dan gelijk aan nul). Als per waarde van x een constant aantal waarden van y aanwezig is, kan men de berekeningen het best uitvoeren met behulp van zg. orthogonale polynomen. Deze methode wordt o.m. uitvoerig beschreven door BENNETT en FRANKLIN (4, blz. 255 e.v.) en zij bezit tevens het voordeel, dat allerlei toetsen eenvoudig zijn te verrichten.

Bij de waarnemingsparen in tabel 18.3 kan men volgens de zojuist beschreven methode de tweede-graads kromme

$$\hat{y} = -579,23 + 20,267x - 0,10232x^2$$

berekenen en zoals uit onderstaande tabel blijkt, past deze zeer bevredigend bij de waarnemingen:

x	40	50	60	70	80	90
y	69	175	272	335	390	415
\hat{y}	67,7	178,3	268,4	338,1	387,3	416,0
$y - \hat{y}$	1,3	-3,3	3,6	-3,1	2,7	-1,0

Men kan volgens de methode van de kleinste kwadraten ook een bestpassende kromme van een andere vorm bepalen. Zo hebben wij bv. voor de waarnemingsparen in tabel 18.7 de best-passende logaritmische groeikromme (van de vorm: $y = a + bx + c \log x$) bepaald.¹ Deze vergelijking is $\hat{y} = -0,08995 - 0,0009361x + 0,6499 \cdot \log x$ en zij past zeer goed bij de waargenomen gemiddelden \bar{y}_j :

x	25	50	80	130	180
\bar{y}_j	0,790	0,984	1,058	1,163	1,209
\hat{y}_j	0,795	0,967	1,072	1,162	1,207
$\bar{y}_j - \hat{y}_j$	-0,005	0,017	-0,014	0,001	0,002

18.3. Correlatie

18.3.1. CORRELATIEPROBLEMEN

Men spreekt van een correlatieprobleem, als men voor een tweedimensionale populatie wil onderzoeken, of er een relatie bestaat tussen twee grootheden x en y en zo ja, van welke aard en sterkte deze co(r)relatie is. De waarnemingsparen (x, y) worden dan verkregen,

¹ Volledigheidshalve vermelden wij de hiervoor op te lossen vergelijkingen, die luiden:

$$\Sigma y = an + b_1 \Sigma x + b_2 \Sigma (\log x)$$

$$\Sigma xy = a \Sigma x + b_1 \Sigma x^2 + b_2 \Sigma (x \cdot \log x)$$

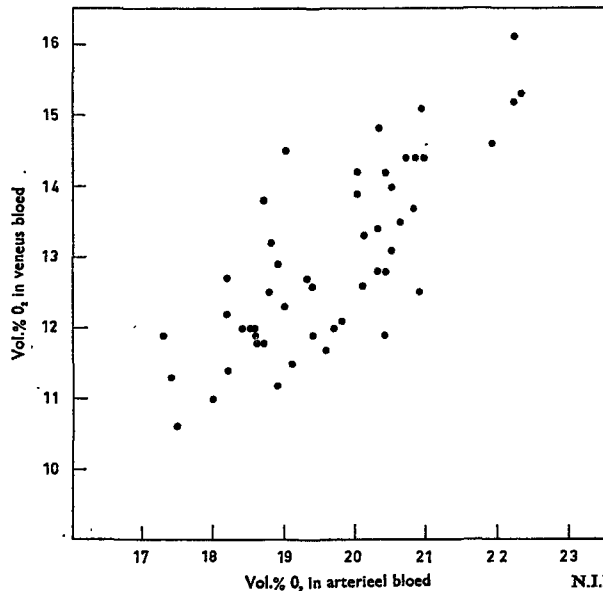
$$\Sigma (y \cdot \log x) = a \Sigma (\log x) + b_1 \Sigma (x \cdot \log x) + b_2 \Sigma [(\log x)^2].$$

door uit de populatie aselect een steekproef van n elementen te trekken en van elk element vast te stellen, welke waarde van x en welke waarde van y het bezit.

Tabel 18.9

Per- soon	Arterie x	Vena y	Per- soon	Arterie x	Vena y	Per- soon	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. FERRIS, 'Introduction to the analysis of experimental data', cursus Biostatistics 202, Columbia University, U.S.A.



N.I.P.G. 2032

Figuur 18.11. Spreidingsdiagram van de waarnemingsparen in tabel 18.9.

*Voorbeelden*¹

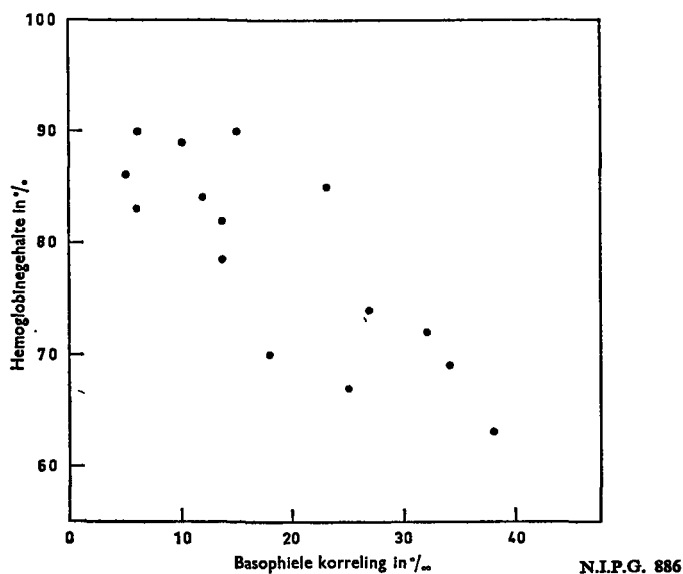
18.7. In tabel 18.9 staan de zuurstofgehalten in vol. % van arterieel bloed (x) en van veneus bloed (y) bij 50 gezonde mannen. Het spreidingsdiagram van deze waarnemingsparen is in figuur 18.11 getekend.

18.8. De waarnemingsparen in tabel 18.10, die in figuur 18.12 zijn getekend, betreffen de basophile korreling (x) en het hemoglobinegehalte (y) van 15 arbeiders met een dreigende loodintoxicatie.

Tabel 18.10

Arbeider	x	y	Arbeider	x	y	Arbeider	x	y
1	38	63	6	5	86	11	12	84
2	23	85	7	32	72	12	6	90
3	14	79	8	14	82	13	6	83
4	10	89	9	18	70	14	34	69
5	25	67	10	15	90	15	27	74

Uit de Afd. Arbeidsgeneeskunde van het N.I.P.G.

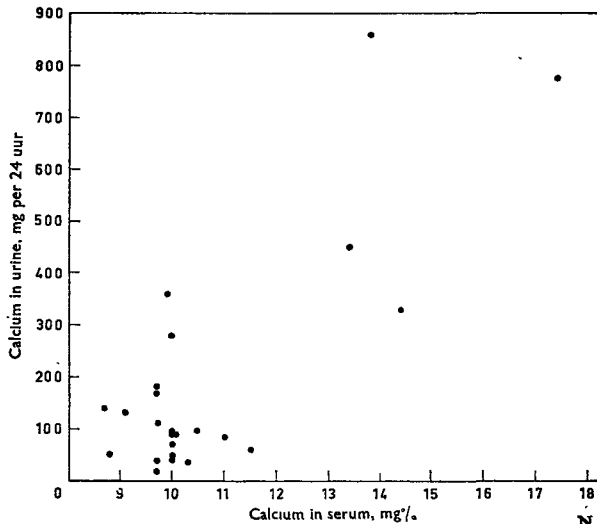


Figuur 18.12. Spreidingsdiagram van de waarnemingsparen in tabel 18.10.

18.9. Beschouw nogmaals de waarnemingsparen in opgave 10.23 (blz. 254), betreffende de calciumuitscheiding in de urine (in mg per 24 uur) en het calciumgehalte in het serum (mg pct) bij 24 patiënten met

¹ Zie ook opgave 2.3 (blz. 42) en voorbeeld 2.6 (blz. 47).

verschillende kwaadaardige aandoeningen, die in figuur 18.13 grafisch zijn voorgesteld.



N.I.P.G. 887

Figuur 18.13. Spreidingsdiagram van de gegevens in opgave 10.23.

18.3.2. DE COVARIANTIE EN DE CORRELATIECOËFFICIËNT

Beschouw een tweedimensionaal verdeelde populatie van N elementen, waarvan elk element een paar waarden (x, y) bezit. De grootheid x bezit op deze populatie een verdeling met gemiddelde μ_x en spreiding σ_x en de grootheid y een verdeling met gemiddelde μ_y en spreiding σ_y . Men definieert nu als de *covariantie* van de grootheden x en y :

$$(18.10.1) \quad \sigma_{xy} = \frac{\sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)}{N}.$$

Voor de berekening van σ_{xy} kan men het best gebruik maken van

$$(18.10.2) \quad \sigma_{xy} = \frac{\sum x_i y_i - \frac{\sum x_i \cdot \sum y_i}{N}}{N} = \frac{\sum x_i y_i}{N} - \mu_x \mu_y.$$

Het is duidelijk dat σ_{xy} groot zal zijn, als grote waarden van x in het algemeen samengaan met grote waarden van y en kleine waarden van x met kleine waarden van y . Een grote waarde x_i gepaard met een grote waarde y_i levert immers in de teller van (18.10) een grote waarde van $x_i - \mu_x$ en een grote waarde van $y_i - \mu_y$ op, zodat ook hun product een grote positieve waarde heeft. Een kleine waarde x_i gepaard met een

kleine waarde y_i leveren een product van twee sterk negatieve factoren, dus eveneens een sterk positieve bijdrage tot de som in de teller op. Verder is wel in te zien, dat σ_{xy} sterk negatief wordt, als grote waarden van x in het algemeen samengaan met kleine waarden van y en omgekeerd. De covariantie is dus een maat voor de co(r)relatie tussen de variabelen x en y . Als σ_{xy} positief is, spreekt men van een *positieve* correlatie, wanneer σ_{xy} negatief is van een *negatieve* correlatie.

De covariantie is echter nog geen geschikte maat voor de correlatie. Zou men bv. door een schaalverandering de spreiding van x en/of y vergroten, zonder hun simultane verdeling te veranderen, dan zal de covariantie naar evenredigheid toenemen. Men kan dit bezwaar echter opheffen door de covariantie door de spreidingen van x en y te delen, d.i. door de *correlatiecoëfficiënt*¹

$$(18.11.1) \quad \rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$$

te gebruiken (als voldoende duidelijk is, op welke variabelen deze coëfficiënt betrekking heeft, laat men de index gewoonlijk achterwege). De correlatiecoëfficiënt kan worden berekend met

$$(18.11.2) \quad \rho = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum (x_i - \mu_x)^2 \cdot \sum (y_i - \mu_y)^2}} = \frac{\sum xy - \frac{\sum x \cdot \sum y}{N}}{\sqrt{[\sum x^2 - (\sum x)^2/N] [\sum y^2 - (\sum y)^2/N]}}$$

Berekent men voor een tweedimensionaal verdeelde populatie de regressielijn van y op x volgens de methode der kleinste kwadraten, dan is volgens (18.7.3) en het voorgaande

$$(18.12) \quad B_{yx} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{\sum (x_i - \mu_x)^2} = \frac{\sigma_{xy}}{\sigma_x^2} = \rho \frac{\sigma_y}{\sigma_x}$$

Berekent men echter de regressielijn van x op y , dan is

$$(18.13) \quad B_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{\sum (y_i - \mu_y)^2} = \frac{\sigma_{xy}}{\sigma_y^2} = \rho \frac{\sigma_x}{\sigma_y}$$

Hieruit volgt, dat men ook kan schrijven:

$$(18.14) \quad \rho^2 = B_{yx} \cdot B_{xy}$$

of $\rho = \pm \sqrt{B_{yx} \cdot B_{xy}}$. Het teken van ρ is positief, als σ_{xy} positief is, dus bij positieve correlatie en dan zijn zowel B_{yx} als B_{xy} positief. Het teken van ρ is negatief, als σ_{xy} negatief is, dus bij negatieve correlatie en dan zijn B_{yx} en B_{xy} negatief. Als de grootheden x en y onderling

¹ ρ = de Griekse letter rho.

onafhankelijk zijn, is $\sigma_{xy} = 0$, zodat ook $\rho = 0$. Is echter $\rho = 0$, dan behoeven de grootheden x en y niet onafhankelijk te zijn: deze uitkomst kan bv. optreden als de relatie tussen deze grootheden *niet* rechtlijnig is. De correlatiecoëfficiënt blijkt dan ook slechts een goede maat voor de correlatie van x en y te zijn, als zowel de regressie van y op x als de regressie van x op y rechtlijnig is. Bestaat er tussen x en y een functioneel verband van de vorm $y = A_{yx} + B_{yx}x$, zodat alle punten (x, y) op één rechte liggen (de regressielijn van y op x samenvalt met die van x op y), dan geldt, dat $x = y/B_{yx} - A_{yx}/B_{yx}$, zodat $B_{xy} = 1/B_{yx}$. Volgens (18.14) is dan $\rho^2 = 1$, of $\rho = \pm 1$. Wanneer $\rho = 1$ spreekt men van een *volkomen positieve* relatie en bij $\rho = -1$ van een *volkomen negatieve* correlatie. De absolute waarde van de correlatiecoëfficiënt is groot, als de simultane verdeling van x en y een sterke concentratie in de buurt van een rechte lijn vertoont en neemt af, naarmate deze concentratie minder sterk wordt.

Voorbeelden

18.10. Beschouw de tweedimensionale verdeling in tabel 18.1. Wegens de symmetrie van de verdelingen van x en y geldt hiervoor, dat $\mu_x = 12$ en $\mu_y = 27$. Verder is $\Sigma xy = \Sigma yx = 1(10)(29) + 2(11)(29) + 4(12)(29) + \dots + 1(14)(25) = 32400$, zodat volgens (18.10.2)

$$\sigma_{xy} = \frac{32400}{100} - (12)(27) = 0,$$

zodat ook $\rho = 0$. Daar in dit geval direct is in te zien, dat zowel de regressie van y op x als de regressie van x op y rechtlijnig is, volgt hieruit dat de grootheden x en y onderling onafhankelijk zijn (op welke andere wijze blijkt dit?).

18.11. Beschouw de nevenstaande tweedimensionale verdeling. Hiervoor is $\mu_x = \mu_y = 2$ en $\Sigma xy = \Sigma yx = 3(2)(3) + 1(1)(2) + \dots + 1(3)(1) = 40$, zodat

$$\sigma_{xy} = \frac{40}{10} - (2)(2) = 0, \quad \rho = 0.$$

$y \backslash x$	1	2	3	Totaal
3		3		3
2	1	2	1	4
1	1	1	1	3
Tot.	2	6	2	10

De grootheden x en y zijn echter niet onafhankelijk verdeeld. De regressie van x op y is wel, die van y op x niet rechtlijnig.

18.12. Voor de hierna volgende tweedimensionale verdeling is: $\Sigma x = 61$, $\Sigma x^2 = 207$, $\Sigma y = 138$, $\Sigma y^2 = 1036$, $\Sigma xy = 379$ (controleer deze uitkomsten),

zodat:

$$\Sigma (x - \mu_x)^2 = 207 - (61)^2/20 = 20,95,$$

$$\Sigma (y - \mu_y)^2 = 1036 - (138)^2/20 = 83,8,$$

$$\Sigma (x - \mu_x)(y - \mu_y) = 379 - (61)(138)/20 = -41,9.$$

Uit (18.11.2) volgt:

$$\rho = \frac{-41,90}{\sqrt{(20,95)(83,8)}} = -1.$$

$x \backslash y$	1	2	3	4	5	Totaal
11	1					1
9		5				5
7			8			8
5				4		4
3					2	2
Tot.	1	5	8	4	2	20

Verder is

$$B_{yx} = -41,9/20,95 = -2, \text{ en } B_{xy} = -41,9/83,8 = -\frac{1}{2}.$$

Onderstel, dat men uit een tweedimensionaal verdeelde populatie van de omvang N met covariantie σ_{xy} en correlatiecoëfficiënt ρ aselekt een steekproef trekt van n elementen (met teruglegging, of zonder teruglegging met $n \ll N$). Uit de n waarnemingsparen (x, y) kan men dan een zuivere schatting van σ_{xy} verkrijgen door te nemen

$$(18.15) \quad s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}}{n-1}.$$

Een zuivere schatting van ρ is:

$$(18.16) \quad r = \frac{s_{xy}}{s_x s_y} = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma (x_i - \bar{x})^2 \cdot \Sigma (y_i - \bar{y})^2}} = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}}{\sqrt{[\Sigma x^2 - (\Sigma x)^2/n][\Sigma y^2 - (\Sigma y)^2/n]}}.$$

Verder geldt, dat

$$b_{yx} = r \frac{s_y}{s_x} \text{ en } b_{xy} = r \frac{s_x}{s_y}.$$

Ook de grootheid r kan alleen tussen -1 en $+1$ variëren. Bestaat er tussen de grootheden x en y een exact lineair verband, dan zal $|r| = 1$ zijn. Als de grootheden onafhankelijk verdeeld zijn, zal r in het algemeen niet gelijk zijn aan nul, maar een kansverdeling bezitten met gemiddelde $\mu_r = 0$ en een zekere spreiding, die afhangt van n en van het karakter van de tweedimensionale populatieverdeling. Deze kansverdeling is voor enkele bijzondere gevallen berekend, o.m. als de simultane verdeling van x en y een tweedimensionaal normale verdeling is. Wij komen hierop in hoofdstuk 20 terug.

18.4. Opgaven

18.1. Uit WILKS (37).

Bij een reeks proeven betreffende de relatie tussen de concentratie van penicilline in eenheden per ml (x) en de gemiddelde diameter van de remmingszone in mm (y) verkreeg men de volgende uitkomsten:

x	1	2	4	8	16	32
y	15,9	17,8	19,5	21,35	23,1	24,8

- Maak van deze gegevens een grafiek en teken op het oog de regressielijn van y op x .
- Schat hieruit de gemiddelde diameter voor een concentratie van 17 en van 22 eenheden per ml.
- Hoeveel eenheden per ml zijn (ongeveer) nodig om een gemiddelde diameter van 21 mm te verkrijgen?
- Vervang x door $z = \log x$ en voer daarna nogmaals *a* t/m *c* uit.

18.2. Bereken de regressievergelijking van y op x volgens de methode der kleinste kwadraten voor onderstaande gegevens:

a.

x	3,4	4,3	5,4	6,7	8,7	10,6
y	4,5	5,8	6,8	8,1	10,5	12,7

b.

x	38	32	28	27	26	24	20	19	15	12
y	18	20	26	32	36	34	36	42	46	53

18.3. Bereken σ_{xy} en ρ voor de nevenstaande tweedimensionale populatieverdeling.

Zijn de regressies rechtlijnig?

$y \backslash x$	0	1	2	3	4	Totaal
5					1	1
4			1	2	1	4
3	1	2	4	2	1	10
2	1	2	1			4
1	1					1
Tot.	3	4	6	4	3	20

$y \backslash x$	3	4	5	6	Totaal
15			1	2	3
14		1	3		4
13	2	4			6
12		1	3		4
11			1	2	3
Tot.	2	6	8	4	20

18.4. Beschouw nevenstaande tweedimensionale verdeling. Kunt U zonder berekening de waarde van B_{yx} geven? Welke waarde heeft dan ρ ?

Zijn de grootheden x en y onafhankelijk?

18.5. Ieder stel van de onderstaande gegevens heeft betrekking op een tweedimensionale verdeling, maar bevat een discrepantie, waardoor het in werkelijkheid niet verkregen kan zijn, tenzij bij de berekeningen een fout is gemaakt. Geef aan, wat onjuist is.

a. $\hat{y} = 0,96x + 13,3$ en $\hat{x} = 1,42y - 12,5$.

b. $\Sigma (x_i - \bar{x})(y_i - \bar{y}) = 1543$, $\Sigma (x_i - \bar{x})^2 = 1560$, $\Sigma (y_i - \bar{y})^2 = 1136$.

c. $r = 0,65$, $b_{yx} = -0,24$, $b_{xy} = -0,72$.

d. $\rho = 0,5$, $\sigma_x = 4$, $\sigma_y = 10$, $B_{yx} = 0,25$.

18.6. Uit het examen statistisch analyst 1955:

Uit 5 fiches, waarop de cijfers 1, 2, 3, 4 resp. 5 staan vermeld, worden er aselect achtereenvolgens twee getrokken, zonder teruglegging van het eerste. Noemt men de getallen, die op de twee getrokken fiches staan, resp. x en y , dan bezitten x en y een simultane waarschijnlijkheidsverdeling. Bereken de correlatiecoëfficiënt $\rho(x, y)$.

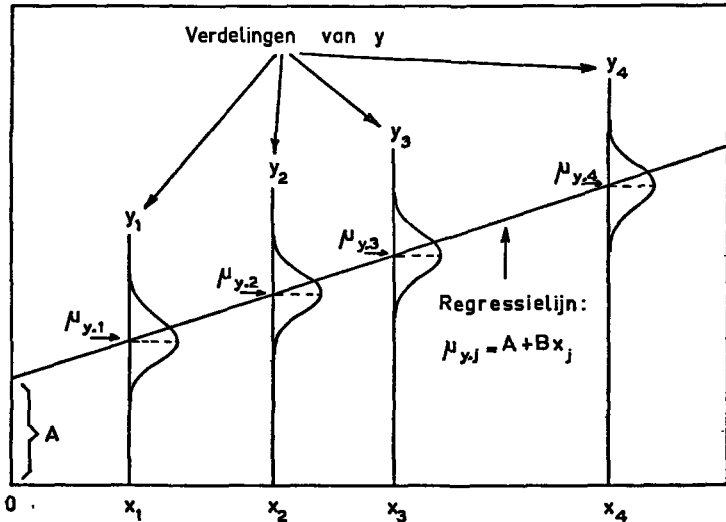
REGRESSIE-ANALYSE

19.1. Lineair-normale regressie

19.1.1. BESCHRIJVING VAN HET MODEL

Wij beschouwen een populatie bestaande uit elementen, die alle een waardenpaar (x, y) dragen. Kiest men hieruit een sub-populatie (deelverzameling), die bestaat uit k waarden x_j ($j = 1, 2, \dots, k$), dan behoort bij elke x_j een verdeling van de grootte y , waarvan wij het gemiddelde met $\mu_{y,j}$ en de variantie met $\sigma_{y,j}^2$ aanduiden. Het gemiddelde en de variantie van alle waarden van y in de deelverzameling noemen wij μ_y , resp. σ_y^2 . Wij onderstellen nu het volgende:

1. De verdelingen van y zijn voor alle waarden x_j normaal en zij bezitten dezelfde variantie $\sigma_{y,x}^2$ (dus: $\sigma_{y,1}^2 = \sigma_{y,2}^2 = \dots = \sigma_{y,k}^2 = \sigma_{y,x}^2$).



N.I.P.G. 889

Figuur 19.1. Schematische voorstelling van het lineair-normale regressiemodel.

2. De regressie van y op x is rechtlijnig, d.w.z. de gemiddelden $\mu_{y,1}, \mu_{y,2}, \dots, \mu_{y,k}$ liggen op een rechte, waarvan de vergelijking luidt

$$(19.1.1) \quad \mu_{y,j} = A + Bx_j,$$

waarin A en B dus parameters zijn¹. Daar $A = \mu_y - B \cdot \bar{x}$ kan men voor deze vergelijking ook schrijven

$$(19.1.2) \quad \mu_{y,j} = \mu_y + B(x_j - \bar{x}).$$

Dit z.g. lineair-normale regressiemodel is in figuur 19.1 schematisch weergegeven. De verdelingen van y zijn hierbij in het platte vlak getekend, maar zij dienen uiteraard loodrecht daarop te worden gedacht.

Men kan bij dit model twee grensgevallen onderscheiden. Het eerste treedt op als $\sigma^2_{y \cdot x} = 0$. Alle waarden van y liggen dan op de regressielijn, zodat het verband tussen y en x functioneel is. Het tweede geval doet zich voor als de regressielijn evenwijdig aan de X -as verloopt, zodat $\mu_{y,1} = \mu_{y,2} = \dots = \mu_{y,k} = \mu_y$ en $B = 0$.

19.1.2. SCHATTING VAN A , B EN $\sigma^2_{y \cdot x}$ UIT EEN STEEKPROEF

Onderstel, dat men bij elke waarde x_j aselect een steekproef van n_j elementen heeft getrokken en voor elk element de waarde van y heeft bepaald, zodat er in totaal $n_1 + n_2 + \dots + n_k = \Sigma n_j = m$ waarnemingsparen verkregen zijn. Een willekeurige waarde van y kan dan worden aangeduid met y_{ij} , waarbij de index j ($j = 1, 2, \dots, k$) aangeeft dat de waarde behoort tot de steekproef die bij x_j is getrokken en de index i ($i = 1, 2, \dots, n_j$) het volgnummer van deze waarde binnen de steekproef aanduidt. Voor het gemiddelde van de y 's in deze steekproef kan men dan bv. schrijven

$$\bar{y}_j = \frac{\sum_{i=1}^{n_j} y_{ij}}{n_j}$$

en voor het gemiddelde van alle m waarden van y

$$\bar{y} = \frac{\sum_{j=1}^k \sum_{i=1}^{n_j} y_{ij}}{m} \quad (m = \Sigma n_j).$$

De totale kwadratensom van de m waarden van y is, volledig geschreven:

$$S_{\text{totaal}} = \sum_{j=1}^k \sum_{i=1}^{n_j} (y_{ij} - \bar{y})^2.$$

Wij zullen echter, wanneer dit geen aanleiding tot misverstand kan geven, voor de som van alle y -waarden eenvoudig zetten Σy en voor hun kwadratensom $\Sigma (y - \bar{y})^2$. Voor de som van alle m waarden van x zetten wij Σx en voor hun kwadratensom $\Sigma (x - \bar{x})^2$. Tenslotte ver-

¹ Deze parameters worden vaak met α en β aangeduid, maar omdat wij deze symbolen reeds bij de toetsingstheorie hebben ingevoerd gebruiken wij A en B .

eenvoudigen wij de formule voor de som van de producten van elke $y_{ij} - \bar{y}$ en de bijbehorende $x_j - \bar{x}$ tot $\Sigma (x - \bar{x}) (y - \bar{y})$.

Uit de totale steekproef, bestaande uit m waarnemingsparen (x, y) , kan men nu met de in 18.2.2 beschreven methode der kleinste kwadraten de best-passende regressielijn

$$(19.2) \quad \hat{y}_j = a + bx_j = \bar{y} + b(x_j - \bar{x})$$

berekenen, waartoe b met formule (18.7) en a met formule (18.8) kan worden verkregen. Wij vermelden — zonder daarvoor het bewijs te leveren — dat dan:

1. a een zuivere schatting is van de parameter A van de bestudeerde subpopulatie.

2. b een zuivere schatting is van de parameter B van deze populatie.

3. Een zuivere schatting van $\sigma^2_{y..x}$, de gemeenschappelijke variantie van de k verdelingen y_j , kan worden verkregen door te nemen

$$(19.3.1) \quad s^2_{y..x} = \frac{\sum_{j=1}^k \sum_{i=1}^{n_j} (y_{ij} - \hat{y}_j)^2}{m - 2},$$

$$\begin{aligned} \text{waarin } \Sigma \Sigma (y_{ij} - \hat{y}_j)^2 &= \Sigma \Sigma [y_{ij} - \{\bar{y} + b(x_j - \bar{x})\}]^2 \\ &= \Sigma \Sigma [(y_{ij} - \bar{y}) - b(x_j - \bar{x})]^2 = \Sigma (y - \bar{y})^2 - b^2 \Sigma (x - \bar{x})^2 \\ &= \Sigma y^2 - a \Sigma y - b \Sigma xy \end{aligned}$$

de (minimale) som is van de kwadraten van de verschillen tussen elke waargenomen y_{ij} en haar met (19.2) op basis van x_j geschatte gemiddelde \hat{y}_j . Om een zuivere schatting te verkrijgen dient men te delen door $m - 2$, daar in de regressievergelijking twee schattingen (nl. van de parameters A en B) voorkomen. Daar volgens (18.7.3) geldt, dat $b = \Sigma (x - \bar{x}) (y - \bar{y}) / \Sigma (x - \bar{x})^2$ kan men ook schrijven

$$(19.3.2) \quad s^2_{y..x} = \frac{\Sigma (y - \bar{y})^2 - \frac{[\Sigma (x - \bar{x}) (y - \bar{y})]^2}{\Sigma (x - \bar{x})^2}}{m - 2}$$

en deze formule leent zich het best voor de berekening van $s^2_{y..x}$.

Wij geven nu eerst enkele voorbeelden van de berekening van een regressievergelijking met (19.2) en van $s^2_{y..x}$ met (19.3.2). Bij deze voorbeelden komt de vraag aan de orde, of het gebruik van het lineair-normale model gerechtvaardigd is. Uit 19.1.1 blijkt, dat daartoe aan drie voorwaarden dient te zijn voldaan:

1. De regressie van y op x moet rechtlijnig zijn. Men kan weliswaar voor elk gegeven regressieprobleem de best-passende regressielijn volgens (19.2) berekenen, maar dit heeft weinig of geen zin (en kan

zelfs zeer misleidend zijn), wanneer er van rechtlijnigheid geen sprake is. Men dient daarom te allen tijde eerst het spreidingsdiagram te tekenen en dit op rechtlijnigheid te inspecteren. Blijkt de regressie duidelijk kromlijinig te zijn, dan zal men een ander model moeten kiezen (zie 19.2), of zich moeten afvragen of een lineaire regressie verkregen kan worden door het toepassen van een transformatie. Zo kan bv. een exponentiële groeicurve van de vorm $y = AB^x$ door het nemen van logaritmen tot een rechte van de vorm $\log y = \log A + (\log B)x$ worden getransformeerd (men vindt deze curve ook in de gedaante $y = Ae^{Bx}$; in dat geval neme men $\ln y = \ln A + Bx$, waarin $\ln = \log_e$; zie deel I, bijlage 2, par. 4). Bestaat er twijfel omtrent de rechtlijnigheid, dan kan men de best-passende rechte op aanpassing toetsen met de in 19.5.1 beschreven methode.

2. De (populatie) verdelingen van y moeten (pseudo) normaal zijn. Beschikt men over een niet te klein aantal waarnemingen van y per x_j , dan kan men de normaliteit van de y -verdelingen nog wel onderzoeken. Treedt scheefheid naar rechts op dan kan men nagaan of een transformatie van de y -waarden tot normaliteit leidt. Bij kleine aantallen waarnemingen zal men moeten volstaan met een inspectie op uitbijters.

3. De (populatie) verdelingen van y dienen dezelfde spreiding $\sigma_{y..}$ te bezitten. Slechts in dat geval heeft het zin $s_{y..}$ als schatting van deze gemeenschappelijke spreiding te berekenen. De juistheid van deze onderstelling kan worden onderzocht met de toets van BARTLETT of als alle n_j gelijk zijn met de toets van HARTLEY.

Voorbeelden

19.1. Men wil voor een nauwkeurig gedefinieerde populatie van jongens van ongeveer 7 t/m 10 jaar, de regressie van de (gemiddelde) lichaamslengte op de leeftijd bestuderen. Men kiest daartoe een subpopulatie, die bestaat uit jongens van precies 7, 8, 9 en 10 jaar, trekt uit elke leeftijd aselect een steekproef van 8 jongens en meet van elke jongen de lichaamslengte in cm. De uitkomsten van dit onderzoek zijn in tabel 19.1 opgenomen en in figuur 19.2 grafisch voorgesteld.¹

Bij dit voorbeeld is de juistheid van het lineair-normale model zeer plausibel. Lengteverdelingen zijn gewoonlijk pseudo-normaal en bij vele onderzoeken van grote omvang² is gebleken, dat deze verdelingen bij de betrokken leeftijden ongeveer dezelfde spreiding bezitten en dat de regressie vrijwel rechtlijnig is. Zoals de lezer zelf kan nagaan is bij het voorbeeld het verschil tussen de waargenomen

¹ In werkelijkheid zal men vanzelfsprekend steekproeven van grotere omvang bij meerdere leeftijden trekken. De fictieve gegevens in tabel 19.1 zijn aan een groter onderzoek ontleend.

² Zie bv. Report on the heights and weights of school pupils in the county of London in 1949, London County Council, Staples Press Ltd, London, 1950.

spreidingen ver van significant; de hypothese, dat de regressie rechtlijnig is zullen wij in 19.1.5 toetsen.

Tabel 19.1. Leeftijd in jaren (x) en lichaamslengte in cm (y) van 32 jongens.

$x = 7$		$x = 8$		$x = 9$		$x = 10$	
y	y^2	y	y^2	y	y^2	y	y^2
127	16129	125	15625	134	17956	137	18769
120	14400	124	15376	131	17161	136	18496
123	15129	119	14161	139	19321	140	19600
118	13924	128	16384	129	16641	134	17956
122	14884	126	15876	137	18769	139	19321
114	12996	121	14641	141	19881	131	17161
125	15625	132	17424	132	17424	145	21025
120	14400	125	15625	135	18225	141	19881
969	117487	1000	125112	1078	145378	1103	152209

$$\Sigma x = 8(7+8+9+10) = 272, \quad \Sigma x^2 = 8[(7)^2+(8)^2+(9)^2+(10)^2] = 2352$$

$$\Sigma (x - \bar{x})^2 = 2352 - (272)^2/32 = 2352 - 2312 = 40$$

$$\Sigma y = 969 + 1000 + 1078 + 1103 = 4150$$

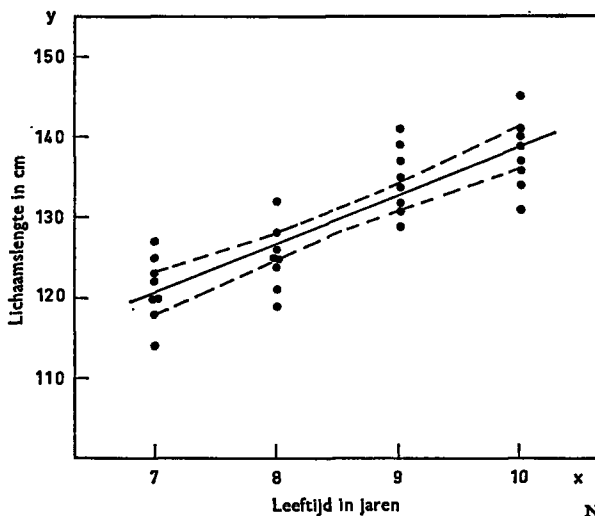
$$\Sigma y^2 = 117487 + 125112 + 145378 + 152209 = 540186$$

$$\Sigma (y - \bar{y})^2 = 540186 - (4150)^2/32 = 540186 - 538203,125 = 1982,875$$

$$\Sigma xy = 7(969) + 8(1000) + 9(1078) + 10(1103) = 35515$$

$$\Sigma x \cdot \Sigma y/n = (272)(4150)/32 = 35275$$

$$\Sigma (x - \bar{x})(y - \bar{y}) = 35515 - 35275 = 240$$



N.I.P.G. 890

Figuur 19.2. Spreidingsdiagram van de gegevens in tabel 19.1 met de regressielijn $y = 78,7 + 6x$, en de verbindingslijnen van de boven- en benedengrenzen van de 95% betrouwbaarheidsintervallen van $\mu_{y,j}$.

In tabel 19.1 zijn de drie kwadratensommen, die wij voor de berekening van de regressielijn nodig hebben [$\Sigma (x - \bar{x})^2$, $\Sigma (y - \bar{y})^2$ en $\Sigma (x - \bar{x})(y - \bar{y})$] reeds berekend. Met (18.7.3) volgt hieruit

$$b = \frac{\Sigma (x - \bar{x})(y - \bar{y})}{\Sigma (x - \bar{x})^2} = \frac{240}{40} = 6$$

en met (18.8)

$$a = \frac{\Sigma y - b \cdot \Sigma x}{m} = \frac{4150 - 6(272)}{32} = 78,7.$$

De regressievergelijking van y op x luidt dus

$$\hat{y} = 78,7 + 6x.$$

Verder volgt uit (19.3.2):

$$s_{y,x}^2 = \frac{1982,875 - \frac{(240)^2}{40}}{30} = \frac{1982,875 - 1440}{30} = 18,0958,$$

zodat

$$s_{y,x} = 4,25.$$

De berekende regressielijn is in figuur 19.2 getekend en daartoe zijn twee waarden van \hat{y} bepaald, nl. voor

$$\begin{cases} x = 7, \hat{y} = 78,7 + 6(7) = 120,7 \\ x = 10, \hat{y} = 78,7 + 6(10) = 138,7. \end{cases}$$

Ter contrôle kan dienen, dat de regressielijn door het punt $(\bar{x}, \bar{y}) = (8,5, 129,7)$ loopt.

De berekende lijn geeft een schatting van de gemiddelde toeneming van de lichaamslengte bij toenemende leeftijd. De waarde $b = 6$ leert, dat voor de beschouwde sub-populatie per jaar ongeveer een lengtegroei van 6 cm optreedt. Wij komen hierop in 19.1.3 terug, maar willen er nu reeds op wijzen, dat men niet zonder meer mag extrapoleren, d.i. de gevonden vergelijking op jongere of oudere leeftijden toepassen. Stelt men bv. in de gevonden regressievergelijking $x = 0$, dan vindt men $\hat{y} = 78,7$, wat wil zeggen dat men dus de gemiddelde lengte bij de geboorte op ongeveer 79 cm zou schatten (terwijl deze in werkelijkheid in de nabijheid van de 50 cm ligt). Deze foutieve schatting is een gevolg van het feit, dat de regressie van 0 tot ongeveer 6 jaar *niet* rechtlijnig verloopt.

19.2. Vervolgens beschouwen wij nogmaals de $n = 31$ waarnemingsparen in tabel 18.4 betreffende de regressie van y (percentage droge stof in het hart van kippeëmbryo's op x (incubatietijd in dagen). Bij dit voorbeeld kan er twijfel bestaan omtrent de juistheid van het

lineair-normale model. Het aantal waarnemingen is te klein om tot een duidelijke uitspraak omtrent de normaliteit en gelijkheid van de varianties der y -verdelingen te komen (er komen echter geen grote uitbijters, noch grote verschillen in variantie voor) en men zou hier een kromlijnige (exponentiële) regressie verwachten. Toch berekenen wij eerst de lineaire regressievergelijking van y op x om te onderzoeken, hoe goed deze 'past'. De berekeningen verlopen dan als volgt:

$$\begin{aligned}\Sigma x &= 4(9) + 3(10) + \dots + 2(20) = 422 \\ \Sigma x^2 &= 4(9)^2 + 3(10)^2 + \dots + 2(20)^2 = 6082 \\ \Sigma (x - \bar{x})^2 &= 6082 - (422)^2/31 = 6082 - 5744,6452 = 337,3548 \\ \Sigma y &= 12,10 + 12,16 + 12,26 + \dots + 16,79 + 16,87 = 439,45 \\ \Sigma y^2 &= (12,10)^2 + (12,16)^2 + \dots + (16,87)^2 = 6285,8975 \\ \Sigma (y - \bar{y})^2 &= 6285,8975 - (439,45)^2/31 = 6285,8975 - 6229,5581 \\ &= 56,3394 \\ \Sigma xy &= 9(12,10 + 12,16 + 12,26 + 12,26) + \dots + 20(16,79 + 16,87) \\ &= 6118,02 \\ \Sigma x \cdot \Sigma y/n &= (422)(439,45)/31 = 5982,1903 \\ \Sigma (x - \bar{x})(y - \bar{y}) &= 6118,02 - 5982,1903 = 135,8297\end{aligned}$$

Uit deze gegevens volgt:

$$b = \frac{\Sigma (x - \bar{x})(y - \bar{y})}{\Sigma (x - \bar{x})^2} = \frac{135,8297}{337,3548} = 0,4026,$$

en

$$a = \frac{\Sigma y - b \cdot \Sigma x}{m} = \frac{439,45 - (0,4026)(422)}{31} = 8,70.$$

De vergelijking van de regressielijn luidt dus

$$\hat{y} = 8,70 + 0,403x.$$

Tenslotte is (volgens formule 19.3.2):

$$s_{y \cdot x}^2 = \frac{56,3394 - \frac{(135,8297)^2}{337,3548}}{29} = \frac{56,3394 - 54,6893}{29} = 0,056900,$$

zodat

$$s_{y \cdot x} = 0,239.$$

De berekende lijn is in figuur 18.6 getekend. Hoewel enkele duidelijke afwijkingen van deze lijn voorkomen (bv. bij $x = 14$ en bij $x = 18$), blijkt zij verder redelijk te passen. In 19.1.5 zullen wij voor dit regressieprobleem de toets op de rechtlijnigheid toepassen.

19.3. Bij voorbeeld 18.4 inzake de regressie van de sicca-bepalingen op bloedmonsters met vijf nauwkeurig bekende gehalten kan o.i. het

lineair-normale model zonder bezwaar worden gebruikt. De berekeningen van de regressielijn van y op x leveren op:

$$\begin{aligned} \Sigma x &= 308,5, \Sigma x^2 = 3843,95, \Sigma(x-\bar{x})^2 = 3843,95 - (308,5)^2/25 = 37,06, \\ \Sigma y &= 310,5, \Sigma y^2 = 3895,61, \Sigma(y-\bar{y})^2 = 3895,61 - (310,5)^2/25 = 39,20, \\ \Sigma xy &= 3868,94, \Sigma(x-\bar{x})(y-\bar{y}) = 3868,94 - (308,5)(310,5)/25 = 37,37, \end{aligned}$$

zodat

$$b = \frac{37,37}{37,06} = 1,008365 = 1,0084,$$

en

$$a = \frac{310,5 - (1,008365)(308,5)}{25} = -0,0232.$$

De regressievergelijking luidt dus

$$\hat{y} = -0,0232 + 1,0084x.$$

Verder is

$$s^2_{y \cdot x} = \frac{39,20 - \frac{(37,37)^2}{37,06}}{23} = \frac{39,20 - 37,682593}{23} = 0,065974,$$

$$s_{y \cdot x} = 0,257.$$

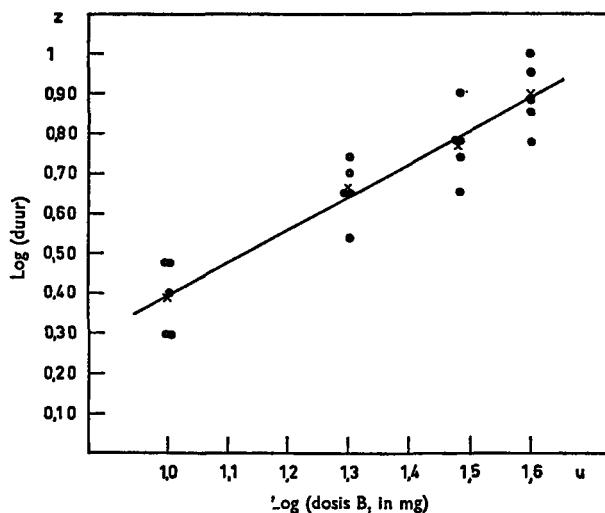
19.4. Bij voorbeeld 18.5 is reeds uit het spreidingsdiagram (in figuur 18.8) te zien, dat de spreidingen der y -verdelingen *niet* gelijk zijn, zodat het lineair-normale model niet zonder meer kan worden gebruikt. In tabel 19.2 hebben wij echter de waarden van $u = \log x$ en $z = \log y$ opgenomen en hiervan is in figuur 19.3 het spreidingsdiagram getekend.

Tabel 19.2

log (dosis) $u = \log x$	u^2	log (duur afwezigheid bradycardie) $z = \log y$					Σz_j	Σz_j^2	$u_j \Sigma z_j$
1,000	1,000000	0,301	0,301	0,398	0,477	0,477	1,954	0,794664	1,954000
1,301	1,692601	0,544	0,653	0,653	0,699	0,740	3,289	2,184955	4,278989
1,477	2,181529	0,653	0,740	0,778	0,778	0,903	3,852	2,999986	5,689404
1,602	2,566404	0,778	0,845	0,875	0,954	1,000	4,452	3,995050	7,132104
5,380	7,440534	$\Sigma u = 5(5,380) = 26,900$					13,547	9,974655	19,054497
		$\Sigma u^2 = 5(7,440534) = 37,202670$					Σz	Σz^2	Σuz

Op het oog lijkt de regressie van z op u nu rechtlijnig en de varianties van de vier verdelingen van z zijn, in volgorde van opklimmende u , resp. 0,007760, 0,005363, 0,008101 en 0,007747 en verschillen dus weinig. Omtrent de normaliteit van z valt echter ook in dit geval weinig te zeggen.

Een deel van de berekeningen, die nodig zijn voor het verkrijgen van de regressievergelijking van z op u is reeds in tabel 19.2 uitge-



N.I.P.G. 891

Figuur 19.3. Spreidingsdiagram van de gegevens in tabel 19.2 met de regressielijn $z = -0,42 + 0,8157u$.

voerd (merk op, dat deze bij gelijke n , kunnen worden vereenvoudigd!). Hieruit volgen de kwadratsommen

$$\Sigma (u - \bar{u})^2 = 37,20267 - (26,9)^2/20 = 1,022170,$$

$$\Sigma (z - \bar{z})^2 = 9,974655 - (13,547)^2/20 = 0,798595,$$

en de productsom

$$\Sigma (u - \bar{u})(z - \bar{z}) = 19,054497 - (26,9)(13,547)/20 = 0,833782,$$

zodat

$$b = \frac{0,833782}{1,022170} = 0,8157$$

en

$$a = \frac{13,547 - (0,8157)(26,9)}{20} = -0,4198 = -0,42 .$$

De regressielijn $z = -0,42 + 0,8157u$ is in figuur 19.3 getekend en past zo goed bij de waargenomen gemiddelden, dat de toets op de rechtlijnigheid hier wel achterwege kan blijven. Verder is

$$s_{z,u}^2 = \frac{0,798595 - (0,833782)^2/1,022170}{18} = 0,006582 ,$$

$$s_{z,u} = 0,0811.$$

19.5. Tenslotte beschouwen wij de waarnemingen betreffende de stikstofgehalten in het serum van ratten bij verschillende leeftijden in tabel 18.7. Hoewel uit het spreidingsdiagram in figuur 18.9 reeds

duidelijk blijkt, dat de regressie van y op x kromlijinig is, voeren wij de berekeningen voor de regressielijn $\hat{y} = a + bx$ uit om later de toets op de rechtlijngheid ook op dit voorbeeld te kunnen toepassen. Daar hier $n_j = 10$ ($m = \sum n_j = 50$) verloopt de berekening op dezelfde wijze als bij het voorgaande voorbeeld. De uitkomsten zijn:

$$\begin{aligned} \Sigma x &= 4650, \quad \Sigma x^2 = 588250, \quad \Sigma (x - \bar{x})^2 = 588250 - (4650)^2/50 = 155800, \\ \Sigma y &= 52,04, \quad \Sigma y^2 = 55,4756, \quad \Sigma (y - \bar{y})^2 = 55,4756 - (52,04)^2/50 = 1,312368, \\ \Sigma xy &= 5224, \quad \Sigma (x - \bar{x})(y - \bar{y}) = 5224 - (4650)(52,04)/50 = 384,28. \end{aligned}$$

Hieruit volgt, dat

$$b = \frac{384,28}{155800} = 0,0024665,$$

en

$$a = \frac{52,04 - (0,0024665)(4650)}{50} = 0,8114,$$

terwijl

$$s_{y \cdot x}^2 = \frac{1,312368 - (384,28)^2/155800}{48} = 0,00759465,$$

$$s_{y \cdot x} = 0,0871.$$

19.1.3. TOETSEN EN INTERVALSCHATTINGEN

Alle steekproeven van de omvang m , met dezelfde aantallen waarnemingen n_j per x_j , die uit de bestudeerde sub-populatie kunnen worden gekozen, bezitten hetzelfde gemiddelde \bar{x} en dezelfde spreiding s_x , zodat de grootheid x *niet* stochastisch is. Bij aselechte trekking van zo'n steekproef zijn echter de steekproeffuncties $s_{y \cdot x}^2$, \hat{b} en \hat{y} , stochastische grootheden, waarvan de kansverdelingen bij lineair-normale regressie kunnen worden afgeleid.

De verdeling van $s_{y \cdot x}^2$

Bij lineair-normale regressie volgt de grootheid

$$(19.4) \quad \frac{(m - 2) s_{y \cdot x}^2}{\sigma_{y \cdot x}^2}$$

een χ^2 -verdeling met $\nu = m - 2$. Men kan dan een $100(1 - \alpha)\%$ betrouwbaarheidsinterval van $\sigma_{y \cdot x}^2$ verkrijgen door te nemen

$$(19.5) \quad \frac{(m - 2) s_{y \cdot x}^2}{\chi_{1-\alpha}^2} < \sigma_{y \cdot x}^2 < \frac{(m - 2) s_{y \cdot x}^2}{\chi_{\frac{1}{2}\alpha}^2}.$$

De verdeling van \underline{b}

Bij linear-normale regressie volgt de grootheid \underline{b} een normale verdeling met gemiddelde $\mu_b = B$ en met spreiding

$$(19.6) \quad \sigma_b = \frac{\sigma_{y \cdot x}}{\sqrt{\sum (x - \bar{x})^2}},$$

zodat de grootheid

$$(19.7) \quad \underline{T} = \frac{(\underline{b} - B)}{\sigma_b} = \frac{(\underline{b} - B) \sqrt{\sum (x - \bar{x})^2}}{\sigma_{y \cdot x}}$$

standaardnormaal verdeeld is. Een zuivere schatting van σ_b is

$$(19.8) \quad s_b = \frac{s_{y \cdot x}}{\sqrt{\sum (x - \bar{x})^2}},$$

en verder geldt, dat

$$\frac{(m-2) s_b^2}{\sigma_b^2} = \frac{(m-2) s_{y \cdot x}^2}{\sigma_{y \cdot x}^2},$$

zodat volgens (13.10) de grootheid

$$(19.9) \quad \underline{t} = \frac{\frac{(\underline{b} - B) \sqrt{\sum (x - \bar{x})^2}}{\sigma_{y \cdot x}}}{\sqrt{\frac{(m-2) s_{y \cdot x}^2}{\sigma_{y \cdot x}^2} / (m-2)}} = \frac{(\underline{b} - B) \sqrt{\sum (x - \bar{x})^2}}{s_{y \cdot x}}$$

een STUDENT-verdeling volgt met $m-2$ vrijheidsgraden. Via deze grootheid kan men dus hypothesen toetsen van de vorm $B = B_0$. Toetst men de hypothese $H_0: B = 0$ (d.w.z. $\mu_{y \cdot 1} = \mu_{y \cdot 2} = \dots = \mu_{y \cdot k}$, er is geen regressie van y op x), dan gaat (19.9) over in

$$(19.10) \quad \underline{t} = \frac{\underline{b} \sqrt{\sum (x - \bar{x})^2}}{s_{y \cdot x}}.$$

Men kan een 100 $(1 - \alpha)$ % betrouwbaarheidsinterval van B verkrijgen door te nemen

$$(19.11) \quad \underline{b} + t_{\frac{1}{2}\alpha} \cdot \frac{s_{y \cdot x}}{\sqrt{\sum (x - \bar{x})^2}} < B < \underline{b} + t_{1-\frac{1}{2}\alpha} \cdot \frac{s_{y \cdot x}}{\sqrt{\sum (x - \bar{x})^2}}.$$

Toetst men voor voorbeeld 19.1 (lengte-leeftijd) de hypothese $H_0: B = 0$, dan vindt men een zeer significante uitkomst:

$$t_0 (y = 30) = \frac{6 \sqrt{40}}{4,25} = 8,93.$$

Voor dit voorbeeld vindt men als 95% betrouwbaarheidsinterval van B :

$$6 + (-2,042) \frac{4,25}{\sqrt{40}} < B < 6 + (2,042) \frac{4,25}{\sqrt{40}},$$

$$4,63 < B < 7,37.$$

Voor voorbeeld 19.3 (hemoglobinebepalingen) toetsen wij vervolgens via (19.9) de hypothese $H_0: B = 1$. Uit (19.9) volgt dan:

$$t_0 = \frac{(1,0084 - 1) \sqrt{37,06}}{0,257} = \frac{0,0511}{0,257} = 0,2.$$

H_0 kan dus niet worden verworpen en de lage waarde van t wijst er op, dat de overeenstemming tussen de waargenomen en de te verwachten waarde van B bijzonder goed is.

Uit (19.8) blijkt, dat s_b kleiner wordt naarmate n groter wordt en naarmate $\Sigma (x - \bar{x})^2$ groter wordt. Het laatste houdt in, dat men bij een bepaalde m een betrouwbaarder schatting van B verkrijgt, naarmate de spreiding van de gekozen x_j groter is. Hieruit volgt, dat het nuttig is de waarden van x zoveel mogelijk gespreid te kiezen over het traject van deze grootheid, waarvoor men de regressie van y op x wil bestuderen (zie echter voor enkele andere desiderata betreffende deze keuze 19.1.6). Als de regressie rechtlijnig is verkrijgt men de meest betrouwbare schatting van B , door twee waarden van x aan de uiteinden van het te onderzoeken traject te kiezen en bij elke waarde $\frac{1}{2}m$ waarnemingen te verrichten. Deze procedure kan men echter alleen volgen, als men a priori weet, dat de regressie lineair is, omdat men dan niet meer op rechtlijnigheid kan toetsen.

De verdeling van \hat{y}_j

Voor het lineair-normale model volgt de grootheid \hat{y}_j een normale verdeling met gemiddelde

$$(19.12) \quad \mu_{y,j} = \mu_y + B(x_j - \bar{x})$$

en variantie

$$(19.13) \quad \text{var}(\hat{y}_j) = \frac{\sigma_{y \cdot x}^2}{m} + \sigma_{y \cdot x}^2 \frac{(x_j - \bar{x})^2}{\Sigma (x - \bar{x})^2} = \sigma_{y \cdot x}^2 \left[\frac{1}{m} + \frac{(x_j - \bar{x})^2}{\Sigma (x - \bar{x})^2} \right].$$

De grootheid

$$(19.14) \quad t = \frac{\hat{y}_j - \mu_{y,j}}{s_{y \cdot x} \sqrt{\frac{1}{m} + \frac{(x_j - \bar{x})^2}{\Sigma (x - \bar{x})^2}}}$$

volgt dus een STUDENT-verdeling met $m - 2$ vrijheidsgraden. Een

100(1 - α)% betrouwbaarheidsinterval van $\mu_{y,j}$ kan worden verkregen met

$$(19.15) \quad \hat{y}_j + t_{\frac{1}{2}\alpha} \cdot s_{y \cdot x} \sqrt{\frac{1}{m} + \frac{(x_j - \bar{x})^2}{\sum (x - \bar{x})^2}} < \mu_{y,j} < \hat{y}_j + t_{1-\frac{1}{2}\alpha} \cdot s_{y \cdot x} \sqrt{\frac{1}{m} + \frac{(x_j - \bar{x})^2}{\sum (x - \bar{x})^2}}.$$

Bij voorbeeld 19.1 vindt men met (19.15) het volgende 95% betrouwbaarheidsinterval van $\mu_{y,j}$ (gebruik makend van het feit, dat $\hat{y} = 78,7 + 6x$):

$$78,7 + 6x_j - (2,042) (4,25) \sqrt{\frac{1}{32} + \frac{(x_j - 8,5)^2}{40}} < \mu_{y,j} < 78,7 + 6x_j + (2,042) (4,25) \sqrt{\frac{1}{32} + \frac{(x_j - 8,5)^2}{40}}.$$

In tabel 19.3 zijn voor $x_j = 7, 8, 8,5, 9$ en 10 de corresponderende waarden van \hat{y}_j en de beneden- en bovengrenzen van de 95% betrouwbaarheidsintervallen van de $\mu_{y,j}$ gegeven, die uit voorgaande ongelijkheid volgen.

Tabel 19.3

x_j	\hat{y}_j	95% betr. interval van $\mu_{y,j}$	
		Benedengrens	Bovengrens
7	120,7	118,1	123,3
8	126,7	125,0	128,4
8,5	129,7	128,2	131,2
9	132,7	131,0	134,4
10	138,7	136,1	141,3

In figuur 19.2 zijn de benedengrenzen en de bovengrenzen van de 95% betrouwbaarheidsintervallen van $\mu_{y,j}$ per x_j uitgezet en door gebroken lijnen verbonden. Het interval blijkt de kleinste breedte te bezitten wanneer $x_j = \bar{x}$, want dan is $(x_j - \bar{x})^2 = 0$ (terwijl $\hat{y}_j = \bar{y}$), zodat (19.15) overgaat in

$$(19.16) \quad \bar{y} + t_{\frac{1}{2}\alpha} \cdot \frac{s_{y \cdot x}}{\sqrt{m}} < \mu_y < \bar{y} + t_{1-\frac{1}{2}\alpha} \cdot \frac{s_{y \cdot x}}{\sqrt{m}}.$$

De breedte van het interval neemt toe naarmate x_j verder van \bar{x} verwijderd is; dit is begrijpelijk, daar dan de toevallige variatie van \hat{b} een steeds grotere invloed gaat uitoefenen.

Toch wordt nog vaak aangegeven dat de betrouwbaarheidsgrenzen van $\mu_{y,j}$ evenwijdig aan de waargenomen regressielijn verlopen en dat men ze kan verkrijgen door te nemen $\hat{y}_j \pm T_{\frac{1}{2}\alpha} s_{y,x} / \sqrt{m}$. Laatstgenoemde grenzen gelden slechts bij benadering voor zeer grote steekproeven.

Voor de populatie waarop het lineair-normale regressiemodel is gebaseerd (zie 19.1.1) geldt de uitspraak, dat (bv.) 95% der y -waarden bij x_j , tussen de grenzen

$$\mu_{y,j} \pm 1,96 \sigma_{y,x} \text{ [of: } (A + Bx_j) \pm 1,96 \sigma_{y,x}]$$

is gelegen. Hieruit volgt echter *niet* — zoals eveneens vaak wordt gedacht — dat men uit een steekproef van m paren (x, y) als 95% betrouwbaarheidsinterval van (de enkele waarneming) y_j kan nemen

$$\hat{y}_j \pm 1,96 s_{y,x} \text{ [of: } (a + bx_j) \pm 1,96 s_{y,x}].$$

De juiste grenswaarden zijn in dit geval:

$$\hat{y}_j \pm t_{0,025} s_{y,x} \sqrt{1 + \frac{1}{m} + \frac{(x_j - \bar{x})^2}{\sum (x - \bar{x})^2}},$$

waarin t weer $m-2$ vrijheidsgraden bezit. Voor zeer grote steekproeven kan men echter t door T vervangen en is de vorm onder het wortelteken bij benadering gelijk aan 1, zodat dan de eerder genoemde grenzen bij benadering gelden.

Stelt men $x_j = 0$, dan verkrijgt men met (19.15) een 100 $(1 - \alpha)$ % betrouwbaarheidsinterval van $\mu_{y,0} = A$, t.w.

$$(19.17) \quad a + t_{\frac{1}{2}\alpha} s_{y,x} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x - \bar{x})^2}} < A < a + t_{1-\frac{1}{2}\alpha} s_{y,x} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x - \bar{x})^2}}.$$

Tevens volgt uit het voorgaande, dat voor het lineair-normale model de grootheid a normaal verdeeld is met gemiddelde $\mu_a = A$ en met variantie

$$\sigma_a^2 = \sigma_{y,x}^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum (x - \bar{x})^2} \right], \text{ zodat de grootheid}$$

$$(19.18) \quad t = \frac{a - A}{s_{y,x} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum (x - \bar{x})^2}}}$$

een STUDENT-verdeling met $m - 2$ vrijheidsgraden volgt. Men dient zich echter te realiseren, dat de waargenomen a_0 slechts reële betekenis bezit (en voor het toetsen van hypothesen betreffende A of het geven van een schattingsinterval van A kan worden gebruikt), indien de waarde $x = 0$ binnen of dichtbij het door de gekozen waarden x_1 t/m x_k bestreken gebied ligt, resp. indien het verantwoord kan worden geacht de bestudeerde regressie tot $x = 0$ te extrapoleren. Zo weet men bv. bij voorbeeld 19.1, dat de regressie van de lengte op de leeftijd van de geboorte tot ongeveer 6 jaar *niet* rechtlijnig is, zodat

men de voor 7 t/m 10 jaar gevonden regressielijn niet mag doortrekken tot $x = 0$ en de hiervoor geldende $a_0 = 78,7$ niet kan gebruiken voor het schatten van de gemiddelde lichaamslengte bij de geboorte. Bij voorbeeld 19.2 ontbreekt elke informatie betreffende de regressie van y op x bij een incubatietijd van 0 t/m 9 dagen, zodat ook hier een schatting van A op grond van a_0 niet verantwoord is. Maar bij voorbeeld 19.3 is het aannemelijk, dat ook voor het traject van 0 t/m 10 g% de regressie van y op x rechtlijnig is met $b = 1$, zodat men bv. de hypothese $H_0: A = 0$ wel kan toetsen op basis van $a_0 = -0,0232$. Men vindt dan met (19.18)

$$t_0 = \frac{-0,0232}{\sqrt{0,065974 [1/_{25} + (12,34)^2/37,06]}} = -0,044 ,$$

zodat van het verwerpen van H_0 geen sprake is.

19.1.4. SCHATTING VAN B ALS $A = 0$

Als men weet of mag aannemen, dat $A = 0$, dan gaat voor het lineair-normale model de regressievergelijking (19.1) over in

$$(19.19) \quad \mu_{y,j} = Bx_j .$$

In dit geval kan men een zuivere schatting van B verkrijgen door te berekenen

$$(19.20) \quad b = \frac{\sum xy}{\sum x^2} ,$$

terwijl een zuivere schatting van $\sigma^2_{y \cdot x}$ volgt uit

$$(19.21) \quad s^2_{y \cdot x} = \frac{\sum y^2 - (\sum xy)^2 / \sum x^2}{m - 1} = \frac{\sum y^2 - b \cdot \sum xy}{m - 1} ,$$

waarbij dus door $m - 1$ wordt gedeeld, omdat slechts één parameter (B) behoeft te worden geschat. Nu volgt de grootheid

$$(19.22) \quad \frac{(m - 1) s^2_{y \cdot x}}{\sigma^2_{y \cdot x}}$$

een χ^2 -verdeling met $m - 1$ vrijheidsgraden. Analoog aan het behandelde in 19.1.3 vindt men, dat de grootheid

$$(19.23) \quad t = \frac{(b - B) \sqrt{\sum x^2}}{s_{y \cdot x}}$$

een STUDENT-verdeling volgt met $m - 1$ vrijheidsgraden terwijl men een $100(1 - \alpha)\%$ betrouwbaarheidsinterval van B kan verkrijgen door te nemen

$$(19.24) \quad b + t_{\frac{1}{2}\alpha} \cdot \frac{s_{y \cdot x}}{\sqrt{\sum x^2}} < B < b + t_{1-\frac{1}{2}\alpha} \cdot \frac{s_{y \cdot x}}{\sqrt{\sum x^2}} .$$

De formules, die nu voor \hat{y}_j gelden, kunnen worden verkregen door in de desbetreffende formules in 19.1.3 in plaats van $\sum (x - \bar{x})^2$ te nemen $\sum x^2$.

Beschouw voorbeeld 18.1 betreffende de colorimeteraflezingen (y) bij verschillende concentraties van α -resin (x). Hier mag worden aangenomen, dat de regressielijn van y op x door de oorsprong loopt, zodat $A=0$. Voor de 6 waarnemingen geldt: $\sum x = 7,75$, $\sum x^2 = 13,6635$, $\sum y = 562$, $\sum y^2 = 71890$ en $\sum xy = 991,01$, zodat volgens (19.20)

$$b = \frac{991,01}{13,6635} = 72,53$$

en volgens (19.21)

$$s^2_{y,x} = \frac{71890 - (991,01)^2/13,6635}{5} = 2,4620 ,$$

$$s_{y,x} = 1,57 .$$

Het 95% betrouwbaarheidsinterval van B is dus

$$72,53 - (2,571) \frac{1,57}{\sqrt{13,6635}} < B < 72,53 + (2,571) \frac{1,57}{\sqrt{13,6635}} ,$$

$$71,4 < B < 73,6.$$

19.1.5. TOETS OP DE RECHTLIJNIGHEID VAN DE REGRESSIE

Wanneer in de deelverzameling (zie 19.1.1), die bij een regressieprobleem wordt bestudeerd, de regressie van y op x lineair is, liggen de gemiddelden $\mu_{y,j}$ op de regressielijn $\mu_{y,j} = \mu_y + B(x_j - \bar{x})$. De gemiddelden \bar{y}_j van de bij de waarden x_j getrokken steekproeven zullen echter tengevolge van toevallige variatie niet precies gelijk zijn aan hun parameters en zij liggen dus niet (alle) op de regressielijn $\hat{y}_j = \bar{y} + b(x_j - \bar{x})$. Het zijn nu de verschillen $\bar{y}_j - \hat{y}_j$ die informatie verschaffen omtrent de wijze, waarop deze regressielijn 'past' en wel door middel van een variantie-analyse.

Hiertoe beginnen wij de totale kwadratensom $\sum (y - \bar{y})^2$ volgens 17.2.1 in twee componenten te splitsen:

$$(19.25) \quad \sum_{j=1}^k \sum_{i=1}^{n_j} (y_{ij} - \bar{y})^2 = \sum_{j=1}^k n_j (\bar{y}_j - \bar{y})^2 + \sum_{j=1}^k \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2 ,$$

of:

$$S_{\text{totaal}} (m-1 \text{ vr. gr.}) = S_{\text{tussen}} (k-1 \text{ vr. gr.}) + S_{\text{binnen}} (m-k \text{ vr. gr.}) .$$

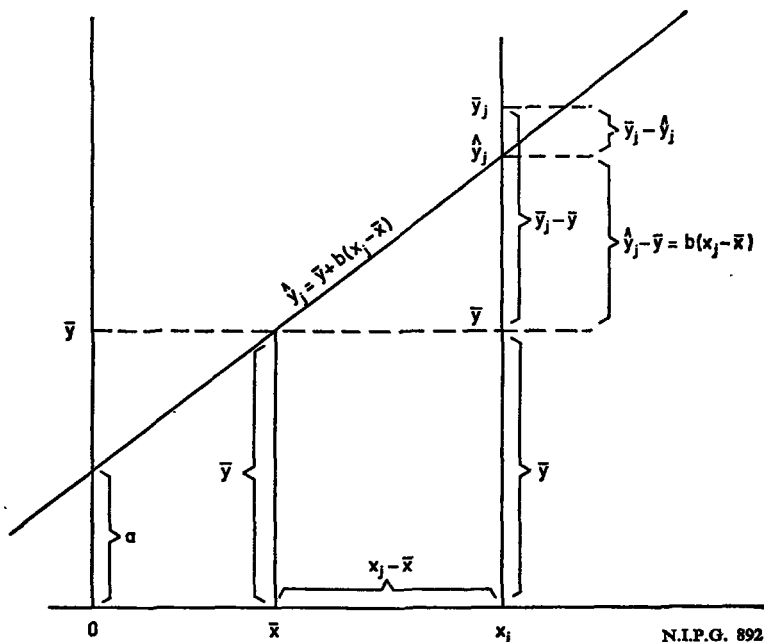
Zoals wij reeds eerder hebben gezien kan men deze kwadratensommen

het best berekenen door te nemen:

$$S_{\text{totaal}} = \Sigma y^2 - \frac{(\Sigma y)^2}{m} \quad (m = \Sigma n_j)$$

$$\begin{aligned} S_{\text{tussen}} &= \frac{(\Sigma y_1)^2}{n_1} + \frac{(\Sigma y_2)^2}{n_2} + \dots + \frac{(\Sigma y_k)^2}{n_k} - \frac{(\Sigma y)^2}{m} \\ &= \sum_{j=1}^k \frac{(\Sigma y_j)^2}{n_j} - \frac{(\Sigma y)^2}{m} \end{aligned}$$

$$S_{\text{binnen}} = S_{\text{totaal}} - S_{\text{tussen}} .$$



Figuur 19.4. Lineaire regressie: $\hat{y}_j = (\bar{y}_j - \hat{y}_j) + (\hat{y}_j - \bar{y}) + \bar{y}$.

Wanneer er een lineaire regressie van y op x aanwezig is, dan geldt echter (zie figuur 19.4):

$$\bar{y}_j - \bar{y} = (\bar{y}_j - \hat{y}_j) + (\hat{y}_j - \bar{y})$$

of [zie formule (19.2)]

$$(19.26) \quad \bar{y}_j - \bar{y} = (\bar{y}_j - \hat{y}_j) + b(x_j - \bar{x}).$$

Hierin representeert $b(x_j - \bar{x})$ de deviatie tengevolge van de regressie en $\bar{y}_j - \hat{y}_j$, de deviatie ten opzichte van de regressielijn. Uit (19.26)

volgt, dat S_{tussen} als volgt kan worden gesplitst:

$$\begin{aligned} \sum_{j=1}^k n_j (\bar{y}_j - \bar{y})^2 &= \sum_{j=1}^k n_j [(\bar{y}_j - \hat{y}_j) + b(x_j - \bar{x})]^2 \\ &= \sum_{j=1}^k n_j [(\bar{y}_j - \hat{y}_j)^2 + 2b(\bar{y}_j - \hat{y}_j)(x_j - \bar{x}) + b^2(x_j - \bar{x})^2] . \end{aligned}$$

Daar de som van de dubbele producten binnen de grote haken gelijk is aan nul en daar

$$\sum_{j=1}^k n_j b^2 (x_j - \bar{x})^2 = b^2 \cdot \Sigma (x - \bar{x})^2$$

kan men dus schrijven

$$(19.27) \quad \sum_{j=1}^k n_j (\bar{y}_j - \bar{y})^2 = \sum_{j=1}^k n_j (\bar{y}_j - \hat{y}_j)^2 + b^2 \cdot \Sigma (x - \bar{x})^2 .$$

Op deze wijze is bij lineaire regressie S_{tussen} dus eveneens in twee componenten te splitsen. De tweede component, die wij met $S_{\text{regressie}}$ aanduiden, bezit 1 vrijheidsgraad en representeert het deel van de variatie tussen de gemiddelden, dat ‘verklaard’ kan worden door de regressie van y op x . Deze kwadratensom kan het best worden berekend door te nemen

$$S_{\text{regressie}} = b^2 \cdot \Sigma (x - \bar{x})^2 = \frac{[\Sigma (x - \bar{x})(y - \bar{y})]^2}{\Sigma (x - \bar{x})^2} .$$

De eerste component, die wij $S_{\text{gemiddelden}}$ noemen, bezit $k - 2$ vrijheidsgraden en betreft de resterende variatie van de gemiddelden om de regressielijn.

De verschillende kwadratensommen zijn in tabel 19.4 samengevat en hierbij is tevens aangegeven, hoe en in welke volgorde hun berekening verloopt en welke variantieschattingen zij opleveren.

Tabel 19.4. Kwadratensommen bij lineaire regressie

	Kwadratensom		Vrijh. graden	Volgorde berekening	Variantie-schatting
	Formule	Berekening			
Regressie	$b^2 \Sigma (x - \bar{x})^2$	$\frac{[\Sigma (x - \bar{x})(y - \bar{y})]^2}{\Sigma (x - \bar{x})^2}$	1	4	$s^2_{\text{regressie}}$
Gemiddelden om regressie	$\sum_{j=1}^k n_j (\bar{y}_j - \hat{y}_j)^2$	$S_{\text{tussen}} - S_{\text{regressie}}$	$k - 2$	5	$s^2_{\text{gemiddelden}}$
Tussen de steekproeven	$\sum_{j=1}^k n_j (\bar{y}_j - \bar{y})^2$	$\sum_{j=1}^k \frac{(\Sigma y_j)^2}{n_j} - \frac{(\Sigma y)^2}{m}$	$k - 1$	2	s^2_{tussen}
Binnen de steekproeven	$\sum_{j=1}^k \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2$	$S_{\text{totaal}} - S_{\text{tussen}}$	$m - k$	3	s^2_{binnen}
Totaal	$\sum_{j=1}^k \sum_{i=1}^{n_j} (y_{ij} - \bar{y})^2$	$\Sigma y^2 - \frac{(\Sigma y)^2}{m}$	$m - 1$	1	

Onder de hypothese, dat de regressie van y op x lineair is, zijn s^2_{gem} en s^2_{binnen} zuivere schattingen van $\sigma^2_{y,x}$, zodat dan de grootheid

$$(19.28) \quad \underline{F} = s^2_{\text{gem}}/s^2_{\text{binnen}}$$

een F -verdeling volgt met $\nu_1 = k - 2$ en $\nu_2 = m - k$. Men kan deze hypothese bij een drempelwaarde α toetsen door $F_0 = s^2_{\text{gem}}/s^2_{\text{binnen}}$ te berekenen en H_0 kan dan worden verworpen, als

$$F_0 \geq F_{1-\alpha}(\nu_1, \nu_2), \text{ d.i. voor } P_R(F_0) \leq \alpha.$$

Opmerking. Als het lineair-normale model geldig is, levert (19.3) een zuivere schatting van σ^2 . Uit tabel 19.4 blijkt, dat men dan ook kan nemen

$$(19.29) \quad s^2 = \frac{S_{\text{totaal}} - S_{\text{regressie}}}{m - 2}$$

De hypothese, dat de waargenomen regressie van y op x toevallig is ($B = 0$) kan ook worden getoetst via de grootheid

$$(19.30) \quad \underline{F} = \frac{s^2_{\text{regressie}}}{s^2}$$

met $\nu_1 = 1$ en $\nu_2 = m - 2$. Deze toets is identiek aan de toets van STUDENT in 19.1.3, daar geldt dat F [volgens formule 19.30] = t^2 [volgens formule 19.9].

Wij gaan nu deze toets op de rechtlijnigheid van de regressie van y op x op enkele der voorgaande voorbeelden toepassen. Bij voorbeeld 19.1 (regressie van lengte op leeftijd bij jongens van 7, 8, 9 en 10 jaar) is (zie tabel 19.1):

$$\begin{aligned} S_{\text{tussen}} &= \frac{(969)^2 + (1000)^2 + (1078)^2 + (1103)^2}{8} - \frac{(4150)^2}{32} \\ &= 539706,75 - 538203,125 = 1503,625 \end{aligned}$$

en

$$S_{\text{regressie}} = \frac{(240)^2}{40} = 1440,$$

zodat de kwadratensommen en variantieschattingen zijn:

	S	Vrijheidsgraden	Var. schatting
Regressie	1440	1	1440
Gemiddelden om regressie	63,625	2	31,8125
Tussen de steekproeven	1503,625	3	
Binnen de steekproeven	479,250	28	17,1161
Totaal	1982,875	31	

Hieruit volgt: $F_0 = 31,8125/17,1161 = 1,86$ met $\nu_1 = 2$ en $\nu_2 = 28$. De hypothese, dat de regressie van y op x rechtlijnig is, kan dus niet worden verworpen, daar uit tabel F blijkt dat $F_{0,95} \cong 3,35$.

Vervolgens beschouwen wij voorbeeld 19.2 (zie tabel 18.4 en de berekeningen op blz. 494):

$$S_{\text{tussen}} = \frac{(48,78)^2}{4} + \frac{(38,34)^2}{3} + \dots + \frac{(33,66)^2}{2} - \frac{(439,45)^2}{31}$$

$$= 6285,1434 - 6229,5581 = 55,5853,$$

$$S_{\text{regressie}} = \frac{(135,8297)^2}{337,3548} = 54,6893,$$

zodat de kwadratensommen en variantieschattingen zijn:

	S	Vrijheidsgraden	Var. schatting
Regressie	54,6893	1	54,6893
Gemiddelden om regressie	0,8960	9	0,0996
Tussen de steekproeven	55,5853	10	0,0377
Binnen de steekproeven	0,7541	20	
Totaal	56,3394	30	

$F_0 = 0,0996/0,0377 = 2,64$; daar $F_{0,95}(9, 20) = 2,39$ wordt de hypothese dat de regressie linear is bij een 5% drempel verworpen.

Tenslotte passen wij de toets toe op voorbeeld 19.5 (zie tabel 18.7 en de berekeningen op blz. 497):

$$S_{\text{tussen}} = \frac{(7,90)^2 + (9,84)^2 + \dots + (12,09)^2}{10} - \frac{(52,04)^2}{50}$$

$$= 55,259700 - 54,163232 = 1,096468,$$

$$S_{\text{regressie}} = \frac{(384,28)^2}{155800} = 0,947825.$$

De kwadratensommen en variantieschattingen zijn dus:

	S	Vrijheidsgraden	Var. schatting
Regressie	0,947825	1	0,947825
Gemiddelden om regressie	0,148643	3	0,049548
Tussen de steekproeven	1,096468	4	0,004798
Binnen de steekproeven	0,215900	45	
Totaal	1,312368	49	

Hieruit volgt, dat $F_0 = 0,049548/0,004798 = 10,3$, terwijl $F_{0,995}$ met $\nu_1 = 3$ en $\nu_2 = 45$ ongeveer gelijk is aan 4,9. De hypothese, dat de regressie lineair is wordt dus verworpen.

19.1.6. VEREENVOUDIGING VAN DE BEREKENINGEN

De berekeningen van de regressie-analyse kunnen worden vereenvoudigd, wanneer men de k waarden x_j equidistant kiest en per x_j eenzelfde aantal waarnemingen n van de grootheid y verricht. Hierbij kunnen twee gevallen worden onderscheiden:

a. k is oneven

In dit geval vervangt men de grootheid x door een hulpgrootheid u , waarvan het nulpunt bij de middelste waarde van x ligt en die bij iedere hogere (lagere) waarde van x met 1 opklimt (afdaalt): zie tabel 19.5 met $k = 7$. Dan is $u = (x - \bar{x})/c$ en $x = \bar{x} + cu$, waarin c de constante afstand tussen de waarden van x is.

Tabel 19.5. Regressie-analyse met equidistante waarden van x en n waarnemingen per x_j (fictieve gegevens)

1	2	3	4					5	6
x_j	u_j	u_j^2	y					Σy_j	$u_j \Sigma y_j$
3	-3	9	3	4	5	5	8	25	-75
6	-2	4	4	6	8	8	9	35	-70
9	-1	1	6	7	8	9	10	40	-40
12	0	0	7	7	9	10	12	45	0
15	+1	1	8	10	12	12	13	55	+55
18	+2	4	11	14	14	15	16	70	+140
21	+3	9	14	15	16	16	19	80	+240
	0	28						350	+250
	Σu_j	Σu_j^2						Σy	$\Sigma u_j y$

Daar de grootheid u zodanig is gekozen, dat $\bar{u} = 0$ en er per x_j , dus ook per u_j , n waarnemingen verricht zijn, geldt

$$\Sigma (x - \bar{x})^2 = n \sum_{j=1}^k (x_j - \bar{x})^2 = nc^2 \sum_{j=1}^k (u_j - \bar{u})^2 = nc^2 \sum_{j=1}^k u_j^2 = nc^2 \Sigma u_j^2$$

en

$$\Sigma (x - \bar{x})(y - \bar{y}) = c \cdot \Sigma (u - \bar{u})(y - \bar{y}) = c \left[\Sigma u y - \frac{\Sigma u \cdot \Sigma y}{kn} \right] = c \cdot \Sigma u y.$$

Hieruit volgt:

$$(19.31) \quad b = \frac{c \cdot \Sigma u y}{nc^2 \Sigma u_j^2} = \frac{\Sigma u y}{nc \cdot \Sigma u_j^2}.$$

Men vindt Σuy door per regel u_j met Σy_j te vermenigvuldigen en de verkregen producten te sommeren (zie de kolommen 5 en 6) en Σu^2 is eveneens eenvoudig te berekenen (zie kolom 3). Voor het voorbeeld is

$$\text{dus } b = \frac{250}{5 \times 3 \times 28} = 0,6, \text{ zodat } a = \frac{350}{35} - 0,6(12) = 2,8.$$

De regressievergelijking van y op x luidt: $\hat{y} = 2,8 + 0,6x$.

b. k is even

Nu kiest men de hulpgrootheid u zodanig, dat $u = \frac{2(x - \bar{x})}{c}$ en $x = \bar{x} + \frac{1}{2}cu$, bv. bij $k = 8$:

x_j	3	6	9	12	15	18	21	24
u_j	-7	-5	-3	-1	+1	+3	+5	+7

De lezer kan zelf verifiëren dat nu geldt

$$(19.32) \quad b = \frac{2 \Sigma uy}{nc \cdot \Sigma u_j^2}.$$

Voorbeeld

19.6. De gegevens in tabel 19.6 zijn ontleend aan BLISS (130) en zij betreffen een onderzoek van P. R. BURKHOLDER.

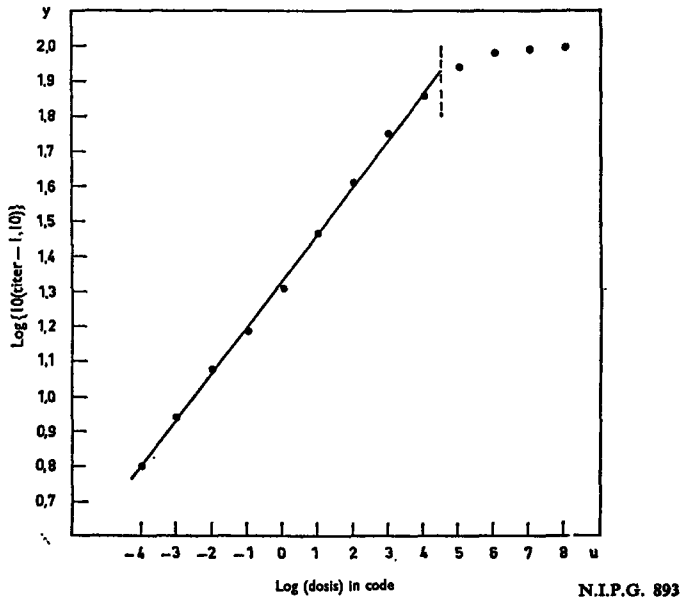
Tabel 19.6. Titers betreffende de groei van *Lactobacillus arabinosis* in 72 uur bij verschillende doses van niacine

Dosis $\mu\text{g}/$ buisje	Log dosis x	Titers in drie buisjes			u_j	$y = \log\{10(\text{titer} - 1,10)\}$			Σy_j	$u_j \Sigma y_j$	\hat{y}_j
0		1,10	1,08	1,13							
0,0250	-1,6021	1,71	1,78	1,70	-4	0,79	0,83	0,78	2,40	-9,60	0,80
0,0354	-1,4515	1,98	1,95	1,98	-3	0,94	0,93	0,94	2,81	-8,43	0,94
0,0500	-1,3010	2,23	2,39	2,31	-2	1,05	1,11	1,08	3,24	-6,48	1,08
0,0707	-1,1505	2,68	2,68	2,60	-1	1,20	1,20	1,18	3,58	-3,58	1,19
0,1000	-1,0000	3,02	3,23	3,23	0	1,28	1,33	1,33	3,94	0	1,31
0,1414	-0,8495	4,07	4,03	4,00	+1	1,47	1,47	1,46	4,40	4,40	1,47
0,2000	-0,6990	5,27	5,15	5,10	+2	1,62	1,61	1,60	4,83	9,66	1,61
0,2828	-0,5485	6,74	6,73	6,74	+3	1,75	1,75	1,75	5,25	15,75	1,75
0,4000	-0,3979	8,45	8,35	8,38	+4	1,87	1,86	1,86	5,59	22,36	1,86
0,5657	-0,2474	9,78	9,86	9,73		1,94	1,94	1,94			1,94
0,8000	-0,0969	10,70	10,84	10,76		1,98	1,99	1,98			1,98
1,1314	+0,0536	10,92	10,89	10,56		1,99	1,99	1,98			1,99
1,6000	+0,2041	10,94	11,13	11,17		1,99	2,00	2,00			2,00

$$\Sigma u_j = 0 \quad \Sigma u_j^2 = 2(16 + 9 + 4 + 1) = 60 \quad \Sigma y = 36,04 \quad \Sigma uy = 24,08$$

Bij vele doseringsproeven blijkt de regressie van de responsie (resp.

van de log responsie) op de *logarithme van de dosis* rechtlijnig te zijn.¹ Bij deze proef vormen dan ook de logarithmen der doses een equidistante reeks, waarin elke dosis met $\sqrt{2}$ is vermenigvuldigd om de naasthogere dosis te verkrijgen. Bij iedere dosis zijn drie bepalingen verricht in drie verschillende reageerbuisjes. Na de incubatieperiode van 72 uur zijn de bacillen door verhitting gedood en is ieder buisje getitreerd met 0,1 N NaOH. Dit proces leverde de titers op, die links in tabel 19.6 zijn opgenomen. Daar de regressie van de titers op de log dosis uitgesproken kromlijinig is, zijn vervolgens de waarden van $y = \log \{10 (\text{titer} - 1,10)\}$ berekend [van elke waargenomen titer is het gemiddelde van de drie contrôletiters, $(1,10 + 1,08 + 1,13)/3 = 1,10$ afgetrokken en het verschil is met 10 vermenigvuldigd om negatieve waarden van y te vermijden]. In de plaats van de grootheid $x = \log$



Figuur 19.5. Dosis-werkingskromme van niacine (zie tabel 19.6).

dosis is de hulpgrootheid u ingevoerd en in figuur 19.5 zijn de punten (u, y) getekend. Hoewel de regressie van y op x ook nu niet geheel rechtlijnig is, blijkt wèl de responsie voor de negen doses van $0,0250 \mu\text{g}$ tot en met $0,4000 \mu\text{g}$ ongeveer rechtlijnig te zijn. Voor dit deel van de gegevens verloopt de berekening van de regressievergelijking van y op x als volgt (zie tabel 19.6 en bedenk dat $n = 3$ en $c = 0,1505$):

$$b = \frac{24,08}{(3)(0,1505)(60)} = 0,8889 \quad (\text{formule 19.31}).$$

¹ Zie ook hoofdstuk 23.

Daar het gemiddelde van de logarithmen van de doses gelijk is aan $\bar{x} = \log(0,1000) = -1$ en $\bar{y} = 36,04/27 = 1,3348$, is

$$a = 1,3348 - (0,8889)(-1) = 2,2237.$$

De regressievergelijking van $y = \log\{10(\text{titer} - 1,10)\}$ op $x = \log$ dosis luidt dus: $\hat{y} = 2,2237 + 0,8889x$.

Om de regressielijn in figuur 19.5 eenvoudig te kunnen tekenen, dient men de regressievergelijking van y op u (de gecodeerde log dosis) te berekenen. Hiervoor geldt, daar nu $c = 1$:

$$b' = \frac{24,08}{(3)(60)} = 0,1338$$

en daar $\bar{u} = 0$ vindt men dat

$$a = 1,3348.$$

Met deze regressievergelijking, $\hat{y} = 1,3348 + 0,1338u$, is de regressielijn in figuur 19.5 getekend.

Volledigheidshalve toetsen wij nu nog op rechtlijnigheid. Daar $\Sigma y^2 = (0,79)^2 + (0,83)^2 + \dots + (1,86)^2 = 51,3366$ en $\Sigma y = 36,04$, is $\Sigma(y - \bar{y})^2 = 51,3366 - (36,04)^2/27 = 51,3366 - 48,106726 = 3,229874$. De kwadratensom tussen de steekproeven is

$$\frac{(2,40)^2 + (2,81)^2 + \dots + (5,59)^2}{3} - \frac{(36,04)^2}{27} = 3,224341$$

en de kwadratensom voor de regressie is $(24,08)^2/(3)(60) = 3,221369$, zodat de volgende tabel kan worden samengesteld:

	S	Vrijheids- graden	Variantie schatting
Regressie	3,221369	1	
Gemiddelden om regressie	0,002972	7	0,000425
Tussen de steekproeven	3,224341	8	
Binnen de steekproeven	0,005533	18	0,000307
Totaal	3,229874	26	

Hieruit blijkt, dat $F_0 = 18(0,002972)/7(0,005533) = 1,38$. Daar $F_{0,95}$ met $\nu_1 = 7$ en $\nu_2 = 18$ gelijk is aan 2,58 kan de hypothese, dat de regressie van y op x rechtlijnig is bij een 5% drempel niet worden verworpen.

19.2. Uitbreiding van de regressie-analyse

In 18.2.3 is reeds aangegeven, welke berekeningsmethoden men kan volgen als de regressie van y op x *kromlijnig* is. Daarbij is gebleken, dat kromlijnige regressieproblemen moeilijker te hanteren zijn dan

rechtlijnige en het is daarom begrijpelijk dat men in de praktijk bij het optreden van kromlijnigheid eerst zal trachten of deze door een transformatie van x en/of y kan worden opgeheven. Verschillende voorbeelden hebben laten zien, dat dit veelvuldig lukt.

De regressie-analyse kan ook worden uitgebreid tot meer dan twee grootheden. Bij deze zg. *meervoudige regressie-analyse* kan men bv. de regressievergelijking van y op de grootheden x_1 en x_2 berekenen. Een inleiding tot deze methode kan men vinden in SNEDECOR (49) en WALKER en LEV (35) en een uitvoeriger bespreking in BENNETT en FRANKLIN (4).

Tenslotte noemen wij nog de *covariantie-analyse*, die ahw. een synthese van de regressie-analyse en de variantie-analyse vormt. Deze methode kan o.m. worden toegepast, als men beschikt over k aselechte steekproeven, ieder bestaande uit n waarnemingsparen (x, y) . Als er een lineaire regressie van y op x aanwezig is, kan men de gemiddelden \bar{y} , vergelijken, rekening houdend met het feit dat de gemiddelden \bar{x} , van de steekproeven verschillend zijn. Denk in dit verband bv. aan een proef, waarbij men de invloed van drie verschillende diëten op de groei van ratten wil nagaan. Men stelt aselekt drie groepen van n ratten samen en geeft elke groep gedurende zekere tijd één dieet. Nu kan men wel elk dier dezelfde hoeveelheid voedsel voorzetten, maar dit garandeert niet, dat zij ook alle evenveel eten. Wanneer men nu de hoeveelheid voeding die werkelijk wordt opgenomen registreert, kan men met behulp van de covariantie-analyse de gemiddelde gewichtstoename van de groepen vergelijken met uitschakeling van de eventuele verschillen in de opgenomen hoeveelheden voedsel. Voor een beschrijving van deze techniek verwijzen wij naar DIXON en MASSEY (8), SNEDECOR (49), EDWARDS (55) of WALKER en LEV (35).

19.3. Opgaven

19.1. Uit de cursus: Biostatistics II, John Hopkins University, Balt., U.S.A. (gegevens iets gewijzigd).

One of the factors used in measuring pollution of water is the oxygen demand, which is a measure of the oxygen used by the natural biological agencies in the breaking down of organic material in the water. The following table shows the values of oxygen demand, for the first stage of deoxygenation, in one series of tests on Ohio River water (Pub. Health Bull. 173, p. 112, 1927).

Results of Incubation tests, at 20°C. x = period of incubation (days), y = oxygen demand (parts per million). Sept. 25 et seq.											
x	y	x	y	x	y	x	y	x	y	x	y
1	0,85	2	1,25	3	1,50	5	2,13	6	2,13	7	2,55
1	0,97	2	1,12	4	1,74	5	1,86	6	2,20	8	2,64
1	1,00	3	1,61	4	1,60	5	1,95	7	2,39	8	2,70
2	1,17	3	1,43	4	1,80	6	2,36	7	2,50	8	2,60
										9	2,89
										9	3,01

- Teken het spreidingsdiagram van deze waarnemingsparen.
- Bereken de vergelijking van de best-passende regressielijn van y op x en teken deze in het spreidingsdiagram.
- Bereken $s^2_{y,x}$ en bepaal het 95% betrouwbaarheidsinterval van B . Onder welke voorwaarden bezit dit interval geldigheid?
- Toets de hypothese, dat de regressie rechtlijnig is ($\alpha = 0,05$).

19.2. Betreft opgave 12.2 (blz. 291). Kies bij de toetsingen $\alpha = 0,05$:

- Teken het spreidingsdiagram.
- Bepaal de vergelijking van de best-passende regressielijn van y op x . Teken deze in het spreidingsdiagram.
- Toets de hypothese $B = 0$.
- Toets de hypothese $\sigma_{y,1} = \sigma_{y,2} = \sigma_{y,4} = \sigma_{y,8}$.
- Toets op rechtlijnigheid.

19.3. Naar PERRY (186).

Een groep van 60 muizen met gewichten tussen 13 en 15 gr wordt aselekt in 6 subgroepen van 10 dieren gesplitst. De dieren krijgen vervolgens een intrave-neuze injectie met 0,25 cc van een oplossing van neoarsphenamine, te beginnen met 5,5 pct voor de eerste subgroep en met 0,5 pct opklimmend t/m 8 pct. De overlevingstijden (in minuten) en de logaritmen van deze tijden per subgroep zijn in onderstaande tabel opgenomen [s = sterkte van de oplossing, x = log (sterkte).]

1		2		3		4		5		6	
s	x	s	x	s	x	s	x	s	x	s	x
5,5	0,740	6,0	0,778	6,5	0,813	7,0	0,845	7,5	0,875	8,0	0,903
314	2,497	195	2,290	179	2,253	178	2,250	53	1,724	60	1,778
353	2,548	243	2,386	198	2,297	178	2,250	108	2,033	80	1,903
358	2,554	260	2,415	243	2,386	210	2,322	137	2,137	87	1,940
358	2,554	293	2,467	258	2,412	232	2,365	172	2,236	113	2,053
372	2,571	315	2,498	258	2,412	242	2,384	181	2,258	116	2,064
417	2,620	327	2,515	287	2,458	272	2,435	181	2,258	116	2,064
432	2,635	345	2,538	307	2,487	281	2,449	201	2,303	145	2,161
502	2,701	360	2,556	317	2,501	301	2,479	246	2,391	188	2,274
551	2,741	427	2,630	351	2,545	310	2,491	255	2,407	208	2,318
631	2,800	484	2,685	386	2,587	340	2,531	260	2,415	216	2,334

- Teken het spreidingsdiagram van de overlevingstijd (in minuten) en de sterkte (in pct). Wat blijkt daaruit?
- Als a , maar voor $\log(\text{overlevingstijd}) = y$ en $\log(\text{sterkte}) = x$.

Voer vervolgens op de getransformeerde gegevens de volgende analyse uit:

- Onderzoek, of de varianties van de groepen homogeen zijn.
- Bepaal de vergelijking van de best-passende regressielijn van y op x . Teken deze lijn in het spreidingsdiagram b .
- Voer de toets op de rechtlijnigheid van de regressie uit. Kies bij de toetsingen $\alpha = 0,05$.

CORRELATIEREKENING

20.1. Tweedimensionaal normale correlatie

20.1.1. BESCHRIJVING VAN HET MODEL

Beschouw een populatie, waarvan elk element een waardenpaar (x, y) bezit en waarbij de simultane verdeling van de grootheden x en y tweedimensionaal normaal is. De verdeling van de grootheid x is dan normaal, met gemiddelde μ_x en spreiding σ_x en ook de grootheid y is normaal verdeeld, met gemiddelde μ_y en spreiding σ_y .

Bij zo'n tweedimensionaal normale verdeling blijkt de regressie van y op x rechtlijnig te zijn, evenals de regressie van x op y . De vergelijking van de regressielijn van y op x kan als volgt worden geschreven (zie 19.1.1)

$$(20.1.1) \quad \mu_{y \cdot x} = A_{yx} + B_{yx} x_j,$$

of (zie 18.3.2)

$$(20.1.2) \quad \mu_{y \cdot x} = \mu_y + \rho \frac{\sigma_y}{\sigma_x} (x_j - \mu_x).$$

Voor de vergelijking van de regressielijn van x op y geldt

$$(20.2.1) \quad \mu_{x \cdot y} = A_{xy} + B_{xy} y_i,$$

of

$$(20.2.2) \quad \mu_{x \cdot y} = \mu_x + \rho \frac{\sigma_x}{\sigma_y} (y_i - \mu_y).$$

Er zijn steeds twee regressielijnen, behoudens in het geval dat $\rho = +1$ of -1 : dan vallen beide lijnen samen. Deze lijnen snijden elkaar in het punt (μ_x, μ_y) . Uit het voorgaande blijkt verder, dat een tweedimensionaal normale verdeling volledig gespecificeerd kan worden door vijf parameters, t.w. μ_x , μ_y , σ_x , σ_y en ρ_{xy} .

Bij elke waarde van x treft men een normale verdeling van y -waarden aan met als variantie

$$(20.3) \quad \sigma_{y \cdot x}^2 = \sigma_y^2 (1 - \rho^2).$$

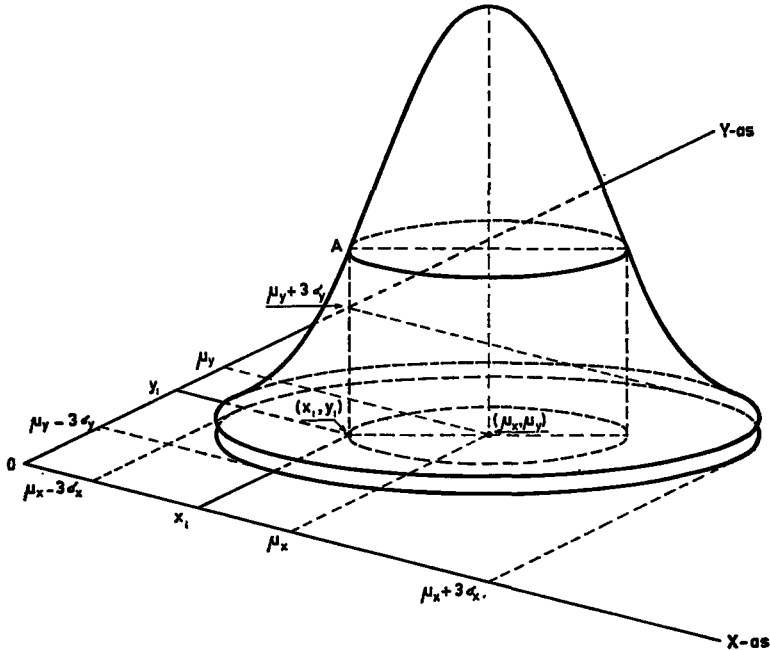
Hieruit volgt, dat $\sigma_{y \cdot x}^2 = \sigma_y^2$ als $\rho = 0$ en dat $\sigma_{y \cdot x}^2 = 0$ als $\rho^2 = 1$. Bij elke waarde van y behoort een normale verdeling van waarden van

x , die als variantie heeft

$$(20.4) \quad \sigma_{x,y}^2 = \sigma_x^2 (1 - \rho^2),$$

zodat $\sigma_{x,y}^2 = \sigma_x^2$ als $\rho = 0$ en $\sigma_{x,y}^2 = 0$ als $\rho^2 = 1$. Uit (20.3) en (20.4) volgt, dat men ook kan schrijven

$$(20.5) \quad \rho^2 = 1 - \frac{\sigma_{y,x}^2}{\sigma_y^2} = 1 - \frac{\sigma_{x,y}^2}{\sigma_x^2}.$$



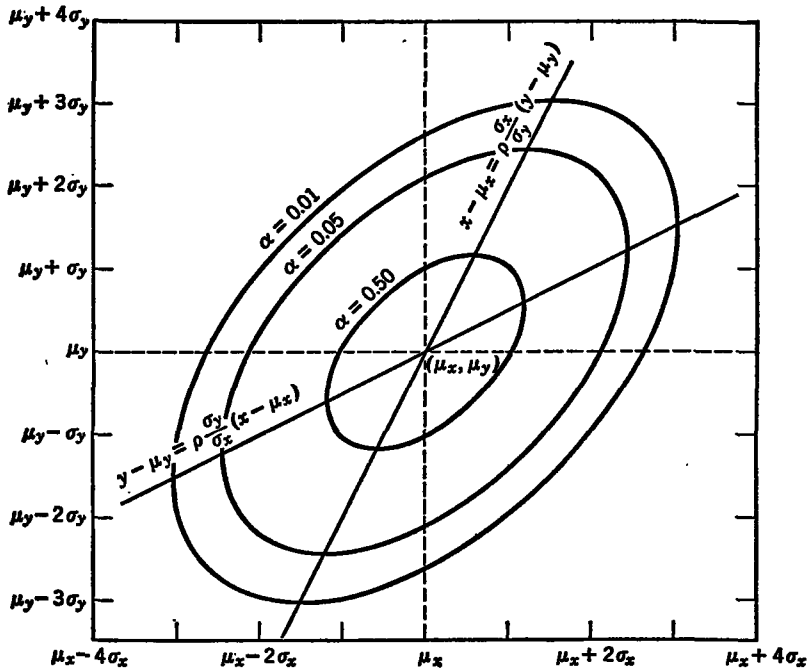
N.I.P.G. 894

Figuur 20.1. Tweedimensionaal normale verdeling.

Beschouw vervolgens de tweedimensionaal normale verdeling in figuur 20.1. De 'berg' in deze tekening is een kansverdeling en de kans van een bepaalde combinatie (x_1, y_1) kan worden bepaald, door in het punt (x_1, y_1) een loodlijn op te richten. Snijdt deze het bergoppervlak in het punt A , dan is de hoogte van dit punt boven het grondvlak een aanduiding van de kans van deze combinatie. De horizontale ellips op het bergoppervlak, die door het punt A gaat, geeft op deze wijze alle combinaties van x - en y -waarden met gelijke kans aan en alle lager op het bergoppervlak gelegen punten duiden combinaties met kleinere kansen aan. Onder de overschrijdskans van de combinatie (x_1, y_1) verstaat men nu de som van de kansen van alle combinaties, die een hoogstens even grote kans bezitten als deze combinatie. Deze kans kan

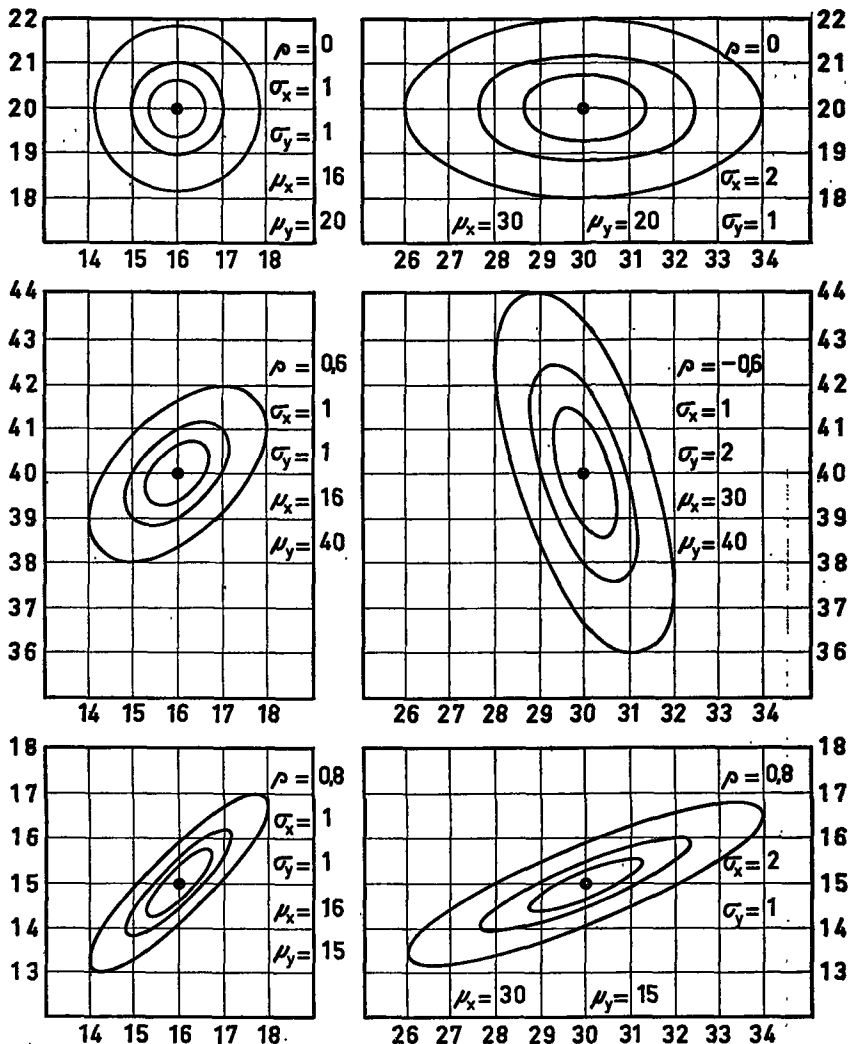
dus worden bepaald (zie figuur 20.1) door de inhoud te bepalen van dat gedeelte van de berg, dat gelegen is buiten de getekende mantel van de elliptische cylinder, gevormd door de ordinaten die behoren bij de horizontale ellips bij A .

Als men nu de ellipsen aanbrengt van de combinaties met een overschrijdingskans van bv. 0,50, 0,05 en 0,01 en deze vervolgens op het grondvlak projecteert, verkrijgt men de tweedimensionale voorstelling van een tweedimensionaal normale verdeling, zoals deze in figuur 20.2 getekend is. In deze tekening zijn tevens de beide regressielijnen aangebracht.



Figuur 20.2. Een tweedimensionaal normale verdeling in het platte vlak.

Om de lezer een indruk te geven van de vormen die deze tweedimensionale voorstelling aanneemt bij verschillende waarden van μ_x , μ_y , σ_x , σ_y en ρ zijn in figuur 20.3 een zestal verschillende tweedimensionaal normale verdelingen getekend. Wij hebben reeds gezien, dat de beide regressielijnen samenvallen als $\rho = +1$ of -1 , dus bij volkomen (positieve of negatieve) correlatie. Uit deze figuren blijkt, dat de hoek tussen de regressielijnen toeneemt, naarmate de correlatie afneemt en dat deze lijnen loodrecht op elkaar staan (met de regressielijn van y op x evenwijdig aan de X -as en de regressielijn van x op y evenwijdig aan de Y -as) als $\rho = 0$, dus bij het ontbreken van correlatie.



N.I.P.G. 981

Figuur 20.8. Ellipsen bij drie overschrijdingskansen voor verschillende tweedimensionaal normale verdelingen.

20.1.2. SCHATTINGEN UIT EEN STEEKPROEF

Als men beschikt over een aselekt getrokken steekproef van n elementen uit een tweedimensionaal normaal verdeelde populatie, dan kan men met formule (18.16) een zuivere schatting r van ρ verkrijgen. Een schatting van de regressielijn van y op x levert de vergelijking

$$(20.6) \quad \hat{y}_j = a_{yx} + b_{yx}x_j = \bar{y} + b_{yx}(x_j - \bar{x}) = \bar{y} + r \frac{s_y}{s_x}(x_j - \bar{x}) .$$

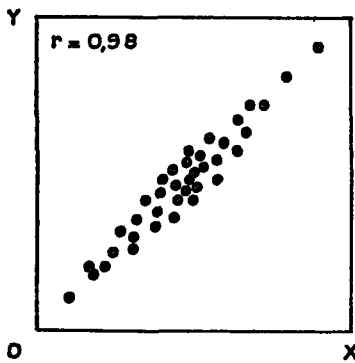
Volgens (19.3) is een zuivere schatting van $\sigma^2_{y..}$:

$$(20.7.1) \quad s^2_{y..} = \frac{\sum (y-\bar{y})^2 - \frac{[\sum (x-\bar{x})(y-\bar{y})]^2}{\sum (x-\bar{x})^2}}{n-2} = \frac{\sum (y-\bar{y})^2 - r^2 \sum (y-\bar{y})^2}{n-2}$$

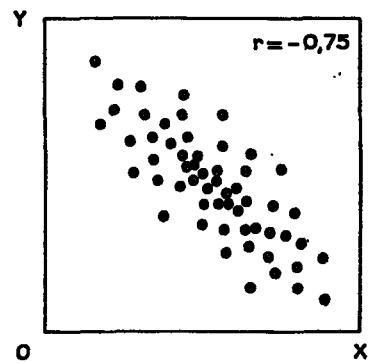
$$= \frac{\sum (y-\bar{y})^2 (1-r^2)}{n-2} = \frac{n-1}{n-2} s_y^2 (1-r^2).$$

Voor niet te kleine n geldt dus bij benadering [vergelijk met (20.3)]:

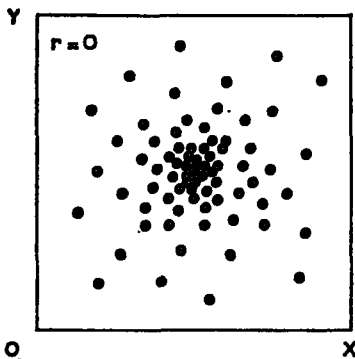
$$(20.7.2) \quad s^2_{y..} = s_y^2 (1-r^2).$$



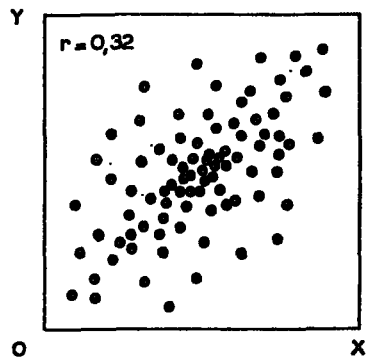
A. Sterke positieve correlatie



B. Negatieve correlatie



C. Geen correlatie



D. Zwakke positieve correlatie

N.I.P.G. 895

Figuur 20.4. Puntenwolken bij vier waarden van de correlatiecoëfficiënt.

Een schatting van de regressielijn van x op y kan worden verkregen door te berekenen

$$(20.8) \quad x_i = a_{xy} + b_{xy}y_i = \bar{x} + b_{xy}(y_i - \bar{y}) = \bar{x} + r \frac{s_x}{s_y}(y_i - \bar{y}).$$

Een zuivere schatting van $\sigma^2_{x \cdot y}$ is

$$(20.9.1) \quad s^2_{x \cdot y} = \frac{\Sigma(x - \bar{x})^2 - \frac{[\Sigma(x - \bar{x})(y - \bar{y})]^2}{\Sigma(y - \bar{y})^2}}{n - 2} = \frac{n-1}{n-2} s_x^2 (1 - r^2).$$

Voor niet te kleine n geldt dus bij benadering [vergelijk met (20.4)]:

$$(20.9.2) \quad s^2_{x \cdot y} = s_x^2 (1 - r^2).$$

In figuur 20.4 zijn voor vier verschillende tweedimensionaal normale populaties de spreidingsdiagrammen getekend, die bij steekproeven van vrij grote omvang werden verkregen. Bij aanwezigheid van correlatie vertonen de hierin staande *puntenwolken* een ellipsvormige concentratie, die sterker is naarmate de correlatie hoger is (vergelijk met figuur 20.3).

Ook het spreidingsdiagram in figuur 20.5 (betreffende de zuurstofgehalten in het arteriële en het veneuze bloed) vertoont zo'n typische ellipsvormige puntenwolk en wij nemen dan ook aan, dat het tweedimensionaal normale model bij benadering geldt. Voor deze waarnemingsparen vindt men (uit tabel 18.9):

$$\begin{aligned} \Sigma x &= 982,0, \quad \bar{x} = 19,64, \quad \Sigma x^2 = 19362,40, \quad \Sigma(x - \bar{x})^2 = 75,92, \quad s_x^2 = 1,5494, \\ \Sigma y &= 647,7, \quad \bar{y} = 12,95, \quad \Sigma y^2 = 8471,79, \quad \Sigma(y - \bar{y})^2 = 81,4842, \quad s_y^2 = 1,6629, \\ \Sigma xy &= 12782,45, \quad \Sigma(x - \bar{x})(y - \bar{y}) = 61,622, \end{aligned}$$

zodat uit (18.16) volgt

$$r = \frac{61,622}{\sqrt{(75,92)(81,4842)}} = \frac{61,622}{78,653} = 0,783.$$

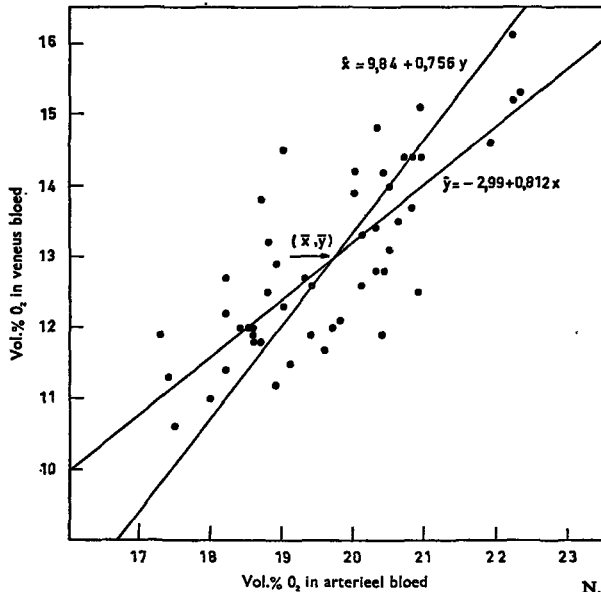
Volledigheidshalve voeren wij voor dit voorbeeld ook de berekeningen betreffende de regressielijnen uit:

$$\begin{aligned} b_{yx} &= \frac{61,622}{75,92} = 0,8117, & b_{xy} &= \frac{61,622}{81,4842} = 0,7562, \\ a_{yx} &= \frac{647,7 - (0,8117)(982)}{50} = -2,99, & a_{xy} &= \frac{982 - (0,7562)(647,7)}{50} = 9,84, \\ s^2_{y \cdot x} &= \frac{81,4842 - \frac{(61,622)^2}{75,92}}{48} = 0,6556, & s^2_{x \cdot y} &= \frac{75,92 - \frac{(61,622)^2}{81,4842}}{48} = 0,6108. \end{aligned}$$

(Ter controle kan dienen, dat $r = \sqrt{b_{yx} \cdot b_{xy}} = \sqrt{(0,8117)(0,7562)} = \sqrt{0,613808} = 0,783$. Verder dienen de regressielijnen elkaar in het

punt (\bar{x}, \bar{y}) te snijden, zodat voor $x = \bar{x}$ geldt, dat $\hat{y} = \bar{y}$ en voor $y = \bar{y}$, dat $\hat{x} = \bar{x}$. In het eerste geval vindt men bv. dat $\hat{y} = -2,99 + (0,8117)(19,64) = 12,95$.)

De beide berekende regressielijnen zijn in figuur 20.5 in het spreidingsdiagram van de waarnemingsparen getekend. De regressielijn van y op x kan dienen, om het gemiddelde van de grootheid y bij een bepaalde waarde van x te schatten. Met de regressielijn van x op y kan het gemiddelde van x voor een bepaalde y worden geschat. Voor elke regressielijn gelden hierbij de toetsen en schattingen die in 19.1.3 zijn besproken.



Figuur 20.5. Spreidingsdiagram van de waarnemingsparen in tabel 18.9 met de beide geschatte regressielijnen.

Beschouw vervolgens de 15 waarnemingsparen in tabel 18.10, waarvan het spreidingsdiagram in figuur 18.12 is getekend. De puntenwolk vertoont ook hier globaal een ellipsvorm, maar de juistheid van het tweedimensionaal normale model is bij dit kleine aantal paren veel minder zeker. De regressie lijkt echter rechtlijnig te zijn en men vindt:

$$\Sigma x = 279, \Sigma x^2 = 6789, \Sigma (x - \bar{x})^2 = 1599,60,$$

$$\Sigma y = 1183, \Sigma y^2 = 94431, \Sigma (y - \bar{y})^2 = 1131,73,$$

$$\Sigma xy = 20902, \Sigma (x - \bar{x})(y - \bar{y}) = -1101,80,$$

zodat

$$r = \frac{-1101,80}{\sqrt{(1599,60)(1131,73)}} = -0,819 .$$

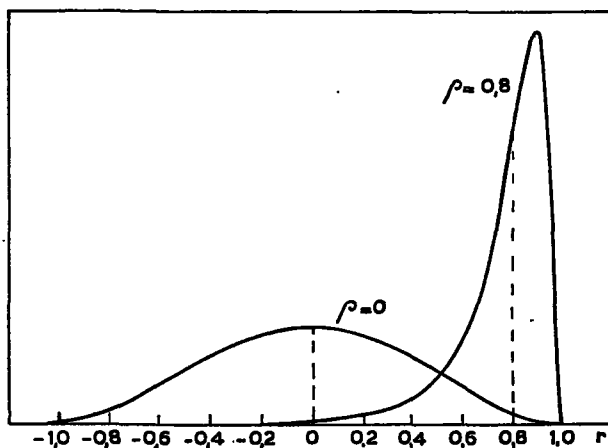
Tenslotte bekijken wij de waarnemingsparen in figuur 18.13, betref-

fende het calciumgehalte in het serum en in de 24-uurs urine bij 24 patiënten met verschillende kwaadaardige aandoeningen. Het is niet verrassend, dat voor dit zeer heterogene materiaal¹ de verdelingen van de beide gehalten *niet* normaal zijn en sterke uitbijters vertonen. Het tweedimensionaal normale model past hier zeker *niet* en men kan in zo'n geval de correlatiecoëfficiënt dan ook beter niet berekenen, daar deze niet alleen een zinloze, maar zelfs een misleidende uitkomst zal opleveren².

Tot dusverre hebben wij de verschillende berekeningen rechtstreeks, d.w.z. met de oorspronkelijke waarnemingen, uitgevoerd. In 20.1.7 geven wij een berekeningsmethode die kan worden gebruikt als men uit een groot aantal waarnemingsparen een correlatietabel heeft samengesteld.

20.1.3. TOETSING VAN DE HYPOTHESE $\rho = 0$

De kansverdeling van de correlatiecoëfficiënt r van een tweedimensionaal normaal verdeelde populatie met correlatiecoëfficiënt ρ is van een wiskundig bekend, maar vrij ingewikkeld type. Deze verdeling blijkt echter alleen afhankelijk te zijn van ρ en van de steekproefomvang n . Voor $\rho = 0$ en niet te kleine n is de verdeling van r bij benadering normaal, maar voor $|\rho| \neq 0$ vertoont deze verdeling een scheefheid, die bij een bepaalde n sterker wordt, naarmate $|\rho|$ verder van 0 verwijderd is (zie figuur 20.6).



N.I.P.G. 896

Figuur 20.6. Kansverdelingen van r voor $n = 8$ en $\rho = 0$ en $0,8$.

¹ Mannen en vrouwen van 23 t/m 81 jaar, met verschillende carcinomen (long, nier, larynx, portio, prostaat, ovarium, enz.) en met andere ziekten (w.o. leukemie, lupus, otitis, osteoporose).

² De correlatiecoëfficiënt voor alle 24 patiënten is 0,80. Laat men de 4 patiënten met hoge gehalten in het serum en in de urine weg (de nummers 1, 2, 3 en 5), dan vindt men $r = -0,126$. Hieruit blijkt, hoe gevoelig de correlatiecoëfficiënt is voor uitbijters.

Bij een bepaalde waarde van ρ nadert de kansverdeling van r echter steeds dichter tot een normale verdeling, naarmate n groter wordt.

Voor een tweedimensionaal normaal verdeelde populatie met $\rho = 0$ blijkt de grootheid

$$(20.10) \quad t = \frac{r \sqrt{n-2}}{\sqrt{1-r^2}} \quad (n \geq 3)$$

een STUDENT-verdeling met $n-2$ vrijheidsgraden te volgen. De hypothese $H_0: \rho = 0$ kan voor zo'n populatie dus op basis van een aselechte steekproef van n waarnemingsparen worden getoetst, door in (20.10) n en de waargenomen r te substitueren en de overschrijdingskans van de waargenomen t_0 in tabel C (met $\nu = n - 2$) op te zoeken. Men kan echter ook gebruik maken van tabel E, waarin voor $\rho = 0$ en verschillende waarden van ν , enkele fractielen van de verdelingen van r zijn opgenomen. Toetst men tweezijdig ($H_1: \rho \neq 0$) dan kan men bij een drempelwaarde α tot het verwerpen van H_0 besluiten, als $P_D(t_0) < \alpha$ is, of als geldt dat $r_0 < r_{\frac{1}{2}\alpha}$ of $r_0 > r_{1-\frac{1}{2}\alpha}$. Bij links éézijdige toetsing ($H_1: \rho < 0$, negatieve correlatie) verwerpt men H_0 als $P_L(t_0) < \alpha$ of als $r_0 < r_\alpha$ is. Bij rechts éézijdige toetsing ($H_1: \rho > 0$, positieve correlatie) wordt H_0 verworpen als $P_R(t_0) < \alpha$ of als $r_0 > r_{1-\alpha}$ is.

Opmerking

Als standaardfout van de correlatiecoëfficiënt (= spreiding van de kansverdeling van r) wordt soms opgegeven: $\sigma_r = (1 - \rho^2)/\sqrt{n-1}$. Deze formule dient te luiden

$$(20.11) \quad \sigma_r = \frac{1 - \rho^2}{\sqrt{n-1}}.$$

Hierin komt echter de (onbekende) parameter ρ voor en daar verder de kansverdeling van r bij betrekkelijk kleine waarden van n en vooral bij $\rho \neq 0$ zeker niet normaal is, heeft men weinig aan deze formule. Bepaald *onjuist* is het om bv. bij een waargenomen $r = 0,8$ (met $n = 37$) als standaardfout op te geven

$$s_r = \frac{1 - (0,8)^2}{\sqrt{37-1}} = 0,06$$

(men schrijft dit dan vaak als: $r = 0,8 \pm 0,06$), daar men hieruit *ten onrechte* zou concluderen, dat bv. het 95% betrouwbaarheidsinterval van ρ is: $0,8 - 1,96(0,06) < \rho < 0,8 + 1,96(0,06)$, $0,68 < \rho < 0,92$. In 20.1.4 bespreken wij, hoe men in zo'n geval wel tewerk kan gaan.

Wanneer echter $\rho = 0$, gaat (20.11) over in

$$(20.12) \quad \sigma_r = \frac{1}{\sqrt{n-1}},$$

en bij grote steekproefomvang geldt dan, daar de verdeling van r bij benadering normaal is met gemiddelde 0, dat de grootheid

$$(20.13) \quad \underline{T} = \frac{r - \varrho}{\sigma_r} = \frac{r}{1} = r \sqrt{n-1} \\ \sqrt{n-1}$$

bij benadering standaardnormaal verdeeld is.

Voorbeelden

20.1. Een aselechte steekproef van 12 waarnemingsparen uit een tweedimensionaal normaal verdeelde populatie levert op $r_0 = 0,40$. Men wil de hypothese $H_0: \varrho = 0$ toetsen en kiest $\alpha = 0,05$. De alternatieve hypothese H_1 luidt, dat er correlatie aanwezig is. De vraag, die men hier feitelijk wenst te beantwoorden luidt: kan (bij dit betrekkelijk kleine aantal waarnemingsparen) de waargenomen correlatie nog op toeval berusten?

Er dient tweezijdig getoetst te worden. Met (20.10) vindt men

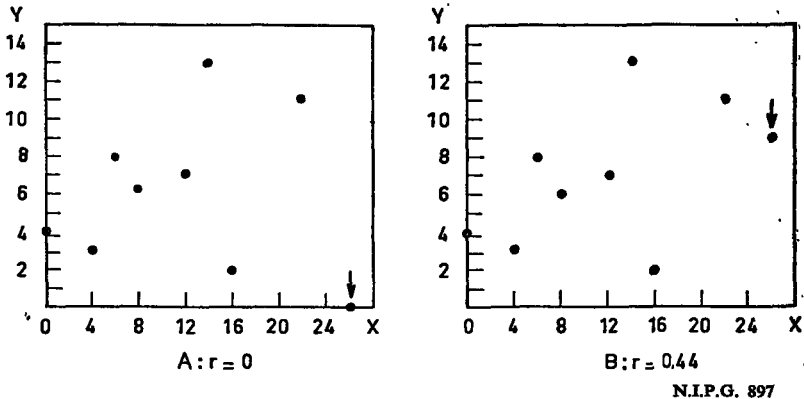
$$t_0 (v = 10) = \frac{0,4 \sqrt{10}}{\sqrt{1 - (0,4)^2}} = 1,38$$

en uit tabel C blijkt, dat $P_D(1,38)$ iets kleiner is dan 20%, zodat H_0 niet kan worden verworpen.

Bij het gebruik van tabel E leest men voor $n - 2 = 10$ en $r_{0,975}$ de kritieke waarde 0,576 af. Dit betekent, dat men bij de gekozen 5% drempel tot verwerpen van H_0 zou kunnen overgaan, als $|r_0| \geq 0,576$. Vanzelfsprekend komt men dus ook met deze tabel tot de conclusie, dat H_0 niet kan worden verworpen.

Uit dit voorbeeld blijkt, dat men voorzichtig moet zijn met de interpretatie van een correlatiecoëfficiënt, berekend uit een betrekkelijk kleine steekproef (of met het trekken van conclusies uit een spreidingsdiagram, waaruit 'duidelijk de aanwezigheid van correlatie blijkt'). Een steekproef van beperkte omvang uit een tweedimensionaal normale populatie met $\varrho = 0$ kan nog heel goed een vrij hoge correlatiecoëfficiënt te zien geven. Figuur 20.7 demonstreert de grote invloed, die bij kleine steekproefomvang de ligging van één waarnemingspaar op de waarde van de correlatiecoëfficiënt heeft. Beide spreidingsdiagrammen hebben 8 waarnemingsparen gemeen, maar bij A is het negende paar (26, 0) en bij B is dit (26, 9).

20.2. Voor de 15 waarnemingsparen in tabel 18.10 (figuur 18.12) vonden wij $r_0 = -0,819$. Wij toetsen met $\alpha = 0,01$ de hypothese $H_0: \varrho = 0$ met als alternatieve hypothese, dat negatieve correlatie



Figuur 20.7.

aanwezig is. Er dient dus links éézijdig getoetst te worden. De toetsing verloopt het eenvoudigst met tabel E: voor $n - 2 = 13$ lezen wij hierin af, dat $r_{0,01} = -0,592$ en daar r_0 aanzienlijk kleiner is, wordt H_0 verworpen. Uit de tabel blijkt ook, dat $r_{0,0005} = -0,760$, zodat de linkse kans van r_0 zelfs kleiner is dan $0,0005$.

20.3. Voor een steekproef van 240 waarnemingsparen uit een tweedimensionaal normale populatie vindt men $r_0 = 0,17$. Men wil met $\alpha = 0,01$ de hypothese $\rho = 0$ toetsen. Bij deze steekproefomvang kan gebruik worden gemaakt van formule (20.13) en deze levert op

$$T_0 = 0,17 \sqrt{239} = 2,628 ,$$

met een tweezijdige overschrijdingskans van ongeveer 0,0084 (tabel A), zodat H_0 wordt verworpen.

20.1.4. TOETSING EN SCHATTINGEN INDIEN $\rho \neq 0$: DE z -TRANSFORMATIE

Door DAVID (140) is de kansverdeling van r berekend voor $\rho = 0,0$ (0,1) 0,9 en $n = 3$ (1) 25, 50, 100, 200 en 400. Zij stelde ook grafieken samen, waarmee men voor waarden van n tussen 3 en 400 op basis van een waargenomen r bij benadering een 90%, 95%, 98% en 99% betrouwbaarheidsinterval voor ρ kan aflezen. Deze zijn voor 95% en 99% o.m. opgenomen in de BIOMETRIKA TABLES I (74, tabel 15).

Wanneer $\rho \neq 0$ kan men voor het toetsen van hypothesen en voor het bepalen van betrouwbaarheidsgrenzen betreffende ρ ook gebruik maken van de zg. z -transformatie, die afkomstig is van FISHER. Deze vond, dat de grootheid

$$(20.14) \quad z = \frac{1}{2} \ln \frac{1+r}{1-r} = 1,1513 [\log(1+r) - \log(1-r)]$$

ook voor kleine steekproeven en voor dicht bij ± 1 gelegen waarden van ϱ ongeveer normaal verdeeld is, met een spreiding die uitsluitend afhangt van de steekproefomvang:

$$(20.15) \quad \sigma_z = \frac{1}{\sqrt{n-3}}$$

en met gemiddelde

$$(20.16.1) \quad \mu_z = \frac{1}{2} \ln \frac{1+\varrho}{1-\varrho} + \frac{\varrho}{2(n-1)}.$$

De term rechts in (20.16.1) is klein in vergelijking tot σ_z en kan daarom zonder bezwaar verwaarloosd worden, zodat bij benadering

$$(20.16.2) \quad \mu_z = \frac{1}{2} \ln \frac{1+\varrho}{1-\varrho} = 1,1513 [\log(1+\varrho) - \log(1-\varrho)].$$

Uit het voorgaande volgt, dat de grootheid

$$(20.17) \quad \underline{T} = \frac{(z - \mu_z)}{\sigma_z} = (z - \mu_z) \sqrt{n-3}$$

bij benadering standaardnormaal verdeeld is. De hypothese $H_0: \varrho = \varrho_0$ kan dus via deze grootheid worden getoetst door in (20.17) de waarde z_0 die correspondeert met r_0 (de waargenomen correlatiecoëfficiënt) en de waarde μ_z die correspondeert met ϱ_0 (de te toetsen waarde van de correlatiecoëfficiënt) en de steekproefomvang n te substitueren. Voor de transformatie van r (ϱ) naar z (μ_z) en vice versa kan men gebruik maken van tabel U.

Voor de praktijk is echter van meer belang, dat men via de grootheid z betrouwbaarheidsgrenzen voor ϱ kan verkrijgen. De hierbij te volgen procedure is:

1. Zet de waargenomen r_0 met tabel U om in een z_0 .
2. Bepaal 100 $(1-\alpha)$ % betrouwbaarheidsgrenzen van μ_z door te nemen

$$(20.18) \quad z_0 + \frac{T_{\frac{1}{2}\alpha}}{\sqrt{n-3}} < \mu_z < z_0 + \frac{T_{1-\frac{1}{2}\alpha}}{\sqrt{n-3}},$$

waarin de fractielen van \underline{T} uit tabel B kunnen worden afgelezen.

3. Zet met tabel U de voor μ_z gevonden grenzen om in grenzen voor ϱ .

Voorbeelden

20.4. Onderstel dat een steekproef van 228 waarnemingsparen oplevert $r_0 = 0,760$. Men wil de hypothese $H_0: \varrho = 0,80$ toetsen met als alternatieve hypothese $H_1: \varrho < 0,80$ en men kiest $\alpha = 0,05$. Uit tabel U blijkt, dat met $r_0 = 0,760$ correspondeert $z_0 = 0,996216$ en dat met

$\rho = 0,800$ correspondeert $\mu_x = 1,098613$. Substitutie in (20.17) levert op

$$T_0 = (0,996216 - 1,098613) \sqrt{228 - 3} = -1,54.$$

Deze uitkomst heeft een linkse kans van 0,0618, zodat H_0 niet wordt verworpen.

Opmerking

Daar de verdeling van r niet normaal is, is het *onjuist* om te nemen [zie formule (20.11)]:

$$T_0 = \frac{r_0 - \rho_0}{\sigma_r} = \frac{(r_0 - \rho_0) \sqrt{n-1}}{1 - \rho_0^2} = \frac{-0,04 \sqrt{227}}{1 - (0,8)^2} = -1,67,$$

(met een linkse kans van 0,0475).

20.5. In 20.1.2 vonden wij voor de 50 waarnemingsparen in figuur 20.5 een $r_0 = 0,783$. Men wil op grond van deze uitkomst een 95% betrouwbaarheidsinterval van ρ samenstellen. Uit tabel U leest men af, dat $z_0 = 1,0531$ en daar $T_{0,975} = 1,96$, is volgens (20.18):

$$1,0531 - \frac{1,96}{\sqrt{50-3}} < Z < 1,0531 + \frac{1,96}{\sqrt{50-3}},$$

$$0,7672 < Z < 1,3390.$$

Deze grenzen voor μ_x kunnen tenslotte met tabel U weer worden omgezet in grenzen voor ρ . Men vindt dan in twee decimalen nauwkeurig

$$0,65 < \rho < 0,87.$$

20.1.5. TOETSING VAN DE HYPOTHESE $\rho_1 = \rho_2$

Beschouw twee onafhankelijke tweedimensionaal normaal verdeelde populaties met correlatiecoëfficiënt ρ_1 , resp. ρ_2 . Uit de eerste populatie trekt men aselect een steekproef van n_1 waarnemingsparen en hiervoor is de correlatiecoëfficiënt r_1 . Een aselecte steekproef van n_2 waarnemingsparen uit de tweede populatie heeft als correlatiecoëfficiënt r_2 . Noemt men nu de met r_1 en r_2 corresponderende waarden van z resp. z_1 en z_2 en de met ρ_1 en ρ_2 corresponderende waarden van μ_x resp. μ_{x_1} en μ_{x_2} , dan volgt de grootheid

$$(20.19) \quad T = \frac{(z_1 - z_2) - (\mu_{x_1} - \mu_{x_2})}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}}$$

bij benadering een standaardnormale verdeling. Onder de hypothese $H_0: \rho_1 = \rho_2$ is ook $\mu_{x_1} = \mu_{x_2}$. Deze hypothese kan dus worden getoetst

via de (bij benadering) standaardnormaal verdeelde grootheid

$$(20.20) \quad T = \frac{z_1 - z_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

Voorbeeld

20.6. De onderstaande gegevens betreffen de correlatie tussen de systolische tensie en het lichaamsgewicht, waargenomen bij twee steekproeven van gezonde kinderen van 7, 8 en 9 jaar uit verschillende populaties. De te toetsen hypothese luidt, dat het waargenomen verschil tussen de correlatiecoëfficiënten toevallig is, d.w.z. dat $\rho_1 = \rho_2$. De berekeningen spreken voor zichzelf:

	n	r	z	$n-3$	$1/(n-3)$
Steekproef 1	209	0,410	0,435611	206	0,004854369
Steekproef 2	324	0,322	0,333877	321	0,003115265
			$z_1 - z_2 = 0,101734$		0,007969634
			$T_0 = \frac{0,101734}{\sqrt{0,007969634}} = \frac{0,101734}{0,089273} = 1,14$		

De tweezijdige overschrijdingskans van deze uitkomst is ongeveer 0,25, zodat H_0 niet wordt verworpen.

20.1.6. TOETSING VAN DE HYPOTHESE $\rho_1 = \rho_2 = \dots = \rho_k$

Men kan met behulp van de z -transformatie voor k tweedimensionaal (bij benadering) normale populaties ook de hypothese $H_0: \rho_1 = \rho_2 = \dots = \rho_k = \rho$ toetsen. Men kan dan gebruik maken van het feit, dat de grootheid

$$(20.21) \quad \underline{X}^2 = \Sigma (n-3) z^2 - \frac{[\Sigma (n-3) z]^2}{\Sigma (n-3)}$$

onder H_0 bij benadering een χ^2 -verdeling volgt met $k-1$ vrijheidsgraden. In tabel 20.1 is een voorbeeld gegeven van de berekeningen, die bij de uitvoering van deze toets dienen plaats te vinden. De toetsing van de hypothese H_0 is hier gebaseerd op de waargenomen correlatiecoëfficiënten van vier aselecte steekproeven.

Tabel 20.1. Berekeningen bij toetsing van de hypothese $H_0: \rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho$

Steekproef	n	$n-3$	r	z	$(n-3)z$	z^2	$(n-3)z^2$
1	33	30	0,530	0,590145	17,7044	0,348271	10,4481
2	58	55	0,621	0,726631	39,9647	0,527993	29,0396
3	42	39	0,705	0,877173	34,2097	0,769432	30,0078
4	47	44	0,456	0,492250	21,6590	0,242310	10,6616
Totaal	180	168			113,5378		80,1571

Uit tabel 20.1 volgt met (20.21)

$$X_0^2 = 80,1571 - \frac{(113,5378)^2}{168} = 80,1571 - 76,7311 = 3,4260 ,$$

een uitkomst die, met 3 vrijheidsgraden, een overschrijdingskans bezit die groter is dan 0,3 en dus geen reden vormt tot het verwerpen van H_0 . Het gewogen gemiddelde van de vier waarden van z bedraagt $\bar{z} = 113,5378/168 = 0,67582$ en uit tabel U blijkt, dat hiermee ongeveer correspondeert $r = 0,589$. Deze uitkomst is een schatting van de (gemeenschappelijke) correlatiecoëfficiënt ρ , die 168 vrijheidsgraden bezit.

Opmerking

De verkregen schatting is systematisch te hoog, daar men feitelijk van elke waarde van z een correctie ter grootte van $\rho/2(n-1)$ zou moeten aftrekken. De moeilijkheid is, dat de hiervoor benodigde waarde van ρ niet bekend is. Deze fout is echter van weinig betekenis, als het aantal steekproeven klein en hun omvang (vrij) groot is. Men kan zo nodig de correctie bij benadering toepassen, door voor ρ een waarde in de plaats te stellen, die iets lager is dan de uit \bar{z} verkregen r (bij het voorgaande voorbeeld dus iets lager dan 0,589, bv. 0,58). Voor de gegevens in tabel 20.1 vindt men dan

Steekproef	z	correctie	z'	$(n-3) z'$
1	0,590145	0,009063	0,581082	17,4325
2	0,726631	0,005088	0,721543	39,6849
3	0,877173	0,007073	0,870100	33,9339
4	0,492250	0,006304	0,485946	21,3816
Totaal	2,686199	0,027528	2,658671	112,4329

Hieruit volgt, dat $\bar{z}' = 112,4329/168 = 0,669243$ en $r = 0,584$ — een waarde die slechts weinig van de zonder correctie berekende verschilt.

20.1.7. BEREKENING VAN r UIT EEN CORRELATIETABEL

Als de waarnemingsparen in een correlatietabel zijn samengevat, waarin de verdelingen van de grootheden x en y een constante klassebreedte bezitten, kan de berekening van de correlatiecoëfficiënt (en van de beide regressiecoëfficiënten) door het gebruik van hulpgrootheden worden vereenvoudigd. Hierbij worden de (klasse) waarden van x (y) vervangen door waarden van de hulpgrootheid u (v) en wel zodanig dat:

$$u = \frac{x - x_0}{c_x} \quad (\text{zodat } x = x_0 + c_x u), \quad s_u = \frac{s_x}{c_x} \quad (\text{zodat } s_x = c_x s_u)$$

en

$$v = \frac{y - y_0}{c_y} \quad (\text{zodat } y = y_0 + c_y v), \quad s_v = \frac{s_y}{c_y} \quad (\text{zodat } s_y = c_y s_v) .$$

Tabel 20.2. Berekening van de correlatiecoëfficiënt uit een correlatietabel door middel van hulpgrootheden

$x \backslash y$		Gewicht in kg													1	2	3	4	5	6	
		19	20	21	22	23	24	25	26	27	28	29	30	31	32	f_v	v	fv	fv^2	Σfv	$v\Sigma fv$
Systolische tensie in mm Hg	130											1			1	+4	+4	16	4	16	
	125												1	1	2	+3	+6	18	12	36	
	120							1		2			1		4	+2	+8	16	9	18	
	115					2	1	1	2				1	1	8	+1	+8	8	7	7	
	110			1		2	3	2	1	1			1		11	0	0	0	-4	0	
	105		1	1	1	2	2	2	1		1				11	-1	-11	11	-14	14	
	100		1		1		1	2			1				6	-2	-12	24	-6	12	
	95				1	1	1								3	-3	-9	27	-6	18	
	90			1	1	1									3	-4	-12	48	-9	36	
	85	1													1	-5	-5	25	-6	30	
1	f_u	1	2	3	4	8	8	8	4	3	2	3	2	1	1	50		-23	193	-13	187
2	u	-6	-5	-4	-3	-2	-1	0	+1	+2	+3	+4	+5	+6	+7						
3	fu	-6	-10	-12	-12	-16	-8	0	4	6	6	12	10	6	7	-13	Σfu				
4	fu^2	36	50	48	36	32	8	0	4	12	18	48	50	36	49	427	Σfu^2				
5	Σfv	-5	-3	-5	-10	-7	-6	-3	+1	+4	-3	+5	+5	+1	+3	-23					
6	$u \Sigma fv$	30	15	20	30	14	6	0	1	8	-9	20	25	6	21	187	$\Sigma fu v$				

Controles

$c_x = 1$
 $x_0 = 25$
 $c_y = 5$
 $y_0 = 110$
 $n = 50$

Deze procedure hebben wij in 2.5.2 reeds beschreven en zij is toegepast op het voorbeeld in tabel 20.2, betreffende het gewicht en de systolische tensie van 50 gezonde jongens van 7 t/m 9 jaar.

Men kan nu eenvoudig aantonen, dat de correlatiecoëfficiënt ook kan worden berekend met

$$(20.22) \quad r = \frac{\Sigma (u - \bar{u}) (v - \bar{v})}{\sqrt{\Sigma (u - \bar{u})^2 \cdot \Sigma (v - \bar{v})^2}} = \frac{\Sigma fuv - \frac{\Sigma fu \cdot \Sigma fv}{n}}{\sqrt{\left[\Sigma fu^2 - \frac{(\Sigma fu)^2}{n} \right] \left[\Sigma fv^2 - \frac{(\Sigma fv)^2}{n} \right]}}$$

en dat voor de regressiecoëfficiënten geldt

$$(20.23) \quad b_{yv} = \frac{c_y}{c_x} b_{vu} = \frac{c_y}{c_x} \cdot \frac{\Sigma fuv - \frac{\Sigma fu \cdot \Sigma fv}{n}}{\Sigma fu^2 - \frac{(\Sigma fu)^2}{n}}$$

en

$$(20.24) \quad b_{xv} = \frac{c_x}{c_y} b_{uv} = \frac{c_x}{c_y} \cdot \frac{\Sigma fuv - \frac{\Sigma fu \cdot \Sigma fv}{n}}{\Sigma fv^2 - \frac{(\Sigma fv)^2}{n}}$$

In tabel 20.2 is een schema gegeven, waarmee de voor de berekening benodigde sommen (Σfu , Σfv , Σfu^2 , Σfv^2 en Σfuv) snel kunnen worden bepaald. Beschouw van dit schema bv. de kolommen 1 t/m 6 (voor de rijen 1 t/m 6 geldt mutatis mutandis hetzelfde). In kolom 1 staat de frequentieverdeling van y , in kolom 2 staan de waarden van de hulp-grootheid v (merk op dat $y_0 = 110$ en dat $c_y = 5$), kolom 3 geeft de producten fv en kolom 4 de producten fv^2 . Sommeert men deze kolommen, dan krijgt men dus Σfv , resp. Σfv^2 . In kolom 5 staat op elke regel de som van de waarden van u van alle waarnemingen, die op deze regel voorkomen. Op de bovenste regel ($y = 130$, $v = +4$) staat slechts één waarneming en wel bij $x = 29$, d.i. $u = +4$. Op deze regel wordt dus in kolom 5 genoteerd: $+4$. Op de tweede regel ($y = 125$, $v = +3$) staan twee waarnemingen en wel één bij $u = +5$ en één bij $u = +7$, zodat in kolom 5 wordt genoteerd $5 + 7 = 12$. En op de vijfde regel ($y = 110$, $v = 0$) staan 11 waarnemingen; van links naar rechts gaande vindt men als som van hun waarden van u :

$$1(-4) + 2(-2) + 3(-1) + 2(0) + 1(+1) + 1(+2) + 1(+4) = -4.$$

Tenslotte worden de getallen in kolom 6 verkregen, door de getallen in de kolommen 2 en 5 met elkaar te vermenigvuldigen. Per regel verschijnt dan het product $v \cdot \Sigma fu = fuv$ en de som van de getallen in kolom 6 is dus Σfuv . Het schema is zo ingericht, dat Σfu , Σfv en Σfuv

tweemaal worden berekend, zodat een contrôle op hun juistheid wordt verkregen.

Voor de gegevens in tabel 20.2 vindt men nu met (20.22)

$$r = \frac{187 - \frac{(-13)(-23)}{50}}{\sqrt{\left[427 - \frac{(-13)^2}{50}\right] \left[193 - \frac{(-23)^2}{50}\right]}} = \frac{181,02}{\sqrt{(423,62)(182,42)}} = 0,651.$$

Verder volgt uit (20.23) en (20.24):

$$b_{yx} = \frac{5}{1} \cdot \frac{181,02}{423,62} = 2,137 \quad \text{en} \quad b_{xy} = \frac{1}{5} \cdot \frac{181,02}{182,42} = 0,198$$

Daar

$\bar{u} = \Sigma fu/n = -13/50 = -0,26$, is $\bar{x} = x_0 + c_x \bar{u} = 25 + (1)(-0,26) = 24,74$
en daar

$\bar{v} = \Sigma fv/n = -23/50 = -0,46$, is $\bar{y} = y_0 + c_y \bar{v} = 110 + (5)(-0,46) = 107,70$.

Zodoende is

$$a_{yx} = \bar{y} - b_{yx} \bar{x} = 107,70 - (2,137)(24,74) = 54,8,$$

$$a_{xy} = \bar{x} - b_{xy} \bar{y} = 24,74 - (0,198)(107,70) = 3,4.$$

De beide regressievergelijkingen zijn dus

$$\hat{y} = 54,8 + 2,14x$$

en

$$\hat{x} = 3,4 + 0,20y.$$

Wij hebben deze regressievergelijkingen berekend om te laten zien, hoe dit in z'n werk gaat. In werkelijkheid zal men zich hier interesseren voor de vraag, of er tussen gewicht en systolische bloeddruk een correlatie aanwezig is en deze vraag kan voor de populatie, waaruit de bestudeerde steekproef stamt bevestigend worden beantwoord, daar uit tabel E blijkt dat de tweezijdige overschrijdingskans van $r = 0,651$ (met $n - 2 = 48$) aanzienlijk kleiner is dan 0,001.

⇒ 20.1.8. ENKELE OPMERKINGEN BETREFFENDE DE CORRELATIECOËFFICIENT

De correlatiecoëfficiënt is primair een toetsingsgrootheid, die kan worden gebruikt als men op basis van een aselechte steekproef uit een tweedimensionaal (bij benadering) normale populatieverdeling van de grootheden x en y de hypothese wil toetsen, dat deze grootheden onderling onafhankelijk zijn. Verder kan men door middel van de correlatiecoëfficiënten voor twee (h) tweedimensionaal normaal ver-

deelde populaties de hypothese $\rho_1 = \rho_2$ ($\rho_1 = \rho_2 = \dots = \rho_k$) toetsen, waarbij dan gebruik dient te worden gemaakt van de z -transformatie.

Indien men een significante correlatiecoëfficiënt heeft waargenomen, geeft het teken van deze coëfficiënt informatie omtrent de richting van de relatie, terwijl men een schattingsinterval kan berekenen, dat een indruk geeft van de grootte van ρ . Hierbij dient men echter te bedenken, dat de correlatiecoëfficiënt zelf als graadmeter voor de correlatie feitelijk minder geschikt is. Het is nl. *niet zo*, dat een correlatiecoëfficiënt van bv. 0,80 op een tweemaal zo sterke correlatie wijst als een correlatiecoëfficiënt van 0,40. Beschouwt men echter het kwadraat van de correlatiecoëfficiënt, dan volgt uit (20.5)

$$(20.25) \quad \rho^2 = \frac{\sigma_y^2 - \sigma_{y.x}^2}{\sigma_y^2}.$$

Hierin is σ_y^2 de variantie van de grootte y en $\sigma_{y.x}^2$ de variantie om de regressielijn van y op x , zodat $\sigma_y^2 - \sigma_{y.x}^2$ de variantie van y is, die 'verklaard' wordt door deze regressie. Hieruit volgt, dat ρ^2 kan worden gezien als de fractie van de totale variantie van y , die door de regressie wordt verklaard. Is er geen correlatie, dan is $\rho^2 = 0$ en $\sigma_{y.x}^2 = \sigma_y^2$: de variantie van de bij een waarde x , behorende y -verdeling is gelijk aan de variantie van alle waarden van y . Is er een volkomen (positieve of negatieve) correlatie, dan is $\rho^2 = 1$ en dus $\sigma_{y.x}^2 = 0$; alle waarden van y liggen op de regressielijn van y op x en de gehele variabiliteit van de grootte y wordt door de regressie verklaard. Een analoge redenering geldt voor de regressie van x op y .

Men noemt ρ^2 wel de 'coëfficiënt of determination' en maakt op grond van deze coëfficiënt de in tabel 20.3 vermelde kwalitatieve indeling naar de grootte van de correlatiecoëfficiënt.

Tabel 20.3

ρ	ρ^2	Kwalificatie
0,90 — 1,00	0,81 — 1,00	zeer hoog
0,78 — 0,89	0,61 — 0,80	hoog
0,64 — 0,77	0,41 — 0,60	middelmatig
0,46 — 0,63	0,21 — 0,40	laag
0,00 — 0,45	0,00 — 0,20	zeer laag

Ten aanzien van het gebruik en de interpretatie van correlatiecoëfficiënten willen wij enkele waarschuwingen laten horen, die wij punt voor punt zullen beschouwen.

a. Correlatie en causaliteit. De aanwezigheid van correlatie tussen de grootheden x en y behoeft niet in te houden, dat er een causaal verband aanwezig is, d.w.z. dat een verandering van x direct verantwoordelijk is voor een verandering van y , of omgekeerd. In de

eerste plaats kan er een reciproke oorzaak-gevolg relatie aanwezig zijn, bv. als een prijsstijging van een product (door toenemende vraag) een vergroting van de productie ten gevolge heeft, die daarna op haar beurt weer een prijsdaling veroorzaakt. Ook kan het zijn, dat één of meer factoren beide grootheden beïnvloeden. Een persoon van flink postuur heeft meestal grote handen en grote voeten, zodat de afmetingen van deze twee lichaamsdelen gecorreleerd zijn door beider relatie tot het geheel. En een voorbeeld van BRADFORD HILL (42): 'For instance, in a series of towns it might be shown that the phthisis death-rate and overcrowding were correlated with one another, the former being high where the latter was high and vice versa. This is not necessarily evidence that phthisis is due to overcrowding. Possibly, and probably, towns with a high degree of overcrowding are also those with a low standard of living and nutrition. This third factor may be the one which is responsible for the level of the phthisis rate, and overcrowding is only indirectly associated with it.' Bijzonder voorzichtig dient men in dit verband te zijn met de correlatie van twee grootheden, die beide in de tijd variëren. Als men voor een reeks van jaren een hoge positieve correlatie vindt tussen het gemiddelde salaris van de Nederlandse hoogleraren en het gebruik van alcohol, dan kan deze correlatie ontstaan zijn zonder dat er één hoogleeraar alcohol heeft gebruikt: in de tijd is immers zowel de salariëring als de alcoholconsumptie toegenomen. Bij deze twee tijdreeksen is het onderstellen van een causaal verband (hogere salarissen leiden tot een groter alcoholverbruik) absurd, maar er zijn vele overeenkomstige gevallen waarin de sprong van een (hoge) correlatie tussen twee tijdreeksen naar een oorzaak-gevolg relatie zo plausibel is, dat hij wél wordt gemaakt. Men denke steeds aan de door PEARL (48) gegeven waarschuwing: 'When ever a correlation is observed, or a correlation coefficient is computed let it be remembered first, last and all the time, that it measures nothing, and means nothing, except statistical association between the variables, no matter how glamorous or seductive the suggestion as to causal relationship may seem to be.'

b. Pseudo-correlatie. Als de grootheden x , y en z onderling geen correlatie vertonen, dan vindt men wel een correlatie tussen deze en ervan afgeleide grootheden (bv. tussen x en $x + y$, x en x/y) of tussen de afgeleide grootheden onderling (bv. tussen x/z en y/z). Tabel 20.4 geeft een voorbeeld waaruit blijkt, dat er bij het ontbreken van correlatie tussen x en y een zeer hoge positieve correlatie tussen x en $x + y$ en een zeer hoge negatieve correlatie tussen x en y/x kan bestaan. Het gevaar is niet denkbeeldig, dat men in de praktijk met dergelijke afgeleide grootheden te maken krijgt zonder hun samenstelling te doorzien en daardoor tot foutieve conclusies komt; RÜMKE (192) geeft van zo'n pseudo-correlatie een fraai voorbeeld.

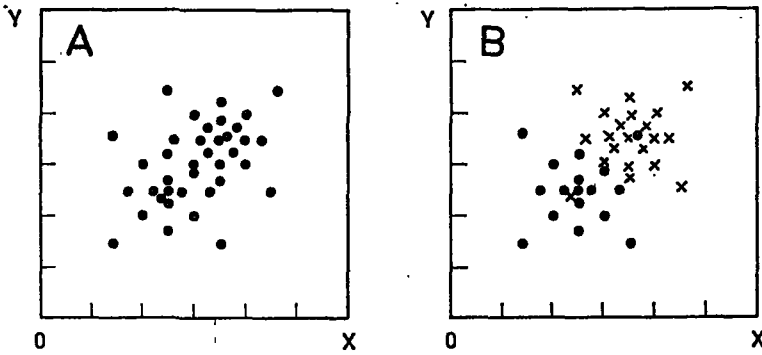
Tabel 20.4. Correlatie van x en y , x en $x + y$ en x en y/x

	x	y	x	$x+y$	x	y/x
	32	18	32	50	32	0,562
	31	13	31	44	31	0,419
	28	22	28	50	28	0,786
	24	17	24	41	24	0,708
	44	11	44	55	44	0,250
	53	19	53	72	53	0,358
	9	16	9	25	9	1,778
	35	23	35	58	35	0,657
	33	23	33	56	33	0,697
	31	18	31	49	31	0,581
Som	320	180	320	500	320	6,796
Kwadratensom	1206	146	1206	1352	1206	1,59887
Productsom		0		1206		-37,516
r		0		0,9445		-0,8543

Uit SNEDECOR (49).

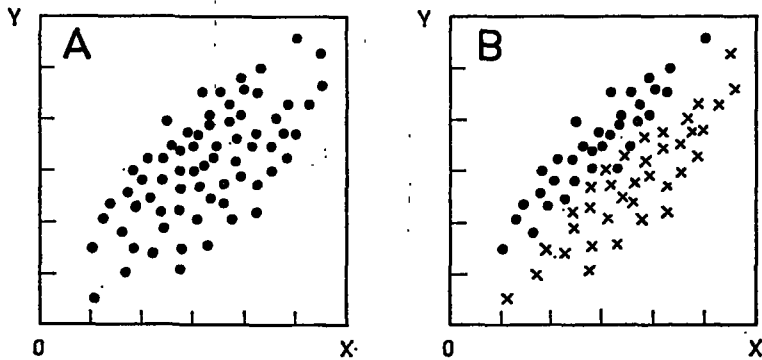
c. Heterogeniteit. Wanneer men de correlatie tussen twee grootheden wil bepalen, dient men attent te zijn op het mogelijke effect van heterogeniteit van het waarnemingsmateriaal (bv. t.a.v. geslacht, het al dan niet aanwezig zijn van een bepaalde afwijking of meerdere vormen van afwijkingen, de aanwezigheid van patiënten in verschillende stadia van een ziekte, het voorkomen van kinderen naast volwassenen, e.d.). Heterogeniteit kan tot een misleidend beeld omtrent de correlatie tussen twee grootheden leiden. In figuur 20.8 staat bij *A* een spreidingsdiagram, dat een duidelijke positieve correlatie vertoont, maar *B* laat zien, dat de puntenwolk uit twee soorten waarnemingsparen is opgebouwd (die met punten, resp. met kruisjes zijn aangegeven) en dat binnen elke soort geen correlatie aanwezig is. Dit geval zou zich bv. kunnen voordoen, als er zowel bij gezonde personen (de punten) als bij patiënten met een bepaalde aandoening (de kruisjes) geen correlatie tussen de systolische bloeddruk en het calciumgehalte van het serum aanwezig is, maar als bij de patiënten de bloeddruk en het calciumgehalte gemiddeld hoger liggen dan bij de gezonden.

Figuur 20.9 geeft een voorbeeld van het tegengestelde effect: de bij *A* geconstateerde correlatie is aanzienlijk lager dan de correlatie, die men bij *B* voor elke groep afzonderlijk vindt. Onderstel, dat ook hier de gezonde personen met punten en de patiënten met kruisjes zijn aangeduid en dat op de *X*-as het calciumgehalte en op de *Y*-as de bloeddruk is afgezet. Het geconstateerde effect is nu een gevolg van het feit, dat bij de patiënten het calciumgehalte gemiddeld hoger, maar de bloeddruk gemiddeld lager is dan bij de gezonde personen.



Figuur 20.8. Correlatie tengevolge van heterogeniteit.

N.I.P.G. 902a



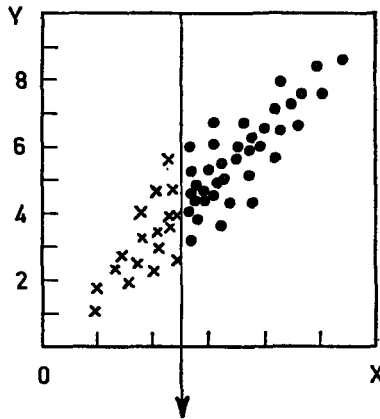
Figuur 20.9. Verlaging van correlatie tengevolge van heterogeniteit.

N.I.P.G. 902b

d. Uitvallers. Onderstel, dat de sollicitanten voor een bepaalde betrekking een medische keuring ondergaan en dat daarbij van alle goedgekeurden de vitale capaciteit wordt bepaald. Men gaat er nu van uit, dat personen waarvan de vitale capaciteit beneden een bepaalde waarde ligt ongeschikt zijn, zodat zij eveneens worden afgewezen. Na verloop van zekere tijd besluit men na te gaan, hoe hoog de correlatie tussen vitale capaciteit en de geschiktheid voor de functie is. Men laat daartoe van een aantal arbeiders, die meer dan een half jaar in dienst zijn en waarvan bij de medische keuring de vitale capaciteit is bepaald, de geschiktheid voor de functie in een schaal van 0 t/m 9 waarden en correleert deze waarderingscijfers met de capaciteitsindex.

Neem nu aan, dat tussen beide grootheden inderdaad een duidelijke correlatie aanwezig is (figuur 20.10). Men mist dan echter in ieder geval de waarderingscijfers van de personen, die wegens een te lage vitale capaciteit zijn afgewezen (deze zijn in de figuur met een kruisje aangeduid), zodat voor het berekenen van de correlatiecoëfficiënt slechts het rechterdeel van de puntenwolk (met de punten) ter beschikking staat en een te lage correlatie wordt waargenomen. Deze verlaging

van de correlatie tengevolge van uitvallers kan zeer aanzienlijk zijn, zodat men van de 'voorspellingswaarde' van de vitale capaciteit een veel te ongunstige indruk krijgt. Overigens bestaat ook de mogelijkheid, dat er van de aangenomen personen inmiddels een aantal zijn uitgevallen, zodat ook de 'punten' incompleet zijn; wij laten het aan de lezer over te overwegen, welke invloed deze uitvallers kunnen hebben.



N.I.P.G. 902c

Figuur 20.10. Verlaging van correlatie tengevolge van uitvallers:

- = x en y bekend,
- x = niet aangenomen personen, y in werkelijkheid onbekend.

Wij willen niet nalaten, hier nogmaals te memoreren, dat de correlatiecoëfficiënt en de hiermee samenhangende technieken gebaseerd zijn op het tweedimensionaal normale model. De kansverdeling van \bar{r} kan vooral bij betrekkelijk kleine steekproefomvang sterk beïnvloed worden door afwijkingen van dit model (in het bijzonder door uitgesproken scheefheid van de verdelingen van x en y en de aanwezigheid van uitbijters). Wanneer dergelijke afwijkingen optreden, zal men er dus goed aan doen de hypothese dat de grootheden x en y onafhankelijk zijn met een verdelingsvrije rangcorrelatiemethode te toetsen.

20.2. Opgaven

20.1. Bereken voor $r_0 = 0,06, 0,20, 0,60$ en voor $n = 5, 26, 82$ en 401 :

a. t_0 met formule (20.10).

b. T_0 met formule (20.13).

Vergelijk de uitkomsten. Wanneer vertonen deze de beste overeenstemming en waarom?

20.2. Bepaal met tabel E, of de volgende correlatiecoëfficiënten significant zijn bij een 5%, resp. 1%, drempel:

a. $r_0 = 0,6, n = 10$

b. $r_0 = 0,6, n = 15$

c. $r_0 = 0,6, n = 20$

d. $r_0 = 0,4, n = 20$

e. $r_0 = 0,4, n = 30$

f. $r_0 = 0,4, n = 50$.

20.3. Betreft de waarnemingsparen in voorbeeld 2.6, blz. 47. Onderstel, dat de populatieverdeling van de grootheden x en y tweedimensionaal normaal is:

a. Toets de hypothese $H_0: \rho = 0$ met $\alpha = 0,05$.

b. Geef het 95% betrouwbaarheidsinterval van ρ .

20.4. Gegeven zijn de volgende waarnemingsparen (x, y) :

x	310	315	368	301	340	308	354	313	322	314	334	377
y	78	67	84	55	78	64	76	72	82	70	67	67
x	257	348	298	287	292	345	380	330	321	265	281	246
y	56	86	60	72	68	88	93	73	72	54	78	52
x	258	230	275	324	282	340						
y	69	48	63	83	56	70						

a. Teken een spreidingsdiagram en bereken r_0 .

b. Onderstel, dat het tweedimensionaal normale model geldt en toets de hypothese $\rho = 0,6$ met $\alpha = 0,05$ (vereenvoudig de berekeningen, door voor elke grootheid de waarden met een constante te verminderen).

c. Bereken de beide regressievergelijkingen en teken de regressielijnen in het spreidingsdiagram.

20.5. Een aselechte steekproef uit een populatie, die bij benadering tweedimensionaal normaal verdeeld is, levert de volgende simultane verdeling van de grootheden x en y op:

$y \backslash x$	71	72	73	74	75	76	77	f_y
113							1	1
112					1	2		3
111			1	2	4	1		8
110			2	5	2			9
109		1	3	3				7
108		2	3					5
107	1	1						2
f_x	1	4	9	10	7	3	1	35

Bepaal het 95% en het 99% betrouwbaarheidsinterval van ρ .

20.6. Uit QUENOUILLE (189), gegevens van M. RICHARDS en A. GREIG. De bepalingen betreffen het hemoglobinegehalte (in %) en het gewicht van het hart (in % van het lichaamsgewicht) bij jonge ratten. Gegeven zijn:

$H = 100 [\log (\text{hemoglobine}) - 1]$ en $W = 100 [\log (\text{hartgewicht}) + 1]$

H	82	77	69	83	71	69	67	57	67	67	62	60
W	79	81	82	83	85	86	88	89	90	91	91	92
H	59	66	59	57	57	51	65	47	45	41	43	38
W	92	93	93	94	94	94	95	98	100	101	102	102

- Teken een spreidingsdiagram en bereken r_0 .
- Geef het 95% betrouwbaarheidsinterval voor ρ . Onder welke onderstelling is dit geldig?
- Bereken de regressievergelijking van W op H . Teken de regressielijn in het spreidingsdiagram.

Gegeven is:

$$\Sigma x = 1459, \quad \Sigma x^2 = 92115, \quad \Sigma y = 2195, \quad \Sigma y^2 = 201735, \quad \Sigma xy = 131736.$$

20.7. Uit PAGE, E. W., A blood test for estimating the week of pregnancy, Science, 105 (1947), 292—293.

x = duur van de zwangerschap in weken, y = 100 (concentratie van pitocinase in eenheden per ml. serum), $z = \log y$.

x	y	$z = \log y$	x	y	$z = \log y$	x	y	$z = \log y$
2	6	0,7782	6,5	15	1,1761	12	120	2,0792
3	5	0,6990	6,5	30	1,4771	12	140	2,1461
3	10	1,0000	7	60	1,7782	13	390	2,5911
3	7	0,8451	8	70	1,8451	13	240	2,3802
3	6	0,7782	8	50	1,6990	13	400	2,6021
3,5	7	0,8451	8	77	1,8865	14	430	2,6335
4	9	0,9542	8,5	77	1,8865	15	900	2,9542
4	10	1,0000	9	50	1,6990	15	500	2,6990
4	7	0,8451	9	80	1,9031	15,5	1100	3,0414
5	14	1,1461	11	150	2,1761	16	740	2,8692
6	28	1,4472	11	170	2,2304	16,5	850	2,9294
6	30	1,4771	12	220	2,3424	16,5	1300	3,1139
6	26	1,4150						

- Teken het spreidingsdiagram van x en y .
- Teken het spreidingsdiagram van x en $z = \log y$.
- Bereken de correlatiecoëfficiënt r_{xy} .
- Bereken de regressielijn van z op x en die van x op z , alsmede $s_{z,x}$ en $s_{x,z}$. Teken beide lijnen in het spreidingsdiagram (b).

Gegeven is: $\Sigma x = 328,5$, $\Sigma x^2 = 3660,75$, $\Sigma z = 67,3691$, $\Sigma z^2 = 142,97226229$, $\Sigma xz = 719,20525$.

20.8. Van een aselechte steekproef uit een bij benadering tweedimensionaal normaal verdeelde populatie is gegeven: $n = 27$, $\Sigma uv = -83$, $\Sigma u = -45$, $\Sigma u^2 = 219$, $\Sigma v = 81$, $\Sigma v^2 = 412$, $c_u = 5$, $c_v = 3$, $x_0 = 32,5$ en $y_0 = 60$.

- Bereken \bar{x} , \bar{y} , s_x en s_y .
- Bereken r_0 en bepaal het 95% betrouwbaarheidsinterval van ρ .
- Bepaal de regressievergelijking van y op x en bereken $s_{y,x}$.

20.9. In een steekproef van 203 elementen uit populatie A vindt men als correlatiecoëfficiënt van de grootheden x en y de waarde 0,66. In een steekproef van 153 elementen uit populatie B vindt men $r_{xy} = 0,78$. Toets de hypothese $H_0: \rho_A = \rho_B$ met $\alpha = 0,05$.

20.10. a. Toets de hypothese, dat onderstaande aselechte steekproeven stammen uit tweedimensionale populaties met dezelfde correlatiecoëfficiënt (kies $\alpha = 0,05$). Wat onderstelt U hierbij?

Steekproef	1	2	3	4	5
<i>n</i>	10	30	22	60	53
<i>r</i>	0,59	0,32	0,48	0,30	0,25

20.11. Als opgave 20.10, maar voor de volgende steekproeven:

Steekproef	1	2	3	4	5	6	7
<i>n</i>	25	33	28	58	100	88	105
<i>r</i>	0,28	0,30	0,42	0,35	0,44	0,75	0,40

20.3.* De correlatiecoëfficiënt als toetsingsgrootheid¹

Deze paragraaf — die bij eerste lezing zonder bezwaar kan worden overgeslagen — is bedoeld om de lezer te laten zien dat men, uitgaande van de correlatiecoëfficiënt, door verschillende vereenvoudigingen van het oorspronkelijke model, toetsingsmethoden verkrijgt die wij inmiddels reeds geheel los van deze coëfficiënt hebben leren kennen.

20.3.1. DE RANGCORRELATIECOËFFICIËNT VAN SPEARMAN

Beschouw een reeks van *n* waarnemingsparen (*x*, *y*), waarvan zowel de waarnemingen *x*₁, *x*₂, . . . , *x*_{*n*} als de waarnemingen *y*₁, *y*₂, . . . , *y*_{*n*} uit rangnummers bestaan (van 1 t/m *n*, maar niet noodzakelijk in deze volgorde). Dan geldt, dat

$$\Sigma x = \Sigma y = 1 + 2 + 3 + \dots + n = n(n+1)/2,$$

$$\Sigma x^2 = \Sigma y^2 = 1^2 + 2^2 + 3^2 + \dots + n^2 = n(n+1)(2n+1)/6,$$

zodat

$$\Sigma (x-\bar{x})^2 = \Sigma (y-\bar{y})^2 = \frac{n(n+1)(2n+1)}{6} - \frac{1}{n} \left\{ \frac{n(n+1)}{2} \right\}^2 = \frac{n(n+1)(n-1)}{12}.$$

Noemt men *x* - *y* = *d*, dan is

$$\Sigma d^2 = \Sigma [(x-\bar{x}) - (y-\bar{y})]^2 = \Sigma (x-\bar{x})^2 + \Sigma (y-\bar{y})^2 - 2 \Sigma (x-\bar{x})(y-\bar{y}) = 2 [n(n+1)(n-1)/12 - \Sigma (x-\bar{x})(y-\bar{y})],$$

zodat

$$\Sigma (x-\bar{x})(y-\bar{y}) = n(n+1)(n-1)/12 - \Sigma d^2/2.$$

Uit het voorgaande volgt, door substitutie in formule (18.16):

$$r_r = \frac{\frac{n(n+1)(n-1)}{12} - \frac{\Sigma d^2}{2}}{\frac{n(n+1)(n-1)}{12}} = 1 - \frac{6 \Sigma d^2}{n(n^2-1)}.$$

Deze formule is identiek aan (10.27) op blz. 253, de formule voor de rangcorrelatiecoëfficiënt van SPEARMAN, die dus beschouwd kan worden als een speciaal geval van de correlatiecoëfficiënt.

20.3.2. DE *t*-TOETS OP DE GEMIDDELDEN VAN TWEE ASELECTE STEEKPROEVEN

Beschouw de verdelingen van de twee aselecte steekproeven in tabel 20.5. Wanneer zij afkomstig zijn uit ongeveer normaal verdeelde populaties met dezelfde spreiding, dan kan de hypothese $\mu_1 = \mu_2$ volgens STUDENT worden

Naar NASS (183).

getoetst met als toetsingsgrootheid volgens (16.4):

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

met $\nu = n_1 + n_2 - 2 = n - 2$.

Onderstel nu, dat men arbitrair aan alle elementen in de eerste steekproef de waarde $y = 1$ en aan alle elementen in de tweede steekproef de waarde $y = 0$ toekent. Dan is dus $\Sigma y = \Sigma y^2 = n_1$, zodat $\Sigma (y - \bar{y})^2 = \Sigma y^2 - (\Sigma y)^2/n = n_1 - (n_1)^2/n = [n_1(n_1 + n_2) - (n_1)^2]/n = n_1 n_2/n$. Onderscheidt men de waarden van x in de eerste en de tweede steekproef door toekenning van de index 1, resp. 2, dan geldt verder:

$$\Sigma x = \Sigma x_1 + \Sigma x_2 \text{ en } \Sigma xy = \Sigma x_1$$

zodat

$$\begin{aligned} \Sigma (x - \bar{x})(y - \bar{y}) &= \Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n} = \Sigma x_1 - \frac{(\Sigma x_1 + \Sigma x_2) n_1}{n_1 + n_2} \\ &= \frac{(n_1 + n_2) \Sigma x_1 - (\Sigma x_1 + \Sigma x_2) n_1}{n_1 + n_2} = \frac{n_2 \Sigma x_1 - n_1 \Sigma x_2}{n_1 + n_2} \\ &= \frac{n_1 n_2 \bar{x}_1 - n_1 n_2 \bar{x}_2}{n_1 + n_2} = \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_1 - \bar{x}_2). \end{aligned}$$

Substitueert men nu de verkregen uitkomsten in (18.16), dan levert dit op

$$(20.27) \quad r_{\text{bis}} = \frac{n_1 n_2 (\bar{x}_1 - \bar{x}_2) / (n_1 + n_2)}{\sqrt{\Sigma (x - \bar{x})^2 \cdot n_1 n_2 / (n_1 + n_2)}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\Sigma (x - \bar{x})^2}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

Men noemt deze bijzondere vorm van de correlatiecoëfficiënt, die op twee aselechte steekproeven kan worden toegepast, de 'biserial'-correlatiecoëfficiënt¹. r_{bis} is positief voor $\bar{x}_1 > \bar{x}_2$, nul voor $\bar{x}_1 = \bar{x}_2$ en negatief voor $\bar{x}_1 < \bar{x}_2$. De hypothese $\rho_{\text{bis}} = 0$ kan volgens (20.10) worden getoetst door als toetsingsgrootheid te nemen

$$(20.28) \quad t = r_{\text{bis}} \sqrt{n-2} / \sqrt{1 - r_{\text{bis}}^2}, \text{ met } n - 2 \text{ vrijheidsgraden.}$$

Wanneer echter $\rho_{\text{bis}} = 0$, geldt tevens $\mu_1 = \mu_2$. Bij toetsing via r_{bis} van de hypothese $\rho_{\text{bis}} = 0$ verkrijgt men dan ook precies dezelfde uitkomst als bij de in de eerste alinea genoemde toetsing via $\bar{x}_1 - \bar{x}_2$ van de hypothese $\mu_1 = \mu_2$.

20.3.3. DE TOETS VAN WILCOXON VOOR TWEE ASELECHTE STEEKPROEVEN

Bij deze toets bestaan de n -waarden van de grootheid x in de steekproeven 1 en 2 uit de rangnummers 1 t/m n , zodat

$$\begin{aligned} \Sigma x &= 1 + 2 + 3 + \dots + n = (n+1)/2, \\ \Sigma x^2 &= 1^2 + 2^2 + 3^2 + \dots + n^2 = n(n+1)(2n+1)/6. \end{aligned}$$

Hieruit volgt, dat

$$\Sigma (x - \bar{x})^2 = n(n+1)(n-1)/12 \text{ (zie 20.3.1) en } \Sigma x_2 = n(n+1)/2 - \Sigma x_1,$$

¹ Zie voor een uitvoeriger beschouwing betreffende deze coëfficiënt b.v. WALKER & LEV (85)

Tabel 20.5. Frequentieverdelingen van de grootheid x voor twee aselechte steekproeven

x	Steekproef 2	Steekproef 1
	$y = 0$	$y = 1$
8	1	1
7	3	1
6	6	2
5	8	4
4	3	6
3	2	3
2	1	2
1	1	1
	$n_2 = 25$	$n_1 = 20$

zodat

$$\begin{aligned} \bar{x}_1 - \bar{x}_2 &= \frac{\Sigma x_1}{n_1} - \frac{n(n+1)/2 - \Sigma x_1}{n_2} = \frac{n_2 \Sigma x_1 - n_1 n(n+1)/2 + n_1 \Sigma x_1}{n_1 n_2} \\ &= \frac{n}{n_1 n_2} \left[\Sigma x_1 - n_1 (n+1)/2 \right]. \end{aligned}$$

Substitutie van $\Sigma(x - \bar{x})^2$ en $\bar{x}_1 - \bar{x}_2$ in (20.27) levert op

$$(20.29) \quad r_W = \frac{\frac{n}{n_1 n_2} [\Sigma x_1 - n_1 (n+1)/2]}{\sqrt{n(n+1)(n-1)/12}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \frac{\Sigma x_1 - n_1 (n+1)/2}{\sqrt{n_1 n_2 (n+1)(n-1)/12}}.$$

Maakt men voor deze bijzondere vorm van r_{bis} bij toetsing van de hypothese $\rho_{bis} = 0$ gebruik van de normale benadering volgens (20.13), en schrijft men voor Σx_1 , de som van de rangnummers in steekproef 1, S_1 , dan vindt men

$$(20.30) \quad \underline{T} = r_W \sqrt{n-1} = \frac{S_1 - n_1 (n+1)/2}{\sqrt{n_1 n_2 (n+1)/12}}.$$

Dit is echter de formule voor de normale benadering van de toets van WILCOXON, met als toetsingsgroottheid S_1 in plaats van W [zie de formules (10.12) en (10.9)].

20.3.4. DE χ^2 -TOETS OP DE 2×2 TABEL

Wij beschouwen nogmaals de twee verdelingen in tabel 20.5 en zetten deze om in een 2×2 tabel, door aan $x = 5$ t/m 8 de waarde 1 en aan $x = 1$ t/m 4 de waarde 0 te verbinden (zie tabel 20.6).

Tabel 20.6. 2×2 tabel voor de verdelingen in tabel 20.5

	Steekproef 2 $y = 0$	Steekproef 1 $y = 1$	
$x = 1$ (5 t/m 8)	18 (a)	8 (c)	26
$x = 0$ (1 t/m 4)	7 (b)	12 (d)	19
	25	20	45

Nu is: $\Sigma x = a + c = \Sigma x^2$, $\Sigma y = c + d = \Sigma y^2$, $\Sigma xy = c$. Hieruit volgt, dat

$$\Sigma(x - \bar{x})^2 = \Sigma x^2 - (\Sigma x)^2/n = (a + c) - (a + c)^2/n = (a + c)(b + d)/n$$

(bedenk, dat $n = a + b + c + d$). Op overeenkomstige wijze vindt men, dat $\Sigma(y - \bar{y})^2 = (a + b)(c + d)/n$, terwijl verder

$$\Sigma(x - \bar{x})(y - \bar{y}) = c - \frac{(a+c)(c+d)}{n} = \frac{c(a+b+c+d) - (a+c)(c+d)}{n} = \frac{bc - ad}{n}.$$

Door substitutie in (18.16) verkrijgt men nu als speciaal geval van de correlatiecoëfficiënt de zg. phi-coëfficiënt

$$(20.31) \quad r_{\text{phi}} = \frac{(bc - ad)/n}{\sqrt{\frac{(a+c)(b+d)}{n} \cdot \frac{(a+b)(c+d)}{n}}} = \frac{bc - ad}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}.$$

Toetst men de hypothese $\rho_{\text{phi}} = 0$ via de normale benadering, dan volgt

$$(20.32) \quad T = r_{\text{phi}} \sqrt{n-1} = \frac{(bc - ad) \sqrt{n-1}}{\sqrt{(a+b)(c+d)(a+c)(b+d)}};$$

d.i. formule (9.10) voor de normale benadering van de toets van FISHER op de 2×2 verdeling.

20.3.5. DE CORRELATIECOËFFICIËNT EN DE t -TOETS OP VERSCHILLEN VAN GEPAARDE WAARNEMINGEN

Bij de in 15.3 besproken t -toets op verschillen tussen gepaarde waarnemingen is de toetsingsgrootheid

$$t = \frac{\bar{v} \sqrt{n}}{s_{\bar{v}}} = \frac{\bar{v}}{s_{\bar{v}}}, \text{ waarin, } s_{\bar{v}} = \sqrt{\frac{\Sigma(v - \bar{v})^2}{n(n-1)}}.$$

Men kan echter schrijven

$$\begin{aligned} \Sigma(v - \bar{v})^2 &= \Sigma[(x_1 - x_2) - (\bar{x}_1 - \bar{x}_2)]^2 = \Sigma[(x_1 - \bar{x}_1) - (x_2 - \bar{x}_2)]^2 \\ &= \Sigma(x_1 - \bar{x}_1)^2 + \Sigma(x_2 - \bar{x}_2)^2 - 2\Sigma(x_1 - \bar{x}_1)(x_2 - \bar{x}_2). \end{aligned}$$

Beschouwt men echter de gepaarde waarnemingen als de waarnemingsparen (x_1, x_2) , dan is

$$r_{x_1, x_2} = r = \frac{\Sigma(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)}{\sqrt{\Sigma(x_1 - \bar{x}_1)^2 \cdot \Sigma(x_2 - \bar{x}_2)^2}},$$

zodat

$$\Sigma(x_1 - \bar{x}_1)(x_2 - \bar{x}_2) = r \sqrt{\Sigma(x_1 - \bar{x}_1)^2 \cdot \Sigma(x_2 - \bar{x}_2)^2}.$$

Uit het voorgaande volgt:

$$s_{\bar{v}}^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n(n-1)} + \frac{\Sigma(x_2 - \bar{x}_2)^2}{n(n-1)} - 2r \sqrt{\frac{\Sigma(x_1 - \bar{x}_1)^2}{n(n-1)} \cdot \frac{\Sigma(x_2 - \bar{x}_2)^2}{n(n-1)}},$$

en daar $s_{x_1}^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n-1}$ en $s_{x_2}^2 = \frac{\Sigma(x_2 - \bar{x}_2)^2}{n-1}$ is:

$$s_{\bar{v}}^2 = s_{x_1}^2/n + s_{x_2}^2/n - 2r \sqrt{s_{x_1}^2/n \cdot s_{x_2}^2/n}$$

of (zie 13.2)

$$(20.33) \quad s_{\bar{v}}^2 = s^2_{\bar{x}_1} + s^2_{\bar{x}_2} - 2r s_{\bar{x}_1} s_{\bar{x}_2}.$$

Bij de proefopzet met twee verwante steekproeven (gepaarde waarnemingen) kan men dus ook schrijven

$$(20.34) \quad t = \bar{v} \sqrt{s^2_{\bar{x}_1} + s^2_{\bar{x}_2} - 2r s_{\bar{x}_1} s_{\bar{x}_2}} \quad (v = n - 1)$$

en deze formule demonstreert dat men bij bepaalde waarden van s_{x_1} , s_{x_2} en \bar{v} een grotere waarde van t verkrijgt, naarmate er een hogere *positieve* correlatie tussen de beide waarnemingsreeksen aanwezig is.

Bij een proefopzet met twee aselechte steekproeven van dezelfde omvang n geldt (als $\bar{v} = \bar{x}_1 - \bar{x}_2$):

$$(20.35) \quad t = v \sqrt{s^2_{x_1} + s^2_{x_2}} \quad [v = 2(n-1)].$$

Vergelijkt men (20.34) met (20.35) dan blijkt dat de eerste proefopzet doeltreffender is dan de tweede, mits er een zodanige positieve correlatie tussen de gepaarde waarnemingen is, dat de hieruit resulterende lagere waarde van $s_{\bar{v}}$ van meer betekenis is dan het verlies aan vrijheidsgraden [$n - 1$ i.p.v. $2(n - 1)$].

Tenslotte vermelden wij, dat uit (20.33) als alternatieve formule voor de correlatiecoëfficiënt berekend uit n waarnemingsparen (x, y) volgt

$$(20.36) \quad r_{xy} = \frac{s_x^2 + s_y^2 - s_{x-y}^2}{2s_x s_y} = \frac{\Sigma (x-\bar{x})^2 + \Sigma (y-\bar{y})^2 - \Sigma (x-y)^2}{2 \sqrt{\Sigma (x-\bar{x})^2 \cdot \Sigma (y-\bar{y})^2}}$$

20.4. Functionele relaties

Het kan zijn dat op grond van theoretische overwegingen vaststaat, dat er tussen de grootheden x en y de functionele relatie $y = A + Bx$ bestaat. Zou men de metingen van deze grootheden kunnen verrichten zonder dat daarbij variatie tengevolge van 'meetfouten' optreedt (of met een variatie die zo klein is, dat zij kan worden verwaarloosd), dan zouden alle punten (x, y) precies op de rechte $y = A + Bx$ liggen. Men kan dan deze (ene) rechte ook aangeven met de vergelijking

$$x = y/B - A/B.$$

In de praktijk zullen de waarnemingen vaak wel een zekere toevallige variatie bezitten. Kan men de waarden van één grootheid, zeg van x , zeer nauwkeurig bepalen of kiezen, zodat in feite slechts de grootheid y variabel is, dan kan men de beste schatting van de functionele relatie tussen x en y verkrijgen door de best-passende regressielijn van y op x te berekenen. Wanneer echter de waarnemingen van beide grootheden een zekere — maar bekende — onbetrouwbaarheid bezitten, kan men volgens LINDLEY (225) de beste schatting van de functionele relatie verkrijgen door te nemen

$$(20.37) \quad b_{yx} = m + \sqrt{m^2 + k}$$

en

$$(20.38) \quad a_{yx} = \bar{y} - b\bar{x},$$

waarin

$$(20.39) \quad m = \frac{\Sigma (y - \bar{y})^2 - k \Sigma (x - \bar{x})^2}{2 \Sigma (x - \bar{x})(y - \bar{y})},$$

$$(20.40) \quad k = \frac{\text{de variantie van de enkele waarneming van } y}{\text{de variantie van de enkele waarneming van } x} = \frac{\text{var}(y)}{\text{var}(x)}.$$

Als voorbeeld beschouwen wij de vijf volgende waarnemingsparen:

x	1,1	2,2	2,9	4,1	5,0
y	0,5	2,5	2,5	4,5	4,8

Hiervoor geldt dat $\Sigma (x - \bar{x})^2 = 9,45$, $\Sigma (x - \bar{x})(y - \bar{y}) = 10,46$, $\Sigma (y - \bar{y})^2 = 12,23$, $\bar{x} = 3,06$, $\bar{y} = 2,96$. Als de waarnemingen van x een zeer grote betrouwbaarheid bezitten vergeleken met die van y ,

kan men het best de regressievergelijking van y op x berekenen. Men verkrijgt dan als beste schatting van de functionele relatie: $y = 1,11 x - 0,43$. Zijn daarentegen de waarnemingen van y zeer betrouwbaar en die van x niet, dan berekent men de regressievergelijking van x op y : $x = 0,855y + 0,53$ en zet deze vervolgens in de vergelijking van y op x om. Men vindt dan $y = 1,17 x - 0,62$.

Zijn beide waarnemingen even onbetrouwbaar, zodat $k = 1$, dan geven de formules van LINDLEY de beste oplossing:

$$m = \frac{12,23 - 9,45}{2 (10,46)} = 0,1329 ,$$

$$b = 0,1329 + \sqrt{(0,1329)^2 + 1} = 1,14,$$

$$a = 2,96 - (1,14) (3,06) = -0,53.$$

De schatting van de lineaire functie luidt dus $y = 1,14 - 0,53$ en hieruit blijkt, dat bij $k = 1$ met de methode van LINDLEY een regressielijn wordt verkregen, die midden tussen de beide andere regressielijnen ligt.

JESSOP (224) maakt omtrent de toepassing van het voorgaande op biologische grootheden de volgende opmerkingen: 'The word regression itself was first used in connection with the following biological example. The data consisted of pairs of measurements of 'height of father' and 'height of son'. Then, calling the former x and the latter y , two distinct lines are obtained. Each regression line serves the purpose of enabling the mean value of one measurement to be found if the other is given. The question of a unique 'true' relationship does not arise; we have no body of theory, as we might have were the study physical, to assert that a functional relationship $y = f(x)$ should exist and that the scatter of points on the graph is the result of observational error; the purpose of the two lines is to describe the characteristics of a bivariate normal population in so far as they can be judged from a sample. No single line could tell us as much as the two lines do. . . . In the preceding, error is very unimportant compared with biological variation; it is the latter that was minimised to get the best estimate of the mean y for a given x and the mean x for a given y . However, when our data are physical measurements biological variation does not enter the picture, and it is in such situations, among others, that the question of a single-line representation does arise.'

Het door JESSOP zo duidelijk beschreven onderscheid tussen de situatie, waarin men met één regressielijn kan werken en de situatie, waarin twee regressielijnen nodig zijn, wordt vaak niet ingezien. Wij geven hiervan enkele voorbeelden.

Voorbeelden

20.7. Uit N.T.v.G., 100 (1956), 2784—2789. Men beschikt over 42 waarnemingsparen betreffende stijgingen en dalingen van de calciumuitscheiding in de urine en van het calciumgehalte in het serum tijdens toediening en na het staken van de toediening van hormoonstoffen¹. De veranderingen in het serumcalciumgehalte zijn in een spreidingsdiagram uitgezet tegen de gelijktijdige veranderingen van de calciumuitscheiding in de urine per 24 uur. Men berekent een correlatiecoëfficiënt $r_0 = 0,87$ met $P_D \ll 0,001$. Vervolgens wordt gezegd: 'Wij hebben verder een lijn geconstrueerd, die zo goed mogelijk aan de puntenwolk is aangepast en die het verband aangeeft tussen de veranderingen in de calciëmie en de calciurie. . . . Aangezien beide grootheden fouten kunnen bevatten, die van eenzelfde orde van grootte zijn, kan de richtingscoëfficiënt van deze lijn niet volgens de veelal gebruikte methode van de som der kleinste kwadraten worden berekend.'

Afgezien van de vraag, óf de onbetrouwbaarheid van de metingen van de beide grootheden van dezelfde grootte-orde is, staat wel vast dat deze grootheden een biologische variatie zullen bezitten die veel groter is dan de variatie tengevolge van toevallige meetfouten. Van een functionele relatie is dan ook geen sprake en men dient wel degelijk met twee regressielijnen te werken.

20.8. Uit Ned. Mil. Gen. Tijdschrift, 12 (1959), 100—113. Dit artikel handelt in hoofdzaak over de correlatie tussen de lichaamslengte (L) en de zithoogte (Z) bij vliegers. De auteur vindt bij 137 vliegers voor de correlatie tussen L en Z de waarde $r_0 = 0,75$ en als regressievergelijking van L op Z : $\hat{L} = 1,6Z + 29,0$ (cm). Deze vergelijking zet hij vervolgens om in de vergelijking van Z op L : $1,6Z = L - 29,0$, $Z = L/1,6 - 29,0/1,6$, $\hat{Z} = 0,625L - 18,13$ (cm).

Wij hebben hier te doen met een typisch correlatieprobleem, waarbij kan worden aangenomen dat de metingen van de lengte en de zithoogte relatief zeer betrouwbaar zijn, zodat de punten (L, Z) tengevolge van biologische variatie niet op een rechte liggen. De regressievergelijking van Z op L kan *niet* door omzetting uit die van L op Z worden verkregen, maar dient afzonderlijk te worden berekend. Maakt men hierbij gebruik van de gegevens in de figuur op blz. 106 (waarbij rekening dient te worden gehouden met het feit, dat x en y hulp-grootheden zijn, zodat $Z = 0,5x + 86$ en $L = 0,5y + 165$), dan vindt men: $\hat{Z} = 0,348L + 31,23$ (cm).

Dit artikel bevat overigens nog een groot aantal andere min of meer ernstige fouten. Wij laten het gaarne aan de lezer over om de foutieve gegevens onderaan blz. 109 te corrigeren, de denkfout bovenaan blz. 110 te vinden en de benaderingsvergelijking te verbeteren (en te beoordelen), die later op deze blz. wordt opgesteld. Verder wijzen wij erop, dat de formules I t/m V op blz. 111 weliswaar de symbolen x en y bevatten, maar dat de achter deze formules geplaatste uitkomsten niet op de reeds genoemde hulpgrootheden x en y slaan, maar op de grootheden Z en L betrekking hebben. De formules III en IV geven niet de standaardderivaties (l), maar de varianties van Z en L , maar de uitkomsten achter deze formules zijn wel standaarddeviaties. Formule VI (waarin men waarschijnlijk voor F_p moet lezen F_ρ en voor σ_p moet lezen σ_ρ) is blijkens de er achter geplaatste uitkomst bedoeld als formule voor $s_{Z,L}$. Het is in ieder geval niet de 'standaard-fout van de correlatie-coëfficiënt.

¹ Wij laten buiten beschouwing, dat de 42 waarnemingen 25 patiënten betreffen.

VERDELINGSVRIJE EN KLASSIEKE METHODEN

21.1. Inleiding

De normale verdeling geniet een grote bekendheid en deze heeft wel tot de misvatting geleid, dat het voorkomen van dit verdelingstype 'normaal' is. Nu zijn er wel een aantal biologische eigenschappen, die ongeveer normaal verdeeld blijken te zijn, maar afwijkingen van de normaliteit zijn in het geheel geen uitzondering. GALTON constateerde reeds in 1878 dat naar rechts scheve verdelingen (die echter door het nemen van logaritmen pseudo-normaal worden en daarom wel log-normaal worden genoemd) in de biologie veelvuldig voorkomen (zie ook 23.2). Voorbeelden hiervan zijn o.m. de verdelingen van bloeddrukken, bilirubinegehalten, gewichten, enz. Verder kan men op ieder toepassingsgebied van de statistiek verdelingen vinden, die sterk van de normale afwijken. Enkele sprekende voorbeelden zijn de leeftijdsverdeling van de bevolking in een grote stad en de verdeling van de verzuimfrequenties in een bepaalde periode van alle arbeiders in een bedrijf. Overigens dient men te bedenken, dat duidelijke afwijkingen van de normaliteit ook een gevolg kunnen zijn van het feit dat de bestudeerde populatie heterogeen is, of kunnen ontstaan door bepaalde systematische fouten in het waarnemingsproces.

Uit het voorgaande volgt, dat de statistiek zich soms van de op normaliteit gebaseerde klassieke methoden zal kunnen bedienen en in andere gevallen van de verdelingsvrije methoden gebruik zal dienen te maken. In dit hoofdstuk vergelijken wij eerst enkele aspecten van beide soorten methoden en daarna zullen wij trachten enkele globale richtlijnen te geven, die in de praktijk bij de keuze van een methode behulpzaam kunnen zijn.

21.2. Klassieke en verdelingsvrije methoden

De klassieke methoden zijn afgeleid onder de onderstelling, dat de populatieverdelingen normaal zijn en deze onderstelling behoeft bij de verdelingsvrije methoden niet te worden gemaakt. Dit verschil heeft de volgende consequenties:

1. Voor normaal verdeelde populaties zijn de verdelingsvrije methoden in het algemeen minder doeltreffend dan de klassieke, daar zij van

minder gegevens gebruik maken. Treden duidelijke afwijkingen van de normaliteit op, dan kan precies het tegendeel gelden.

Op blz. 222 t/m 224 hebben wij gezien, dat men bij toepassing van de rang-

Paar	A	B	B-A
1	16,2	17,3	+1,1
2	12,9	16,5	+3,6
3	15,8	17,7	+1,9
4	17,0	19,2	+2,2
5	14,6	14,4	-0,2

tentetoets op onderstaande verschillen tussen de gepaarde waarnemingen A en B een tweezijdige overschrijdingskans vindt van 0,125. Uit voorgaande onderzoeken is echter gebleken, dat deze verschillen gewoonlijk ongeveer normaal verdeeld zijn. Past men op grond hiervan de toets van STUDENT toe, dan vindt men $\bar{v} = 1,72$, $s_{\bar{v}} = 0,6272$, $t_0(4) = 2,74$,

met een tweezijdige overschrijdingskans die iets groter is dan 5%, daar $t_{0,975}(4) = 2,776$. De ver-

delingsvrije overschrijdingskans is dus ongeveer tweemaal zo groot als de klassieke.

Het volgende voorbeeld (uit M. GRUBER, Thiamine deficiëntie en koolhydraat stofwisseling, Dissertatie, Utrecht, 1952) betreft de vetgehalten (in pct van het lichaamsgewicht) bij twee groepen duiven, die onder verschillende omstandigheden verkeerden:

Groep IV:	1,4	2,2	1,7	1,7	2,6	2,5	7,8	2,3
Groep V:	3,0	3,4	6,4	2,9	2,2	7,9		

Beide groepen vertonen duidelijke uitbijters en ook op grond van andere overwegingen is de normaliteitsonderstelling aanvechtbaar. Past men nu de toets van STUDENT toe, dan vindt men een tweezijdige overschrijdingskans van ongeveer 0,20, terwijl de toets van WILCOXON een exacte overschrijdingskans in de nabijheid van 0,05 oplevert (0,0426 of 0,0592, afhankelijk van de werkelijke waarden van de twee gelijken ad 2,2).

2. Bij toepassing van een klassieke methode dient men zich af te vragen of het normale model voor de betrokken populatie(s) geldig is. Bij het gebruik van een verdelingsvrije methode is dit niet nodig, maar wij merken hierbij op, dat men bij een volledige statistische analyse eventuele verschillen in de vorm van de verdelingen of verschillen in variabiliteit (die trouwens eveneens op een verschil in vorm kunnen wijzen) niet steeds buiten beschouwing kan laten. Deze verschillen kunnen namelijk op zichzelf van betekenis zijn voor een juiste interpretatie van de waarnemingsuitkomsten.

Wij ontlenen een voorbeeld aan de dissertatie van H. J. DOKTER, Antithyreoïde stoffen en aangeboren misvormingen, Amsterdam, 1958. Het betreft de waargenomen aantallen corpora lutea bij (A) een groep van 13 wijfjesratten, die gedurende de zwangerschap per os methylthiouracil toegediend kreeg, en (B) een contrôlegroep van 13 wijfjesratten, die 'onder dezelfde omstandigheden verkeerde als de proefdieren', maar dit middel niet kreeg. De aantallen zijn, gerangschikt naar grootte:

A. Proefgroep:	7	8	9	9	9	9	11	12	12	14	14	20	23
B. Contrôles :	9	10	10	10	11	12	12	12	13	13	13	13	13.

De onderzoeker geeft als conclusie: 'Toepassing van de toets van WILCOXON op de gegevens . . . gaf als resultaat, dat de aantallen corpora lutea van de twee groepen niet significant verschilden ($P = 0,4$).

De toegepaste toets is echter niet gevoelig voor het duidelijke verschil in spreiding van de twee groepen. De varianties en spreidingen zijn nl.: $s_A^2 = 22,5769$, $s_A = 4,75$, resp. $s_B^2 = 2,0898$, $s_B = 1,45$. Met de F -toets vindt men dat $F_0(12, 12) = 10,8$ met $P_D < 0,001$ en hoewel deze toets slechts bij benadering geldt is deze uitkomst zo uitgesproken, dat zij niet kan worden verwaarloosd. Er zijn nu verschillende conclusies mogelijk, t.w.: (a) de groepen zijn aselekt samengesteld, maar de behandeling met methylthiouracil leidt tot een grotere spreiding van de aantallen corpora lutea, (b) de groepen zijn aselekt samengesteld, maar behalve de experimentele factor (al of niet toedienen van methylthiouracil) zijn nog andere factoren van invloed geweest, en (c) de groepen zijn niet aselekt gekozen. Wij kunnen niet beoordelen, of mogelijkheid (a) plausibel is, maar zowel (b) als (c) houden in, dat men ook het verschil in niveau van de twee groepen met reserve dient te beschouwen.

3. Door het gebruik van het gemiddelde en de spreiding vereisen de klassieke methoden meer rekenwerk dan de verdelingsvrije, waarbij alleen rangnummers moeten worden toegekend. Hiertegenover staat, dat bij de klassieke methoden toetsen en schatten a.h.w. hand in hand gaan en dat men in vele gevallen bij benadering kan vaststellen welke steekproefomvang dient te worden gekozen om een bepaalde uitspraak met tevoren gekozen onbetrouwbaarheid te kunnen doen (zie bv. 15.1.3, 15.2.3, 16.1.3 en 16.5.3). De verdelingsvrije rangorde-methoden kunnen soms wel dienen voor het verkrijgen van een schattingsinterval [zie bv. HEMELRIJK (18)], maar zij zijn toch overwegend toetsings-technieken.

Wanneer het gebruik van een klassieke methode verantwoord wordt geacht is naar onze mening het meerdere rekenwerk geen factor van betekenis. Bij de meeste onderzoeken op medisch-biologisch terrein kost het verzamelen van de waarnemingen weken, maanden of zelfs jaren en het doet er dan weinig toe, of men één of twee dagen (of weken) aan de statistische analyse besteedt. Overigens heeft men bij een handige aanpak van de berekeningen (gebruik van hulpgrootheden, gebruik van een kwadratentafel en in het gunstigste geval van een rekenmachine) voor de klassieke methoden niet zo veel meer tijd nodig, als wel eens wordt gedacht. De toepassing van bv. de toets van WILCOXON op twee steekproeven van grote omvang met veel gelijken vereist ook het gebruik van een rekenschema (zie tabel 10.10), en de berekeningen verlopen niet veel sneller dan die voor de toets van STUDENT.

In tabel 21.1 is een overzicht gegeven van een aantal klassieke methoden in de hoofdstukken 14 t/m 20 en de hiermee corresponderende kwantitatieve verdelingsvrije methoden in hoofdstuk 10. In de meeste gevallen kan men ook kwalitatieve verdelingsvrije methoden gebruiken (zoals de χ^2 -toets), maar daar deze in het algemeen minder doeltreffend zijn dan de rangorde-methoden hebben wij ze niet in de tabel opgenomen. In hoofdstuk 22 geven wij nog een inleiding tot de zg. statistische proefopzetten, waarbij gebruik wordt gemaakt van de variantie-analyse, en die zich niet of niet zo goed lenen voor een analyse langs de verdelingsvrije weg.

Tabel 21.1. Overzicht van enkele corresponderende klassieke en verdelingsvrije toetsen

Waarnemingsmateriaal		Klassiek	Verdelingsvrij	
Verwante steekproeven	2	Toets van STUDENT op het gemiddelde verschil	Rang-tekentoets	1
	h	Tweevoudige variantie-analyse	Toets van FRIEDMAN (m rangschikkingen)	2
Aselecte steekproeven	2	Toets van STUDENT op de gemiddelden	Toets van WILCOXON	3
	h	Enkelvoudige variantie-analyse	Toets van KRUSKAL en WALLIS	4
Waarnemingsparen	Regressie	Regressie-analyse	Toets van TERPSTRA	5
	Correlatie	Correlatierekening	Rangcorrelatie (vlgs. SPEARMAN of KENDALL)	6

21.3. De keuze van een klassieke of verdelingsvrije methode

De klassieke methoden zijn gebaseerd op de onderstelling dat de populatieverdelingen normaal zijn. Voor pseudo-normale verdelingen (waarbij dus de afwijkingen van de normaliteit betrekkelijk klein zijn) geven zij zelfs bij een kleine steekproefomvang uitkomsten, die vrijwel exact gelden. Bij verdelingen met grotere afwijkingen van de normaliteit verkrijgt men bij toepassing van een klassieke methode slechts een uitkomst, die bij benadering juist is. Men kan dan omtrent deze benadering slechts zeggen dat zij beter wordt naarmate deze afwijkingen minder ernstig en de steekproeven groter zijn. Deze laatste uitspraak is noodgedwongen vaag. Niet-normale verdelingen kunnen immers een enorme verscheidenheid van gedaanten vertonen en voor elk verdelingstype geldt dat een bepaalde steekproeffunctie bij iedere steekproefomvang een andere kansverdeling bezit. Nu zijn er wel verschillende steekproefexperimenten uitgevoerd waarbij geconstateerd kon worden, dat de klassieke methoden reeds bij steekproeven van enkele tientallen elementen vrij ongevoelig zijn voor bv. scheefheid, maar deze zijn o.i. nog te fragmentarisch om tot een duidelijker uitspraak te kunnen leiden.¹ Een andere moeilijkheid is, dat men in de regel de populatieverdelingen niet kent, zodat het oordeel omtrent de vorm van deze verdelingen op de waargenomen verdelingen moet worden gebaseerd. Het is duidelijk, dat hierbij de steekproefomvang van grote betekenis is.

¹ Zie bv. BENEPE (127), BARTLETT (122), BOX (131), COCHRAN (137), HEY (168) en PEARSON (185).

Beschikt men over steekproeven van voldoende grote omvang, dan kan men:

1. De hypothese toetsen, dat deze steekproef afkomstig is uit een normaal verdeelde populatie. Wij bespreken enkele van deze *normaliteitstoetsen* in 21.3.

Past men een normaliteitstoets toe, dan betekent het feit dat men de nulhypothese niet kan verwerpen, uiteraard niet dat de populatieverdeling normaal is. Bij het gebruik van een conservatieve onbetrouwbaarheidsdrempel (zoals bv. $\alpha = 0,10$) kan men naar onze mening bij een niet significante uitkomst van zo'n toets zonder bezwaar een klassieke methode kiezen.

2. Nagaan, of één der in 21.4 behandelde transformaties tot een verdeling leidt, die wél ongeveer normaal is.

Wanneer een verdeling niet normaal is en dit ook door een transformatie niet wordt, kan men — indien dit mogelijk is — gewoonlijk beter een verdelingsvrije methode kiezen. Staat zo'n methode niet ter beschikking dan weet men, dat de uitkomst van de toegepaste klassieke methode met reserve moet worden gehanteerd.

Beschikt men uitsluitend over kleine steekproeven, dan kan men geen of slechts een inadequate (te grove) frequentieverdeling opstellen, zodat men zich uit de waarnemingen geen bevredigend beeld van de populatieverdeling kan vormen. De normaliteitstoetsen zijn dan niet toepasbaar of zeer ongevoelig voor afwijkingen van de normaliteit. Weet men uit ervaring, dat de bestudeerde grootheid gewoonlijk bij benadering normaal verdeeld is (of dat zij een verdeling volgt die door transformatie in een normale overgaat) en vertonen de eigen waarnemingen geen duidelijke contra-indicaties (zoals bv. uitbijters), dan kan men de normaliteitsonderstelling nog wel maken. Is dit niet zo, of zijn er aanwijzingen die op afwijkingen van de normaliteit wijzen, dan kiest men zo mogelijk een verdelingsvrije methode.

Uit de voorgaande richtlijnen blijkt dat bij de keuze tussen de twee soorten methoden de steekproefomvang een belangrijke rol speelt. Bij kleine steekproeven zal men eerder een verdelingsvrije methode gebruiken, terwijl bij grote steekproeven in vele gevallen een verantwoord gebruik van een klassieke methode mogelijk zal zijn. In dit verband zouden wij voor $n < 40$ van kleine steekproeven willen spreken en bij $n > 100$ van grote. Voor het tussenliggende gebied ligt de zaak moeilijker. De keuze zal hier vaak afhangen van de kennis die men reeds omtrent de bestudeerde grootheid of grootheden bezit.

Voorbeelden:

21.1. In tabel 21.2a staan waarnemingen die betrekking hebben op tien aselechte steekproeven van 5 elementen uit dezelfde populatie. Het is duidelijk dat één steekproef van deze kleine omvang te weinig informatie verschaft omtrent de vorm van de populatieverdeling. Beziet men de steekproeven af-

zonderlijk, dan zal men bv. geneigd zijn de waarde 29 in steekproef *f* en de waarde 10 in steekproef *g* als uitbijter te kwalificeren. Beschouwt men echter de uitkomsten als één steekproef van 50 elementen, dan kan men de frequentieverdeling in tabel 21.2*b* opstellen.

Tabel 21.2*a*. Tien aselechte steekproeven met $n = 5$ uit dezelfde populatie

<i>a.</i>	17	19	23	13	11	<i>f.</i>	19	29	19	18	20
<i>b.</i>	15	25	22	22	18	<i>g.</i>	20	20	10	19	21
<i>c.</i>	20	14	20	15	22	<i>h.</i>	18	25	23	21	19
<i>d.</i>	16	16	24	19	24	<i>i.</i>	17	22	24	19	18
<i>e.</i>	21	24	17	15	26	<i>j.</i>	20	18	25	16	23

waarde 10 in steekproef *g* als uitbijter te kwalificeren. Beschouwt men echter de uitkomsten als één steekproef van 50 elementen, dan kan men de frequentieverdeling in tabel 21.2*b* opstellen.

Tabel 21.2*b*. Frequentieverdeling van de 50 waarden in tabel 21.2*a*

x	10—12	13—15	16—18	19—21	22—24	25—27	28—30	Totaal
<i>f</i>	2	5	11	16	11	4	1	50

Bij deze — nog niet zo grote — steekproefomvang krijgt men de indruk, dat de populatieverdeling niet sterk van een normale verdeling zal afwijken. Deze indruk is juist, daar de steekproeven aselekt getrokken zijn uit de pseudo-normaal verdeelde modelpopulatie in tabel R.

21.2. In tabel 21.3 vindt men vijf frequentieverdelingen van de diameters van erythrocyten, elke verdeling met $n = 200$. Hier kan men o.i. reeds op het oog concluderen, dat de verdelingen pseudo-normaal zijn en zoals wij in 21.3 zullen zien leiden de toegepaste normaliteitstoetsen (waarop de in de tabel vermelde waarden van g_1 en g_2 betrekking hebben) eveneens tot deze uitspraak.

Tabel 21.3. Frequentieverdelingen van de diameters van 200 (ongewassen) erythrocyten bij 5 proeven

Diameter in μ	Proefnummer				
	1	2	3	4	5
9,00		1			
8,75	2	1			
8,50	0	3	1		
8,25	4	8	1	1	4
8,00	15	23	10	6	6
7,75	23	42	17	16	16
7,50	43	47	30	29	27
7,25	45	31	39	46	45
7,00	38	27	44	44	48
6,75	19	8	33	34	34
6,50	6	7	19	12	14
6,25	4	2	5	11	3
6,00	1		1	1	2
5,75					1
Totaal	200	200	200	200	200
g_1	0,1068	-0,0492	0,2043	-0,0081	0,0830
g_2	0,5177	0,3714	-0,2272	-0,2130	0,3719

Uit: M. A. MIJSBERG, De osmotische resistentie van gewassen erythrocyten, Diss., Amsterdam, 1953, Tabel XXII.

21.3. Bij 38 Alaska konijnen van 4 tot 5 maanden (afkomstig van het Centraal Proefdierenbedrijf T.N.O.) vonden wij voor enkele grootheden de volgende frequentieverdelingen:

Gewicht (gr)		Rest N ₂ (mg%)		Ureum (mg%)		Alk. fosfatase (bod. eenh.)	
x	f	x	f	x	f	x	f
850—999	1	30,0—32,9	1	14,0—15,9	1	1,0—1,9	1
1000—1149	1	33,0—35,9	1	16,0—17,9	—	2,0—2,9	3
1150—1299	4	36,0—38,9	3	18,0—19,9	2	3,0—3,9	7
1300—1449	8	39,0—41,9	4	20,0—21,9	4	4,0—4,9	13
1450—1599	11	42,0—44,9	6	22,0—23,9	5	5,0—5,9	6
1600—1749	8	45,0—47,9	10	24,0—25,9	9	6,0—6,9	3
1750—1899	4	48,0—50,9	5	26,0—27,9	7	7,0—7,9	4
1900—2049	—	51,0—53,9	4	28,0—29,9	5	8,0—8,9	1
2050—2199	1	54,0—56,9	1	30,0—31,9	3		
		57,0—59,9	2	32,0—33,9	1		
		60,0—62,9	1	34,0—35,9	1		
Totaal	38	Totaal	38	Totaal	38	Totaal	38
$\bar{x} = 1512,7$		$\bar{x} = 46,2$		$\bar{x} = 25,5$		$\bar{x} = 4,77$	
$s = 232,33$		$s = 6,56$		$s = 4,10$		$s = 1,60$	

De verdeling der alkalische fosfasaten is enigszins scheef naar rechts, maar de drie andere verdelingen zijn ongeveer symmetrisch en het interval $\bar{x} \pm 2,6 s$ bevat praktisch alle waarnemingen. Wij menen daarom dat voor deze homogene proefdieren reeds op grond van deze kleine steekproef kan worden gezegd, dat de populatieverdelingen van gewicht, rest-stikstof en ureum geen ernstige afwijkingen van de normaliteit zullen vertonen.

De situatie is echter een stuk ongunstiger, als men met een heterogene groep proefdieren werkt. Zo vonden wij voor een groep van 54 dieren van verschillende rassen en leeftijden de volgende tweetoppige en naar rechts scheve verdeling:

Gewicht (gr)	1450-1749	1750-2049	2050-2349	2350-2649	2650-2949	2950-3249	3250-3549	3550-3849	3850-4149	4150-4449
f	1	5	10	5	11	12	3	4	1	2

De conclusie, dat de populatieverdeling niet normaal is kan o.i. stellig worden getrokken, maar zij mist elke reële betekenis, omdat bij zo'n samenraapsel van dieren het begrip populatie reeds zinloos is.

Wij geloven, dat het voorgaande voorbeeld ook op menselijke populaties betrokken kan worden. Plakt men bv. op een te bestuderen populatie van patiënten een zo ruim etiket, dat er te veel verschillende en zelfs allerlei bijzondere gevallen toe behoren, dan kan men a priori zeggen dat sommige grootheden twee- of meertoppig verdeeld zullen zijn of uitbijters zullen vertonen. Het feit dat men dan geen klassieke methoden kan gebruiken is op zichzelf niet zo bezwaarlijk. De heterogeniteit kan echter ook de uitkomsten van het onderzoek beïnvloeden en vertroebelen, bv. doordat een therapie effect heeft bij een subgroep van 'speciale' patiënten en niet bij de overigen.

21.4. Normaliteitstoetsen

21.4.1. NORMAAL WAARSCHIJNLIJKHEIDSPAPIER

Bij een niet te klein aantal waarnemingen kan een grof onderzoek naar de normaliteit worden ingesteld met behulp van normaal waarschijnlijkheidspapier. Wij hebben het gebruik van dit papier niet beschreven, omdat dit onderzoek geen toetsingsmogelijkheid biedt en dus tot een subjectieve beslissing leidt. VEEN (31) heeft echter een duidelijk en met voorbeelden geïllustreerd artikel aan dit papier gewijd.

21.4.2. DE χ^2 -TOETS VOOR AANPASSING

In 10.1 (op blz. 216) is reeds aangegeven, op welke wijze men met de χ^2 -toets voor aanpassing de hypothese kan toetsen, dat de populatieverdeling normaal is. De toets kan worden gebruikt als men over een steekproef beschikt, waarvan de omvang voldoende groot is, zodat een niet te grove frequentieverdeling kan worden samengesteld. PEARSON en HARTLEY (74) schrijven echter over deze methode: 'While the procedure has certain advantages, it may prove somewhat insensitive to real departures from normality, partly because the χ^2 -test only takes into account the magnitude of the differences between observed and expected frequencies and not their sign and arrangement, and partly owing to the necessity of grouping together small tail frequencies when applying the test. If there are less than 50 — 100 observations, this loss of sensitivity through grouping will apply, indeed, to the whole range of the distribution.'

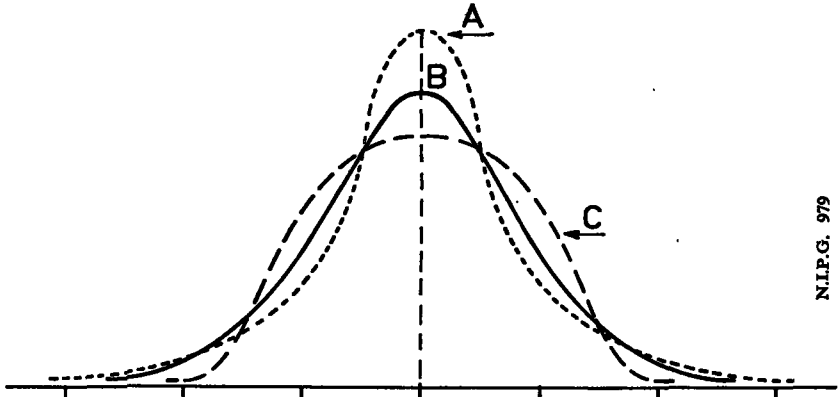
21.4.3. DE TOETS VAN KOLMOGOROV-SMIRNOV

Zoals wij reeds in 10.2 opmerkten, is het grote voordeel van deze toets boven de χ^2 -toets, dat geen frequentieverdeling behoeft te worden gebruikt. Hoewel men strikt genomen de *K-S* toets alleen kan gebruiken, als de parameters μ en σ bekend zijn, zal de toepassing met als schatting van deze parameters \bar{x} en s waarschijnlijk geen ernstige fout opleveren.

21.4.4. DE NORMALITEITSTOETSEN VAN FISHER

Er zijn verschillende toetsen ontworpen, die gericht zijn op het onderscheiden van bepaalde afwijkingen van de normaliteit, t.w. *asymmetrie* (scheefheid) en *kurtosis* ('spitsheid' of 'platheid', zie figuur 21.1). Wij bespreken eerst afzonderlijk de twee toetsen, die hiervoor door FISHER (44) zijn aangegeven en daarna de wijze waarop men deze toetsen kan combineren.¹

¹ Deze toetsen verschillen slechts weinig van de normaliteitstoetsen van PEARSON en zij verdienen o.i. de voorkeur. Zie voor een vergelijking bv. JOHNSON (22).



Figuur 21.1. Kurtosis: A = leptocurve (spitse kromme), B = normale kromme, C = platycurve (platte kromme).

Bij de FISHER-toetsen wordt gebruik gemaakt van de steekproef-functies

$$(21.1) \quad k_2 = \frac{nM_2 - M_1^2}{n(n-1)}$$

$$(21.2) \quad k_3 = \frac{n^2M_3 - 3nM_2M_1 + 2M_1^3}{n(n-1)(n-2)}$$

$$(21.3) \quad k_4 = \frac{(n^3 + n^2)M_4 - 4(n^2 + n)M_3M_1 - 3(n^2 - n)M_2^2 + 12nM_2M_1^2 - 6M_1^4}{n(n-1)(n-2)(n-3)}$$

Hierin is n de steekproefomvang en $M_k = \sum_{i=1}^n f_i x_i^k$. Het is duidelijk, dat $k_2 = s^2$ is en volledigheidshalve vermelden wij, dat $k_1 = M_1/n = \bar{x}$.

Toets tegen scheefheid. Deze toets berust op de grootheid

$$(21.4) \quad g_1 = \frac{k_3}{k_2 \sqrt{k_2}}.$$

Voor grote steekproeven uit een normaal verdeelde populatie is de grootheid g_1 bij benadering normaal verdeeld met gemiddelde $\mu_{g_1} = 0$ en variantie

$$(21.5) \quad \sigma_{g_1}^2 = \frac{6n(n-1)}{(n-2)(n+1)(n+3)},$$

zodat dan de grootheid

$$(21.6) \quad \underline{T} = \frac{g_1}{\sigma_{g_1}}$$

bij benadering standaardnormaal verdeeld is (de benadering is beter, naarmate n groter is). Bij scheefheid naar links neemt de grootheid g_1 een negatieve, bij scheefheid naar rechts een positieve waarde aan.

De hypothese, dat de waargenomen asymmetrie op toeval berust, kan dus via (21.6) worden getoetst. Bij *éénzijdige* toetsing kan men ook gebruik maken van tabel V, waarin voor een aantal waarden van n de rechter kritieke waarden van g_1 voor $\alpha = 0,05$ en $0,01$ zijn opgenomen (de linker kritieke waarden zijn numeriek gelijk aan de rechter, maar zij dragen een minteken).

Toets tegen kurtosis. Hierbij is de toetsingsgrootheid

$$(21.7) \quad g_2 = \frac{k_4}{k_2^2} = \frac{k_4}{s^4} .$$

Voor grote steekproeven uit een normaal verdeelde populatie is de grootheid g_2 bij benadering normaal verdeeld met $\mu_{g_2} = 0$ en

$$(21.8) \quad \sigma^2_{g_2} = \frac{24n(n-1)^2}{(n-2)(n-3)(n+3)(n+5)} ,$$

zodat dan de grootheid

$$(21.9) \quad T = \frac{g_2}{\sigma_{g_2}}$$

bij benadering standaardnormaal verdeeld is. Een positieve waarde van g_2 wijst op een te spitse en een negatieve op een te platte verdeling. Bij *éénzijdige* toetsing kan men gebruik maken van tabel V, waarin voor verschillende waarden van n de linker en rechter kritieke waarden van g_2 voor $\alpha = 0,05$ en $0,01$ gegeven zijn. Daar deze waarden niet symmetrisch liggen t.o.v. μ_{g_2} verdient waar mogelijk het gebruik van deze tabel de voorkeur boven de normale benadering, die hier minder bevredigend is dan bij de voorgaande toets.

Tabel 21.4. Toepassing van de toetsen tegen scheefheid en kurtosis op de waargenomen verdeling van de hemoglobinegehalten van 190 kinderen van 6 en 7 jaar (tabel 10.1)

Klassewaarde	f	u	fu	fu^2	fu^3	fu^4
16,00	2	+6	12	72	432	2592
15,50	3	+5	15	75	375	1875
15,00	11	+4	44	176	704	2816
14,50	19	+3	57	171	513	1539
14,00	26	+2	52	104	208	416
13,50	46	+1	46	46	46	46
13,00	38	0	0	0	0	0
12,50	24	-1	-24	24	-24	24
12,00	15	-2	-30	60	-120	240
11,50	4	-3	-12	36	-108	324
11,00	1	-4	-4	16	-64	256
10,50	1	-5	-5	25	-125	625
Totaal	190		151	805	1837	10753
	n		M_1	M_2	M_3	M_4

Voorbeelden

21.4. In tabel 21.4 zijn de berekeningen opgenomen die nodig zijn, als men beide normaliteitstoetsen toepast op de frequentieverdeling van de hemoglobinegehalten van kinderen in tabel 10.1 (dit is het voorbeeld op blz. 216, waarop wij ook reeds de χ^2 -toets voor aanpassing toepasten).

Berekeningen

$$k_2 = \frac{190(805) - (151)^2}{190(189)} = 3,624310$$

$$k_3 = \frac{(190)^2(1837) - 3(190)(805)(151) + 2(151)^3}{190(189)(188)} = 0,579945$$

teller van

$$k_4 = [(190)^3 + (190)^2]10753 - 4[(190)^2 + 190](1837)(151) + 3[(190)^2 - 190](805)^2 + 12(190)(805)(151)^2 - 6(151)^4 = 2795421924$$

$$k_4 = \frac{2795421924}{(190)(189)(188)(187)} = 2,214279$$

$$g_1 = \frac{0,579945}{3,624310 \sqrt{3,624310}} = 0,08405 \quad g_2 = \frac{2,214279}{(3,624310)^2} = 0,1686$$

$$\sigma^2 g_1 = \frac{6(190)(189)}{188(191)(193)} = 0,031090 \quad \sigma^2 g_2 = \frac{24(190)(189)^2}{187(188)(193)(195)} = 0,123111$$

$$\sigma_{g_1} = \sqrt{0,031090} = 0,1763 \quad \sigma_{g_2} = \sqrt{0,123111} = 0,3509$$

Normale benadering van de

toets tegen scheefheid:

$$T_1 = \frac{0,08405}{0,1763} = 0,477$$

toets tegen kurtosis:

$$T_2 = \frac{0,1686}{0,3509} = 0,480$$

Teneinde de (vrij omvangrijke) berekeningen zoveel mogelijk te vereenvoudigen, is gebruik gemaakt van een hulpgrootheid (u). Een vergelijking van de berekende waarden van g_1 en g_2 met de kritieke waarden in tabel V (die bij tweezijdige toetsing, zoals in dit geval, voor $\alpha = 0,10$ en $0,02$ gelden), toont onmiddellijk aan dat van significantie geen sprake is. De waarden van T bij de normale benadering wijzen bij beide toetsen op een tweezijdige overschrijdingskans in de nabijheid van 0,60.

21.5. Ook voor de vijf verdelingen in tabel 21.3 zijn de waarden van g_1 en g_2 berekend. Wij laten het aan de lezer over om door middel van tabel V vast te stellen, dat geen dezer waarden significant is.

21.4.5. COMBINATIE VAN DE TOETSEN VAN FISHER

In de praktijk zal de hypothese H_0 , die men wenst te toetsen gewoonlijk luiden, dat de populatieverdeling normaal is, met als alternatieve hypothese, dat afwijkingen van de normaliteit optreden (zonder dat men de aard van deze afwijkingen a priori kan aangeven). Er dient dan tweezijdig te worden getoetst en past men nu de toets tegen scheefheid en de toets tegen kurtosis toe, beide met een onbetrouwbaarheidsdrempel α , dan is de kans dat men H_0 ten onrechte verwerpt groter dan α , nl. $2\alpha(1 - \alpha) + \alpha^2 = 2\alpha - \alpha^2$ (= de kans, dat men ten minste bij één van de twee toegepaste toetsen ten onrechte tot verwerpen komt). Kiest men dus bv. bij de afzonderlijke toetsen $\alpha = 0,05$, dan is de kans dat men H_0 ten onrechte verwerpt $2(0,05) - (0,05)^2 = 0,0975$.

MADOW (176) heeft twee manieren aangegeven, waardoor dit bezwaar kan worden opgeheven:

Toets I: Wil men de hypothese H_0 dat de populatieverdeling normaal is tweezijdig toetsen met drempelwaarde α , dan dient men de toetsen tegen scheefheid en tegen kurtosis uit te voeren met als onbetrouwbaarheidsdrempel

$$(21.10) \quad \alpha' = 1 - \sqrt{1 - \alpha} .$$

Kiest men bv. $\alpha = 0,05$, dan is dus $\alpha' = 1 - \sqrt{1 - 0,05} = 0,0253$. Men besluit dan tot het verwerpen van H_0 , als de toets tegen scheefheid, de toets tegen kurtosis, of beide toetsen significant zijn bij een drempelwaarde 0,0253, d.i. dus als

$$|T_1| \geq T_{1-4\alpha'} \quad \text{of/en} \quad |T_2| \geq T_{1-4\alpha'}$$

d.i. voor

$$|T_1| \geq T_{0,9874} \quad \text{of/en} \quad |T_2| \geq T_{0,9874} .$$

Uit tabel A blijkt, dat $T_{0,9874} = 2,24$.

Toets II: Deze toets berust op het feit, dat de grootheid

$$(21.11) \quad \chi^2 = \left(\frac{g_1}{s_{g_1}}\right)^2 + \left(\frac{g_2}{s_{g_2}}\right)^2 = \underline{T}_1^2 + \underline{T}_2^2 .$$

onder H_0 (bij benadering) een χ^2 -verdeling met 2 vrijheidsgraden volgt. De normaliteitshypothese wordt dan bij een drempelwaarde α verworpen, als $\chi^2_0 \geq \chi^2_{1-\alpha}$, dus bv. bij een 5% drempel als $\chi^2_0 \geq 5,991$.

Het is duidelijk, dat geen der gecombineerde toetsen bij toepassing op het voorbeeld in tabel 21.4 tot het verwerpen van de normaliteits-hypothese leidt.

21.4.6. DE TOETS TEGEN KURTOSIS VAN GEARY

Door GEARY (159) is een toets tegen kurtosis aangegeven, die ook bij kleine steekproefomvang kan worden gebruikt. De toetsingsgrootheid is

$$(21.12) \quad \underline{a} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{\sqrt{n \sum_{i=1}^n (x_i - \bar{x})^2}}$$

Bij de berekeningen kan men nemen:

$$\sum_{i=1}^n |x_i - \bar{x}| = 2 \left[(\text{som van de waarden groter dan } \bar{x}) - \bar{x} (\text{aantal waarden groter dan } \bar{x}) \right].$$

In de BIOMETRIKA TABLES (74, tabel 34A) staan voor verschillende waarden van n (vanaf 11) de linker en rechter kritieke waarden van \underline{a} bij éézijdige toetsing met $\alpha = 0,10, 0,05$ en $0,01$.

21.4.7. SCHEEFHEIDSTOETS GEBASEERD OP KWARTIELEN

Door VAN DE PANNE en LANGE (227) is de op kwartielen gebaseerde scheefheidscoëfficiënt

$$(21.13) \quad \underline{S} = \frac{x_{0,75} + x_{0,25} - x_{0,50}}{x_{0,75} - x_{0,25}}$$

voorgesteld. Bij aselechte trekking van n elementen uit een normaal verdeelde populatie is de grootheid \underline{S} bij benadering normaal verdeeld met gemiddelde 0, terwijl de variantie bij benadering $\sigma_S^2 = \frac{1,839}{n}$ is.

Zodoende is de grootheid

$$(21.14) \quad \underline{T} = \frac{\underline{S}}{\sigma_S} = \frac{\underline{S}n}{1,839}$$

bij benadering standaardnormaal verdeeld. Zij merken omtrent deze scheefheidscoëfficiënt het volgende op: 'Aangezien S bepaald wordt door de verhouding van de afstanden van de mediaan tot beide kwartielen, kunnen we stellen dat de door S gemeten scheefheid betrekking heeft op de ligging van de middelste helft van de waarnemingen. Uitschieters beïnvloeden S slechts door de rangschikking (d.w.z. door het feit dat zij links van het benedenkwartiel of rechts van het bovenkwartiel liggen), maar niet numeriek; men vermijdt dus de vraag of ze al of niet buiten beschouwing gelaten moeten worden.'

21.5. Transformaties

In de voorgaande hoofdstukken zijn reeds enkele transformaties ter

sprake gekomen. Het lijkt ons nuttig deze met enkele andere in één paragraaf samen te vatten.

21.5.1. DE LOGARITHMISCHE TRANSFORMATIE: $y = \log x$ OF $y = \log(x+1)$

Zoals wij reeds hebben gezien leidt de logaritmische transformatie bij vele scheef naar rechts verdeelde grootheden tot een ongeveer normale verdeling (zie ook 5.6.3 en 23.2). GADDUM (222) geeft verschillende voorbeelden van dergelijke lognormale verdelingen.

Wanneer men te doen heeft met verdelingen, waarvan de gemiddelden evenredig zijn met de spreidingen (zodat de zg. *variatiëcoëfficiënt* $v = s/\bar{x}$ ongeveer constant is), kan men met de logaritmische transformatie bereiken, dat de varianties homogeen worden. Als neveneffect vindt men dan gewoonlijk, dat de verdelingen de normale verdeling beter gaan benaderen.

21.5.2. DE VIERKANTSWORTEL TRANSFORMATIE: $y = \sqrt{x}$ OF

$$y = \sqrt{x} + \sqrt{x+1}$$

Deze transformatie kan worden gebruikt als men met POISSON-verdelingen te doen heeft, zodat $s^2 \cong \bar{x}$. Zie voor een voorbeeld tabel 16.2 op blz. 419. Om de uitvoering van de laatstgenoemde vorm van deze transformatie te vergemakkelijken hebben wij in tabel BB voor $0 \leq x \leq 50$ de waarden van $\sqrt{x} + \sqrt{x+1}$ en $(\sqrt{x} + \sqrt{x+1})^2$ opgenomen.

21.5.3. DE RECIPROKE TRANSFORMATIE: $y = 1/x$

Deze transformatie leidt tot homogeniteit van de varianties en tot een meer normale vorm van de verdelingen, wanneer $s \cong \bar{x}^2$. Volgens de literatuur is zij o.m. van toepassing op verdelingen van de tijden die nodig zijn om een bepaalde handeling te verrichten en EDWARDS (55) geeft enige voorbeelden op psychologisch gebied.

21.5.4. DE BGSIN TRANSFORMATIE: $y = 2 \text{ BGSIN } \sqrt{\bar{p}}$

Voor aselechte steekproeven van n elementen uit een dichotome populatie met $P(A) = P$ volgt de grootheid $\hat{p}(A) = \hat{p} = x/n$ een binomiale verdeling met parameters $\mu_{\hat{p}} = P$ en $\sigma_{\hat{p}}^2 = P(1-P)/n$ (zie 4.3). De variantie van \hat{p} is dus niet alleen afhankelijk van de steekproefomvang n , maar tevens van de waarde van de parameter P . Zo geldt bv. voor $n = 100$ en $P = 0,1$ dat $\sigma_{\hat{p}}^2 = 0,0009$, maar voor $n = 100$ en $P = 0,5$ geldt dat $\sigma_{\hat{p}}^2 = 0,0025$. Verder is de verdeling van \hat{p} alleen symmetrisch voor $P = 0,5$ en vertoont zij voor $P \neq 0,5$ een scheefheid, die sterker is naarmate P verder van 0,5 verwijderd is.

De verdeling van de grootheid $\underline{y} = 2 \text{ bgsin } \sqrt{\bar{p}}$ volgt echter voor $0,05 < P < 0,95$ en $n \geq 20$ een verdeling, die niet alleen minder

asymmetrisch is dan de verdeling van \hat{p} , maar waarvan de variantie vrijwel onafhankelijk is van P en bij benadering gelijk gesteld kan worden aan¹

$$(21.15) \quad \sigma_y^2 = 1/n.$$

Deze benadering wordt bij grotere steekproefomvang beter en zij geldt dan ook voor waarden van P kleiner dan 0,05 of groter dan 0,95. Tabel AA-1 geeft de waarden van y voor $\hat{p} = 0,001$ (0,001) 0,500 en wegens de symmetrie van de y -verdeling kan men voor $0,501 \leq \hat{p} \leq 0,999$ gebruik maken van de betrekking $y = 3,1416 - 2 \text{ bgsin } \sqrt{1 - \hat{p}}$. Voor de waarde $y(0)$ die correspondeert met $\hat{p} = 0$ kan men volgens BARTLETT (213) het best nemen

$$(21.16) \quad y(0) = 2 \text{ bgsin } \sqrt{\frac{1}{4n}}$$

en voor de waarde $y(1)$ die correspondeert met $\hat{p} = 1$

$$(21.17) \quad y(1) = 3,1416 - 2 \text{ bgsin } \sqrt{\frac{1}{4n}}.$$

Voor $n = 10$ (1) 50 kan men $y(0)$ en $y(1)$ direct uit tabel AA-2 aflezen.

In 23.4.4 komt de toepassing van de bgsin transformatie bij de biologische ijking ter sprake. Verder kan de transformatie worden gebruikt als men beschikt over twee of meer waargenomen binomiale verdelingen (d.w.z. de verdelingen van fracties (percentages), die alle berusten op hetzelfde aantal waarnemingen en waarbij alle tot één verdeling behorende fracties schattingen zijn van dezelfde parameter P). Bij verschillende waarden van P zullen niet alleen de gemiddelden \bar{p} maar ook de varianties s_p^2 van deze verdelingen verschillen en deze heterogeniteit wordt dan door toepassing van de bgsin transformatie opgeheven, terwijl ook de verdelingen beter normaal worden.

Uit het voorgaande volgt, dat de grootheid

$$(21.18) \quad T = \frac{2 \text{ bgsin } \sqrt{\hat{p}} - 2 \text{ bgsin } \sqrt{P}}{\sqrt{1/n}} = (y_p - y_P) \sqrt{n}$$

voor $0,05 < P < 0,95$ en $n \geq 20$ reeds bij benadering standaardnormaal verdeeld is. Men kan dus op basis van een waargenomen fractie $\hat{p}_0 = x_0/n$ de hypothese $H_0: P = P_0$ ook via (21.18) toetsen en deze methode kan reeds gebruikt worden, als de normale benadering in 6.3 (blz. 142—143) nog niet opgaat. ANSCOMBE (212) heeft aangegeven, dat men bij gebruik van (21.18) het best kan nemen

$$(21.19) \quad \hat{p}_0 = \frac{x_0 + \frac{8}{6}}{n + \frac{3}{4}}.$$

¹ Zie EISENHART (8a), waarin deze variantie voor $n = 10, 20$ en 30 en $P = 0,05$ (0,05) 0,95 is opgenomen in vergelijking met de variantie van \hat{p} .

Voorbeeld

21.6. Men toetst de hypothese $P_0 = 0,2$ met als tegenhypothese $P_1 \neq 0,2$. Een aselechte steekproef van 20 elementen levert op $x_0 = 8$. Hoe groot is bij benadering de tweezijdige overschrijdingskans van deze uitkomst?

Nu is $p_0 = (8 + 0,375)/(20 + 0,75) = 0,404$ en uit tabel AA-1 blijkt dat $y_p = 1,3776$ en $y_P = y(P = 0,2) = 0,9273$, zodat volgens (21.18) $T_0 = (1,3776 - 0,9273) \sqrt{20} = 2,014$. Tabel A geeft als tweezijdige overschrijdingskans van deze uitkomst ongeveer 0,0440 en uit voorbeeld 6.9 (blz. 143) blijkt, dat de exacte tweezijdige kans 0,0437 bedraagt. Ondanks de kleine waarde van n en de sterk van 0,5 afwijkende waarde van P is de benadering dus voortreffelijk.

21.5.5. TRANSFORMATIE VAN χ^2 -, F - EN t -VERDELINGEN

In 13.3.1 hebben wij reeds de transformatie van WILSON en HILFERTY voor χ^2 -verdelingen behandeld en wij zullen deze in verband met de plaatsruimte hier niet herhalen. PAULSON (228) heeft echter naar aanleiding van deze transformatie gezocht naar een transformatie voor de grootheid \underline{F} . Deze grootheid is immers slechts voor een beperkt aantal fractielen en voor een aantal waarden van ν_1 en ν_2 getabelleerd. Hij vond dat voor $\nu_2 \geq 3$ de grootheid

$$(21.20) \quad \underline{T} = \frac{\left(1 - \frac{2}{9\nu_2}\right) \underline{F}^{\frac{1}{3}} - \left(1 - \frac{2}{9\nu_1}\right)}{\sqrt{\frac{2}{9\nu_2} \underline{F}^{\frac{2}{3}} + \frac{2}{9\nu_1}}}$$

bij benadering standaardnormaal verdeeld is. Men kan dus de rechtse kans van $F_0(\nu_1, \nu_2)$ bij benadering vinden, door met (21.20) T_0 te berekenen en de rechtse kans van deze waarde in tabel A op te zoeken. Wil men echter een linkse of een tweezijdige kans bepalen, dan dient ν_1 eveneens ≥ 3 te zijn. PAULSON geeft een aantal voorbeelden waaruit blijkt dat de verkregen benadering reeds bij betrekkelijk kleine aantallen vrijheidsgraden zeer nauwkeurig is.

De benadering kan ook worden gebruikt voor STUDENT-verdelingen. Men dient dan in (21.20) in plaats van F_0 te nemen t_0^2 , terwijl $\nu_1 = 1$ is.

Voorbeeld

12.7. Wij gaan uit van $F_0(\nu_1 = 8, \nu_2 = 15) = 4,00$, waarvan volgens tabel F-3 de rechtse kans gelijk is aan 0,01. Voor deze waarde van \underline{F} geldt, dat $\underline{F}^{\frac{1}{3}} = \sqrt[3]{F_0} = \sqrt[3]{4} = 1,5874$ en $\underline{F}^{\frac{2}{3}} = \sqrt[3]{F_0^2} = \sqrt[3]{16} = 2,5198$, zodat volgens (21.20)

$$T_0 = \frac{(1 - 2/9 \times 15)(1,5874) - (1 - 2/9 \times 8)}{\sqrt{2 \times 2,5198/9 \times 15 + 2/9 \times 8}} = \frac{0,5917}{0,2552} = 2,319.$$

Uit tabel A blijkt dat $P_R(2,31) = 0,0099$ en $P_R(2,32) = 0,0102$. De benadering levert dus een bijzonder nauwkeurige uitkomst op.

STATISTIEK EN PROEFOPZET

22.1. Inleiding

In hoofdstuk 1 hebben wij reeds een korte beschouwing gewijd aan de relatie tussen proefopzet en statistiek. Hierin is tot uitdrukking gekomen dat de proefopzet aan een aantal voorwaarden dient te voldoen, wanneer men de statistiek als bewijsmiddel wil gebruiken, d.w.z. als men wil, dat aan de uitkomsten van de statistische analyse een ondubbelzinnig vaststaande betrouwbaarheid kan worden toegekend. De voornaamste van deze voorwaarden kunnen als volgt worden samengevat:

1. De vraagstellingen van het onderzoek dienen tevoren duidelijk te worden geformuleerd.

2. De populatie(s), waarop het onderzoek betrekking heeft, moeten scherp worden gedefinieerd.

3. De steekproeven uit deze populaties dienen op aselecte wijze te worden gekozen. Bij een experiment moet men de verschillende proefgroepen aselect samenstellen.¹

4. Het waarnemingsproces dient nauwkeurig omschreven te worden. Hierbij zijn steeds de volgende vragen van belang: *wat* gaat men waarnemen, *hoe* gaat men dit doen (en hoe betrouwbaar zijn de te gebruiken waarnemingsmethoden), *wanneer* gaat men de waarnemingen verrichten, en *wie* verrichten ze.

5. De proefopzet dient gericht te zijn op een bepaalde methode van statistische analyse. Men dient tevoren te overwegen, welke onbetrouwbaarheid van de te trekken conclusies toelaatbaar is.

De uit het onderzoek volgende conclusies gelden in principe uitsluitend voor de populatie(s), waarop dit onderzoek betrekking heeft. Wanneer men bv. door middel van een goed opgezet onderzoek heeft vastgesteld, dat een geneesmiddel een therapeutisch effect bezit bij Nederlandse mannen van 20 t/m 50 jaar, dan mag men strikt genomen

¹ In 1.3 hebben wij aangegeven, op welke wijze men hiertoe gebruik kan maken van lotings-tabellen (tabellen met aselecte permutaties), maar deze zijn in de tabellenverzameling van deel I niet gegeven. Naar aanleiding van enkele desbetreffende verzoeken hebben wij in dit deel in tabel CC alsnog enkele lotingstabellen opgenomen.

deze conclusie niet op bv. Engelse mannen van deze leeftijden, op Nederlandse jongens, of op Nederlandse vrouwen van 20 t/m 50 jaar betrekken. Er kunnen echter *medische* overwegingen zijn, die de geldigheid van de conclusie ook voor deze populaties plausibel maken, maar een (statistisch) bewijs van deze geldigheid is slechts te leveren als men het onderzoek ook voor deze populaties uitvoert. Het is duidelijk, dat deze materie bij detectieproblemen moeilijker ligt, omdat daarbij de bestudeerde populaties gewoonlijk a priori niet scherp, zijn gedefinieerd, zodat men achteraf moet trachten vast te stellen hoe zij zijn samengesteld. Hieraan wordt zelden voldoende aandacht besteed.

Een extrapolatie in de tijd van de onderzoeksgegevens is mogelijk, indien voldaan is aan de voorwaarde dat de bestudeerde populaties in de tijd geen veranderingen hebben ondergaan. Op het medisch-biologische toepassingsgebied van de statistiek zullen de meeste populaties na verloop van (wat langere) tijd wél veranderen, zodat de conclusies van een nu verricht onderzoek over 10 jaar wellicht niet meer gelden. Zo kan bv. een insecticide nu (en over een jaar) bij een bepaalde insectensoort een uitgesproken effect hebben, maar na verloop van langere tijd kan zich bij deze soort een immuniteit voor deze stof ontwikkelen, waardoor dit effect vrijwel geheel verloren gaat. Ook in dit geval geldt dat men soms op grond van niet-statistische overwegingen aannemelijk kan maken, dat zich in een populatie 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 door anderen verrichte onderzoekingen vaak verre van zinloos.

Bij de toetsing van een hypothese verkrijgt men uiteindelijk een (één- of tweezijdige) overschrijdingskans, die wordt geconfronteerd met de a priori gekozen onbetrouwbaarheidsdrempel α . In principe besluit men tot het verwerpen van de getoetste hypothese, indien deze overschrijdingskans kleiner dan of gelijk is aan α ; is dit niet het geval, dan wordt deze hypothese niet verworpen. In de praktijk dient men echter bij het toepassen van deze regel op de resultaten van een wetenschappelijk onderzoek met de volgende factoren rekening te houden:

a. De berekende overschrijdingskans geldt vaak bij benadering. Bij het gebruik van een klassieke methode heeft men altijd te maken met afwijkingen van het model, zodat de overschrijdingskans niet exact geldt. Maar ook bij het gebruik van een verdelingsvrije methode verkrijgt men gewoonlijk een benaderende kans, hetzij omdat men bij een groot aantal waarnemingen de exacte kansverdeling van de toetsingsgrootheid niet meer kan berekenen, hetzij omdat men bij het optreden van groepen van gelijke waarnemingen van een benadering gebruik

moet maken [het berekenen van de exacte verdeling is dan soms wel mogelijk, maar zeer tijdrovend].

b. De gekozen onbetrouwbaarheidsdrempel is gewoonlijk meer bedoeld als *richtlijn* dan als onherroepelijk vaststaande grens. De keuze van bv. $\alpha = 0,05$ houdt dan in, dat men tot het verwerpen van de getoetste hypothese wil overgaan als de overschrijdingskans ongeveer 0,05 (of kleiner) is. Er is dan geen wezenlijk verschil tussen een waargenomen overschrijdingskans van (bij benadering) 0,04 en 0,06.

Hierbij komt nog, dat een statistisch significante uitkomst nog geen praktische betekenis behoeft te bezitten. Stel bv. dat een proef uitwijst, dat een bepaalde therapie, toegepast op personen met een te laag hemoglobinegehalte (bv. met een gemiddelde van 10 g%) een gemiddelde verhoging van 0,4 g% oplevert. Zelfs wanneer deze uitkomst significant is bij een 1% drempel (bv. een overschrijdingskans van 0,007 bezit), kan men niet zeggen dat deze therapie medisch gezien van waarde is. Maar het omgekeerde is eveneens mogelijk: een niet significante uitkomst (met bv. een overschrijdingskans van 0,09) kan in een bepaalde situatie voor een onderzoeker op grond van zijn kennis en ervaring een wezenlijke betekenis hebben.

Uit de voorgaande overwegingen volgt, dat men in de praktijk de uitkomsten van een statistische analyse met enige souplesse zal moeten hanteren. Om met HAMAKER (163) te spreken: 'Men moet . . . bij de beoordeling van experimentele gegevens 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.'

Wij willen hier ook de aandacht vestigen op enkele niet-statistische publicaties betreffende het wetenschappelijk onderzoek op medisch-biologisch gebied, die zeer lezenswaard zijn. Dit zijn: 'The art of scientific investigation' van BEVERIDGE en 'Principles of research in biology and medicine' van INGLE, twee boeken die vlot geschreven zijn en naast zeer behartigenswaardige algemene opmerkingen, wenken en waarschuwingen, vele aardige voorbeelden bevatten. Veel uitgebreider is 'An introduction to scientific research' van WILSON, waarin ook statistische en meettechnische achtergronden worden belicht. Tenslotte noemen wij een uitgave van de CIOMS¹, waarin een reeks inleidende lezingen van de in maart 1959 te Wenen gehouden conferentie betreffende 'controlled clinical trials' zijn samengevat. Men vindt hierin o.m. bijdragen van A. BRADFORD HILL, L. J. WITTS (The ethics of c.c.t.), C. M. FLETCHER (Criteria for diagnosis and assessment), D. D. REID (The patient as his own control), R. DOLL (The concurrent assessment of several treatments) en P. ARMITAGE (The sequential

¹ Council for International Organizations of Medical Sciences, 2 Place de Fontenoy, Paris, 7 ème.

approach), een groot aantal voorbeelden en een uitvoerige literatuur-opgave.

Deze inleiding biedt ons ook de gelegenheid te wijzen op een aspect van het waarnemingsproces, dat in 1.4 onvoldoende aandacht heeft gekregen, nl. het gevaar dat de waarnemingen worden vertroebeld door een min of meer systematische fluctuatie in de tijd van het waargenomen verschijnsel. Wij maken daarbij gebruik van een voorbeeld betreffende een proef, waarbij men het kaliumgehalte in het parotisspeeksel bij vijf proefpersonen op drie achtereenvolgende dagen bepaalt. Onderstel dat de bedoeling van deze proef is na te gaan, of er systematische verschillen tussen de dagen optreden. Men verricht elke dag één waarneming per persoon en verkrijgt daarbij de in tabel 22.1 vermelde uitkomsten.

Tabel 22.1. Kaliumgehalte (in mg%) in het parotisspeeksel van vijf personen op drie opeenvolgende dagen

Datum	Proefpersoon				
	I	II	III	IV	V
3/4	86	99	99	91	91
4/4	102	96	105	93	99
5/4	113	102	110	95	107

Past men de toets van FRIEDMAN toe dan vindt men $K_0 = 42$, met een overschrijdingskans van 0,0085. De conclusie luidt dus, dat de kaliumgehalten op de drie onderzoekdagen systematisch verschillen, terwijl de in 12.2 beschreven toets tegen verloop op een stijgende trend van de gehalten wijst.

Nu vormen echter de waarnemingen in tabel 22.1 een deel uit een groter onderzoek, waarvan de uitkomsten in tabel 22.2 staan. Zij zijn hieruit aselekt gekozen, maar hierbij is aangenomen dat de proefpersonen op 3 en 4 april om 10 uur òf om 12 uur, en op 5 april om 14 uur òf om 16 uur voor het onderzoek ter

Tabel 22.2. Kaliumgehalte (in mg/100 cc) van het parotisspeeksel bij vijf proefpersonen

Datum	Uur	Proefpersoon				
		I	II	III	IV	V
3/4	10	86	93	99	91	84
	12	98	99	108	105	91
	14	110	107	114	104	98
	16	110	100	118	118	89
4/4	10	103	96	92	93	99
	12	102	95	105	82	106
	14	121	98	123	112	98
	16	108	109	106	107	91
5/4	10	108	92	102	87	104
	12	110	97	104	92	101
	14	113	102	102	107	105
	16	106	103	110	95	107

beschikking stonden. Bestudeert men de volledige gegevens (bv. door de toets van FRIEDMAN op de sommen van de vier waarnemingen per dag/per persoon toe te passen), dan blijkt er geen sprake te zijn van een systematisch verschil tussen de dagen. Wél ziet men, dat binnen elke dag een systematisch verschil tussen de tijdstippen van waarneming optreedt: de ochtendgehalten zijn in het algemeen lager dan die in de middag en dit verschil is voornamelijk verantwoordelijk voor de verschillen tussen de dagen bij de gegevens in tabel 22.1, die deels 's ochtends en deels 's middags verzameld zijn.

Dit voorbeeld leert dat men een aantal waarnemingen, die op verschillende tijdstippen zijn verricht (bv. op verschillende uren van het etmaal of in verschillende maanden van het jaar) slechts als homogeen mag beschouwen, wanneer geen systematisch verloop in de tijd (bv. een dagritme of een seizoensverloop) optreedt. Vaak zal men een afzonderlijk voorbereidend onderzoek moeten uitvoeren om na te gaan, of aan deze voorwaarde is voldaan. Ook hieraan wordt zelden voldoende aandacht besteed.

In de voorgaande hoofdstukken hebben wij een aantal betrekkelijk eenvoudige proefopzetten en de daarbij behorende methoden van statistische analyse leren kennen. Dit hoofdstuk handelt over een categorie van meer gecompliceerde proefopzetten, de z.g. statistische proefschemas en over enkele speciale steekproeftechnieken. Deze onderwerpen worden niet volledig behandeld. Het is slechts onze bedoeling een indruk te geven van de mogelijkheden die er zijn en de weg te wijzen naar de uitgebreide literatuur, die hierover ter beschikking staat.

22.2. Statistische proefschemas bij experimenten

De techniek der variantie-analyse verschaft de mogelijkheid tot het analyseren van de uitkomsten van meer gecompliceerde maar bijzonder efficiënte proefopzetten, waarbij een volledige symbiose van proefopzet en statistiek optreedt. De situatie, waarin men zich eerst nadat de proef was verricht ging afvragen op welke wijze het waarnemingsmateriaal kon worden geanalyseerd is daarbij geheel in het tegendeel overgegaan, daar het experiment wordt gebaseerd op een 'statistisch' proefschemas.

De stoot tot deze ontwikkeling werd gegeven door R. A. FISHER, die in 1925 in zijn 'Statistical methods for research workers' o.m. de variantie-analyse introduceerde en een aantal ideeën betreffende de opzet van experimenten naar voren bracht, die hij verder uitwerkte in zijn boek 'The design of experiments', dat in 1935 verscheen. FISHER studeerde astronomie, maar hij was na de beëindiging van zijn studie sterk geïnteresseerd in de biologie en de genetica en mede daardoor in de statistica. Van 1919 tot 1933 was hij als statisticus verbonden aan het landbouwkundig proefstation te Rothamsted (Engeland), en de eerste toepassingen van zijn methoden betreffen dan ook voornamelijk experimenten op het gebied van de agronomie. In 1933 werd hij benoemd tot hoogleraar in de eugenetica te London en in 1943 tot hoogleraar in de genetica te Cambridge. Sindsdien heeft hij zich nog maar betrekkelijk weinig met zuiver statistische problemen bezig gehouden. Men kan van de betekenis die het werk

van FISHER voor de statistiek heeft gehad een goede indruk krijgen, als men de artikelen van YATES (210), HOTELLING (170), MATHER (177) en YOUNG (211) leest, die in 1950 zijn geschreven ter gelegenheid van het feit dat 'Statistical methods' 25 jaar tevoren verscheen.

22.2.1. KLASSIEKE EN FACTORIËLE PROEFOPZETTEN

Onder een *factor* verstaat men een voorwaarde die bij een experiment kan worden gevarieerd. Men onderscheidt kwalitatieve factoren (mensen, behandelingswijzen, bereidingswijzen, e.d.) en kwantitatieve factoren (tijden, doseringen, concentraties, enz.). Wanneer men bij een experiment een factor op een aantal manieren laat variëren spreekt men van de verschillende *niveau's* (Eng. 'levels') van die factor. Zo bezit de factor 'geslacht' twee niveau's en de factor 'tijdstip' in tabel 22.2 vier niveau's (10, 12, 14 en 16 uur). De bij een experiment gemeten waarden van een variabele noemt men *responsies* (ook wel: scores). In het algemeen heeft een experiment ten doel te onderzoeken welke veranderingen er in de responsies optreden, als men één of meer factoren op twee of meer niveau's per factor varieert.

Volgens deze terminologie kan men de proefopzet bij de enkelvoudige variantie-analyse in 17.1 (aselecte steekproeven uit k populaties) beschouwen als een experiment met één kwalitatieve factor (de factor 'populatie') en k niveau's. De proefopzet bij de tweevoudige analyse in 17.3 is dan te zien als een experiment met twee kwalitatieve factoren, nl. de factor 'rekenaar' met 6 en de factor 'rekenmachine' met 4 niveau's.

Bij de *klassieke methode van experimenteren*, die voortkomt uit de fysica, wordt steeds één factor tegelijk gevarieerd, waarbij dan alle andere bij het experiment betrokken factoren constant worden gehouden. Bij de door FISHER aangegeven *factoriële proefschema's* experimenteert men in principe bij alle mogelijke combinaties van de verschillende niveau's van deze factoren. Hierbij voert men zo nodig m replicaties per combinatie uit. Men spreekt van een 2^p factorieel schema wanneer er p factoren zijn met 2 niveau's per factor. In 22.2.2 geven wij een voorbeeld van een 2^3 factorieel schema zonder replicaties, terwijl 22.2.3 zo'n schema met replicaties bevat. Tabel 22.2 laat een $3 \times 4 \times 5$ factorieel schema zien, dat in 22.2.4 nader wordt besproken. Schema's met meer dan drie factoren en met drie of meer niveau's per factor worden hier niet gegeven, maar zijn bv. in de landbouw geen zeldzaamheid (zie ook 22.2.6). Daar de meeste informatie wordt verkregen als men met replicaties werkt, vereisen dergelijke schema's reeds spoedig een vrij groot aantal waarnemingen. Er zijn daarom ook zg. *fractionele factoriële* schema's ontworpen, waarbij men een aantal combinaties van factoren kan laten vervallen (en dus het benodigde aantal waarnemingen kan reduceren), zonder dat daardoor essentiële

informatie verloren behoeft te gaan. Tenslotte behandelen wij in 22.2.5 nog enkele bijzondere proefschemas, het Latijnse en het Grieks-Latijnse vierkant, die alleen of in combinatie met andere factoriële schemas kunnen worden toegepast.

De factoriële proefopzetten bezitten twee belangrijke voordelen boven de klassieke, t.w. (1) zij verschaffen bij een gegeven aantal waarnemingen meer informatie, en (2) zij bieden de mogelijkheid om te onderzoeken, of er tussen de verschillende factoren interactie optreedt. Wij komen op deze voordelen bij de behandeling van de voorbeelden terug.¹

22.2.2. EEN 2³ FACTORIEEL PROEFSHEMA ZONDER REPLICATIES

Wij beschouwen een onderzoek naar de opbrengst van een nitreringsproces, nl. het nitreren van aniline, waarvan het eindresultaat het grondmateriaal vormt van een uitgebreid gebied van kleurstoffen en medicamenten². Bij dit onderzoek worden 3 factoren gevarieerd op 2 niveau's, t.w.:

- A. De duur van het toevoegen van salpeterzuur: 2 uur of 7 uur.
- B. De roertijd: ½ uur of 4 uur.
- C. Het restant: afwezig of aanwezig.

Het 'restant' wil zeggen een rest van de voorgaande charge, die is achtergebleven in de pan waarin het procédé plaatsvindt. Bij 'restant afwezig' is de pan dus zorgvuldig gereinigd. Door deze factor in te voeren wil men nagaan, in hoeverre een grondige schoonmaak van de pannen nodig is voor het verkrijgen van een goede opbrengst.

Onderstel nu eerst dat voor elk der 8 mogelijke combinaties één proef wordt uitgevoerd en dat daarbij de uitkomsten verkregen worden, die in tabel 22.3 zijn opgenomen. Het effect van iedere factor, d.i. het verschil tussen de opbrengsten bij de twee niveau's van de factor, kan hieruit worden geschat. Zo bedraagt b.v. de gemiddelde opbrengst bij een roertijd van ½ uur:

$$\frac{87,2 + 88,4 + 86,7 + 89,2}{4} = 87,875 \%$$

en bij een roertijd van 4 uur:

$$\frac{82,0 + 83,0 + 83,4 + 83,7}{4} = 83,025 \%$$

zodat het waargenomen effect van de roertijd $87,875 - 83,025 = 4,85 \%$ is. Dit effect is berekend uit alle 8 waarnemingen en het is

¹ Zie ook FISHER (56), par. 37.

² Het voorbeeld is ontleend aan DAVIES c.s. (142). Dit artikel bevat ook enige voorbeelden van eenvoudige fractionele factoriële schemas.

Tabel 22.3. Procentuele opbrengst bij de nitrering van aniline

		C. Restant			
		Afwezig		Aanwezig	
		A. Toevoeging van salpeterzuur			
		2 uur	7 uur	2 uur	7 uur
B. Roertijd	½ uur	87,2	88,4	86,7	89,2
	4 uur	82,0	83,0	83,4	83,7

onafhankelijk van de (eventuele) effecten van de beide andere factoren. Bij beide niveau's van de roertijd zijn er immers twee waarnemingen met en twee zonder restant, alsmede twee waarnemingen waarbij het salpeterzuur 2 uur en twee waarnemingen waarbij het salpeterzuur 7 uur is toegevoegd. De lezer kan zelf nagaan, dat het waargenomen effect van de factor 'restant' 0,60 % en dat van de factor 'duur van de toevoeging van HNO_3 ' 1,25 % bedraagt. Ook deze effecten berusten op alle waarnemingen en zij zijn stuk voor stuk onafhankelijk van de andere effecten. Hier zien wij dus reeds het eerste voordeel van de factoriële opzet: bij elke vergelijking (betreffende niveau's van één factor) worden *alle* verrichte waarnemingen gebruikt.

Door middel van een variantie-analyse kan worden onderzocht, of de waargenomen effecten significant zijn. De berekeningen verlopen dan als volgt (ter vereenvoudiging zijn alle opbrengsten met 80 % verminderd):

$$\Sigma x^2 = (7,2)^2 + (8,4)^2 + \dots + (3,7)^2 = 290,18,$$

$$\Sigma x = 7,2 + 8,4 + \dots + 3,7 = 43,6, \quad (\Sigma x)^2/2 \times 2 \times 2 = 237,62,$$

$$S_{\text{totaal}} = 290,18 - 237,62 = 52,56.$$

$$S_A = \frac{(7,2 + 2,0 + 6,7 + 3,4)^2 + (8,4 + 3,0 + 9,2 + 3,7)^2}{4} - \frac{(43,6)^2}{8}$$

$$= 240,745 - 237,620 = 3,125,$$

$$S_B = \frac{(7,2 + 8,4 + 6,7 + 9,2)^2 + (2,0 + 3,0 + 3,4 + 3,7)^2}{4} - \frac{(43,6)^2}{8}$$

$$= 284,665 - 237,620 = 47,045,$$

$$S_C = \frac{(7,2 + 8,4 + 2,0 + 3,0)^2 + (6,7 + 9,2 + 3,4 + 3,7)^2}{4} - \frac{(43,6)^2}{8}$$

$$= 238,340 - 237,620 = 0,720.$$

Hieruit volgt:

$$S_{\text{residueel}} = 52,560 - (3,125 + 47,045 + 0,720) = 1,670.$$

De uitkomsten van de berekeningen zijn op de gebruikelijke wijze in tabel 22.4 samengevat.

Tabel 22.4. Uitkomsten van de variantie-analyse toegepast op de opbrengsten in tabel 22.3

	S	V.g.	Var.
A. Salpeterzuur	3,125	1	3,125
B. Roertijd	47,045	1	47,045
C. Restant	0,720	1	0,720
Residueel	1,670	4	0,4175
Totaal	52,560	7	

Wij veronderstellen dat elke opbrengst beschouwd kan worden als een aselechte trekking uit een (ongeveer) normaal verdeelde populatie met gemiddelde μ_{ijr} , dat de $2^3 = 8$ populaties dezelfde spreiding σ bezitten en dat ieder gemiddelde als volgt is samengesteld:

$$(22.1) \quad \mu_{ijr} = \mu + A_i + B_j + C_r.$$

Elke component bezit hier dus twee niveau's, bv. $B_1 =$ de component bij een roertijd van $\frac{1}{2}$ uur, $B_2 =$ de component bij een roertijd van 4 uur. Bij dit model is ondersteld dat de componenten der factoren onderling onafhankelijk zijn, zodat geen interactie optreedt. Wanneer men meent of weet, dat er wel interacties aanwezig zijn dient men de in 22.2.3 behandelde proefopzet met replicaties te gebruiken.

Men kan nu de volgende hypothesen toetsen (vergelijk met 17.3.3):

$H_0(1): A_1 = A_2$ — geen effect voor de factor salpeterzuur:

$F_0 = 3,125/0,4175 = 7,49$. Daar $F_{0,95} = 7,71$ bezit deze uitkomst een overschrijdingskans, die even groter is dan 0,05.

$H_0(2): B_1 = B_2$ — geen effect voor de factor roertijd:

$F_0 = 47,045/0,4175 = 116,68$, met een overschrijdingskans die aanzienlijk kleiner is dan 0,005.

$H_0(3): C_1 = C_2$ — geen effect voor de factor restant:

$F_0 = 0,720/0,4175 = 1,72$, met een overschrijdingskans die veel groter is dan 0,05.

De uit het onderzoek te trekken conclusie luidt dus, dat de hoogste opbrengst verkregen wordt als men slechts $\frac{1}{2}$ uur roert. Het aan- of afwezig zijn van het restant blijkt weinig of geen invloed op de opbrengst te hebben, maar er is een *indicatie* dat een toevoeging van het salpeterzuur in 7 uur eveneens de opbrengst verhoogt.

Als men deze proef op de klassieke manier uitvoert, dus steeds één factor varieert en daarbij de twee andere constant houdt, dan kan

men op de per factor verkregen opbrengsten een enkelvoudige variatie-analyse (met $k = 2$) toepassen. Wil men hierbij hetzelfde aantal vrijheidsgraden verkrijgen als bij de factoriële opzet, dan dient elke afzonderlijke proef uit zes waarnemingen te bestaan, zodat men in totaal $3 \times 6 = 18$ waarnemingen moet uitvoeren!

22.2.3. EEN 2^3 FACTORIEEL PROEFSHEMA MET REPLICATIES

De factoriële opzet bezit als tweede voordeel, dat men door het invoeren van replicaties kan nagaan, of er tussen de bestudeerde factoren interacties optreden. De klassieke proefopzet verschaft deze mogelijkheid niet: men verkrijgt daarbij slechts 'disconnected pieces, that cannot easily be pieced together' [COCHRAN en COX (52)].

In tabel 22.5 is een voorbeeld gegeven van een 2^3 factorieel schema met $m = 3$ replicaties. Het betreft wederom het onderzoek naar de opbrengst van het nitreringsproces van aniline en het is verkregen door aan de opbrengsten in tabel 22.3 twee (gefingeerde) replicaties toe te voegen. Alle opbrengsten zijn hier direct met 80 % vermindert.

Tabel 22.5. Een 2^3 factorieel schema met 3 replicaties

		C. Restant			
		Afwezig		Aanwezig	
		A. Toevoeging van HNO_3			
		2 uur	7 uur	2 uur	7 uur
B. Roertijd	1/2 uur	7,2	8,4	6,7	9,2
		7,9	8,1	7,9	8,9
		7,3	7,9	7,2	8,1
		22,4	24,4	21,8	26,2
	4 uur	2,0	3,0	3,4	3,7
		2,8	4,0	2,9	3,1
2,5		3,1	2,4	3,8	
7,3		10,1	8,7	10,6	
Factor A	{ 2 uur:	22,4 + 7,3 + 21,8 + 8,7 = 60,2			
	{ 7 uur:	24,4 + 10,1 + 26,2 + 10,6 = 71,3			
Factor B	{ 1/2 uur:	22,4 + 24,4 + 21,8 + 26,2 = 94,8			
	{ 4 uur:	7,3 + 10,1 + 8,7 + 10,6 = 36,7			
Factor C	{ Afwezig:	22,4 + 24,4 + 7,3 + 10,1 = 64,2			
	{ Aanwezig:	21,8 + 26,2 + 8,7 + 10,6 = 67,3			

Wij zullen het model, dat aan de analyse van dit schema ten grondslag ligt, niet volledig bespreken (het is echter een betrekkelijk eenvoudige uitbreiding van dat bij de tweevoudige analyse met replicaties

in 17.4). Wel zullen wij aangeven, op welke wijze de berekeningen en de toetsingen verlopen. Zoals steeds wordt eerst de totale kwadraten-som berekend:

$$\begin{aligned}\Sigma \Sigma x &= 7,2 + 7,9 + 7,3 + \dots + 3,8 = 131,5, \\ \Sigma \Sigma x^2 &= (7,2)^2 + (7,9)^2 + \dots + (3,8)^2 = 870,89 \\ (\Sigma \Sigma x)^2 / 2^3 \times 3 &= (131,5)^2 / 24 \dots\dots\dots = 720,5104 \\ S_{\text{totaal}} &= 150,3796.\end{aligned}$$

De verkregen opbrengsten kunnen nu in de eerste plaats worden beschouwd als 8 steekproeven van $m = 3$ elementen uit 8 verschillende populaties (waarbij wordt ondersteld, dat deze populaties (ongeveer) normaal verdeeld zijn met dezelfde spreiding). Past men dan de enkelvoudige analyse toe, dan vindt men

$$\begin{aligned}S_{\text{tussen}} &= \frac{(22,4)^2 + (24,4)^2 + \dots + (10,6)^2}{3} - \frac{(131,5)^2}{24} \\ &= 867,3833 - 720,5104 = 146,8729.\end{aligned}$$

Hieruit volgt dat

$$S_{\text{binnen}} = S_{\text{totaal}} - S_{\text{tussen}} = 150,3796 - 146,8729 = 3,5067.$$

De eerste stap van de analyse levert dus de in tabel 22.6 vermelde resultaten op.

Tabel 22.6. Factorieel schema met 3 replicaties:
eerste stap: enkelvoudige variantie-analyse

	S	V.g.
Tussen de steekproeven	146,8729	7
Binnen de steekproeven	3,5067	16
Totaal	150,3796	23

De kwadraten-som tussen de steekproeven en het bijbehorende aantal vrijheidsgraden kunnen nu weer in componenten worden gesplitst. *In het algemeen* geldt dat er bij drie factoren, t.w. A met k niveau's, B met n niveau's en C met r niveau's, tussen de steekproeven $knr - 1$ vrijheidsgraden zijn, die verdeeld kunnen worden zoals in tabel 22.7 is aangegeven.

Zijn er p factoren, elk met twee of meer niveau's, dan zijn er $p/1! = p$ hoofdeffecten, $p(p-1)/2!$ interacties van de eerste orde, $p(p-1) \times (p-2)/3!$ interacties van de tweede orde, enz. De aantallen vrijheidsgraden van deze interacties kunnen steeds worden bepaald, door het product te nemen van de aantallen vrijheidsgraden van de factoren (hoofdeffecten), waarop zij betrekking hebben. Voor het onderhavige

2^3 factoriële schema is $k=n=r=2$, zodat zowel de hoofdeffecten als de interacties één vrijheidsgraad bezitten.

Tabel 22.7. Verdeling van het aantal vrijheidsgraden 'tussen de steekproeven' bij een schema met drie factoren

Factor <i>A</i>	$k-1$	Hoofdeffecten
Factor <i>B</i>	$n-1$	
Factor <i>C</i>	$r-1$	
Interactie $A \times B$	$(k-1)(n-1)$	Interacties van 2 factoren = interacties van de eerste orde
Interactie $A \times C$	$(k-1)(r-1)$	
Interactie $B \times C$	$(n-1)(r-1)$	
Interactie $A \times B \times C$	$(k-1)(n-1)(r-1)$	Interactie van 3 factoren = interactie van tweede orde

De kwadratensommen van de factoren *A*, *B* en *C* kunnen nu op de gebruikelijke wijze worden berekend (zie tabel 22.5):

$$S_A = \frac{(60,2)^2 + (71,3)^2}{12} - \frac{(131,5)^2}{24} = 725,6442 - 720,5104 = 5,1338,$$

$$S_B = \frac{(94,8)^2 + (36,7)^2}{12} - \frac{(131,5)^2}{24} = 861,1608 - 720,5104 = 140,6504,$$

$$S_C = \frac{(64,2)^2 + (67,3)^2}{12} - \frac{(131,5)^2}{24} = 720,9108 - 720,5104 = 0,4004.$$

Na deze tweede stap van de analyse zijn de in tabel 22.8 staande gegevens beschikbaar (vergelijk met tabel 22.6)

Tabel 22.8. Factorieel schema met 3 replicaties: tweede stap van de analyse

	S	V.g.
Factor <i>A</i>	5,1338	1
Factor <i>B</i>	140,6504	1
Factor <i>C</i>	0,4004	1
Hoofdeffecten	146,1846	3
Interacties	0,6883	4
Tussen	146,8729	7
Binnen	3,5067	16
Totaal	150,3796	23

De derde fase van de analyse wordt gevormd door de berekening van de vier afzonderlijke interactie-kwadratensommen. Wij beschouwen

eerst de interactie van de factoren A en B (interactie $A \times B$). De opbrengsten in tabel 22.5 worden daartoe als volgt samengevat (factor C valt uit):

		A. Toevoeging van HNO_3			
		2 uur	7 uur		
B. Roertijd	}	½ uur	44,2	50,6	94,8
		4 uur	16,0	20,7	36,7
			60,2	71,3	131,5

Hierop kan de tweevoudige analyse worden toegepast, waarbij rekening dient te worden gehouden met het feit, dat elk getal in deze tabel op $2 \times 3 = 6$ waarnemingen berust, zodat de rij- en de kolomtotalen op 12 waarnemingen betrekking hebben. Men vindt dan

$$S_{\text{totaal } AB} = \frac{(44,2)^2 + (50,6)^2 + (16,0)^2 + (20,7)^2}{6} - \frac{(131,5)^2}{24}$$

$$= 866,4150 - 720,5104 = 145,9046,$$

$$S_A = \frac{(60,2)^2 + (71,3)^2}{12} - \frac{(131,5)^2}{24} = 5,1338,$$

$$S_B = \frac{(94,8)^2 + (36,7)^2}{12} - \frac{(131,5)^2}{24} = 140,6504.$$

Wij hebben hier volledigheidshalve S_A en S_B nogmaals berekend, maar deze kwadratensommen staan direct ter beschikking, daar zij reeds bij de tweede stap van de analyse berekend zijn. De interactie-kwadraten-som $A \times B$ is dus

$$S_{AB} = 145,9046 - (5,1338 + 140,6504) = 0,1204.$$

De interactie-kwadratensom van de factoren A en C kan berekend worden uit de tabel:

		C. Restant			
		Afwezig	Aanwezig		
A. Toevoeging van HNO_3	}	2 uur	29,7	30,5	60,2
		7 uur	34,5	36,8	71,3
			64,2	67,3	131,5

Hieruit volgt, dat

$$S_{\text{totaal } AC} = \frac{(29,7)^2 + (30,5)^2 + (34,5)^2 + (36,8)^2}{6} - \frac{(131,5)^2}{24}$$

$$= 726,1383 - 720,5104 = 5,6279,$$

zodat

$$S_{AC} = 5,6279 - (5,1338 + 0,4004) = 0,0937.$$

Voor de factoren *B* en *C* wordt de tabel:

		C. Restant		
		Afwezig	Aanwezig	
B. Roertijd	½ uur	46,8	48,0	94,8
	4 uur	17,4	19,3	36,7
		64,2	67,3	131,5

Men vindt dan

$$S_{\text{totaal } BC} = \frac{(46,8)^2 + (48,0)^2 + (17,4)^2 + (19,3)^2}{6} - \frac{(131,5)^2}{24}$$

$$= 861,5817 - 720,5104 = 141,0713,$$

zodat

$$S_{BC} = 141,0713 - (140,6504 + 0,4004) = 0,0205.$$

De interactie-kwadratensom $A \times B \times C$ kan nu het eenvoudigst worden verkregen, door van het totaal voor interacties (zie tabel 22.8) de som van de drie kwadratensommen van de eerste orde af te trekken, dus

$$S_{ABC} = 0,6883 - (0,1204 + 0,0937 + 0,0205) = 0,4537.$$

De volledige uitkomsten van de uitgevoerde analyse zijn nu tenslotte in tabel 22.9 opgenomen.

Tabel 22.9. Factorieel schema met 3 replicaties. Volledige uitkomsten van de analyse, toegepast op de opbrengsten in tabel 22.5

	S	V.g.	Var.
Factor <i>A</i>	5,1338	1	5,1338
Factor <i>B</i>	140,6504	1	140,6504
Factor <i>C</i>	0,4004	1	0,4004
Interactie $A \times B$	0,1204	1	0,1204
Interactie $A \times C$	0,0937	1	0,0937
Interactie $B \times C$	0,0205	1	0,0205
Interactie $A \times B \times C$	0,4537	1	0,4537
Binnen	3,5067	16	0,2192
Totaal	150,3796	23	

De hypothese, dat geen interactie $A \times B \times C$ aanwezig is, kan nu worden getoetst door de waargenomen $s^2_{ABC} = 0,4537$ door middel van de *F*-toets te vergelijken met $s^2_b = 0,2192$. Men vindt dan, dat $F_0 = 0,4537/0,2192 = 2,07$ en als de toetsing bij een 5% drempel wordt uitgevoerd, is de kritieke waarde $F_{0,95}(1, 16) = 4,49$, zodat de getoetste hypothese niet kan worden verworpen. Op dezelfde wijze kan men de nulhypothese betreffende de overige interacties en de hoofdeffecten

uitvoeren. Direct is in te zien, dat slechts de factoren A en B een significante waarde van F opleveren. Ten aanzien van de hoofdeffecten komen wij dus met deze proefopzet tot dezelfde conclusie als bij de proef zonder replicaties in 22.2.2, maar wij kunnen nu daaraan toevoegen, dat geen interacties van betekenis zijn geconstateerd.¹

22.2.4. EEN $3 \times 4 \times 5$ FACTORIEEL SCHEMA

Als volgend voorbeeld van een factoriële proefopzet beschouwen wij de kaliumgehalten van het parotisspeeksel in tabel 22.2, die volgens een $3 \times 4 \times 5$ schema (zonder replicaties) zijn verzameld. Deze gehalten, verminderd met 100, zijn in tabel 22.10 nogmaals opgenomen, met het merendeel der sub-totalen die voor de berekeningen nodig zijn.

Tabel 22.10. Een $3 \times 4 \times 5$ factorieel schema zonder replicaties. Kaliumgehalten van het parotisspeeksel uit tabel 22.2 (verminderd met 100)

Datum (A)	Uur (B)	Proefpersoon (C)					Totaal
		I	II	III	IV	V	
3/4	10	-14	-7	-1	-9	-16	-47
	12	-2	-1	+8	+5	-9	+1
	14	+10	+7	+14	+4	-2	+33
	16	+10	0	+18	+18	-11	+35
			+4	-1	+39	+18	-38
4/4	10	+3	-4	-8	-7	-1	-17
	12	+2	-5	+5	-18	+6	-10
	14	+21	-2	+23	+12	-2	+52
	16	+8	+9	+6	+7	-9	+21
			+34	-2	+26	-6	-6
5/4	10	+8	-8	+2	-13	+4	-7
	12	+10	-3	+4	-8	+1	+4
	14	+13	+2	+2	+7	+5	+29
	16	+6	+3	+10	-5	+7	+21
			+37	-6	+18	-19	+17
Totaal		+75	-9	+83	-7	-27	+115

Totalen per tijdstip:

$$10 \text{ uur: } (-47) + (-17) + (-7) = -71$$

$$12 \text{ uur: } (+1) + (-10) + (+4) = -5$$

$$14 \text{ uur: } (+33) + (+52) + (+29) = +114$$

$$16 \text{ uur: } (+35) + (+21) + (+21) = +77$$

¹ Hierbij dient te worden opgemerkt, dat deze bevindingen niet met de werkelijkheid behoeven overeen te stemmen, daar de replicaties fictief zijn.

De berekeningen verlopen nu als volgt:

$$\begin{aligned} \Sigma\Sigma x^2 &= (-14)^2 + (-7)^2 + (-1)^2 + \dots + (7)^2 = 4955 \\ (\Sigma\Sigma x)^2 / 3 \times 4 \times 5 &= (115)^2 / 60 \dots\dots\dots = 220,4167 \\ S_{\text{totaal}} &= 4734,5833, \end{aligned}$$

$$S_A = \frac{(22)^2 + (46)^2 + (47)^2}{20} - \frac{(115)^2}{60} = 240,45 - 220,4167 = 20,0333,$$

$$\begin{aligned} S_B &= \frac{(-71)^2 + (-5)^2 + (114)^2 + (77)^2}{15} - \frac{(115)^2}{60} = 1599,4 - 220,4167 \\ &= 1378,9833, \end{aligned}$$

$$\begin{aligned} S_C &= \frac{(75)^2 + (-9)^2 + \dots + (-27)^2}{12} - \frac{(115)^2}{60} = 1114,4167 - 220,4167 \\ &= 894. \end{aligned}$$

In tabel 22.11 zijn de tot dusver verkregen uitkomsten van de analyse samengevat.

Tabel 22.11. Resultaten van de eerste fase van de analyse van een 3 x 4 x 5 schema

	S	V.g.
Factor A (dagen)	20,0333	2
Factor B (uren)	1378,9833	3
Factor C (personen)	894,0000	4
Hoofdeffecten	2293,0166	9
Interacties (totaal)	2441,5667	50
Totaal	4734,5833	59

De totale kwadratensom van de interacties, die door aftrekking verkregen is, dient nu nog in vier componenten te worden gesplitst. De kwadratensom voor de interactie A x B (met 2 x 3 = 6 vrijheidsgraden) volgt op de gebruikelijke wijze uit onderstaande tabel:

	Factor B				Totaal	
	10	12	14	16		
Factor A {	3/4	-47	+1	+33	+35	+22
	4/4	-17	-10	+52	+21	+46
	5/4	-7	+4	+29	+21	+47
Totaal	-71	-5	+114	+77	+115	

$$\begin{aligned} S_{\text{totaal } AB} &= \frac{(-47)^2 + (1)^2 + (33)^2 + \dots + (21)^2}{5} - \frac{(115)^2}{60} \\ &= 1881 - 220,4167 = 1660,5833, \end{aligned}$$

$$S_{AB} = 1660,5833 - (20,0333 + 1378,9833) = 261,5667.$$

De interactie-kwadratensom $A \times C$ (met $2 \times 4 = 8$ vrijheidsgraden) wordt berekend uit de tabel:

		Factor C					Totaal
		I	II	III	IV	V	
Factor A	3/4	+ 4	- 1	+39	+18	-38	+22
	4/4	+34	- 2	+26	- 6	- 6	+46
	5/4	+37	- 6	+18	-19	+17	+47
Totaal		+75	- 9	+83	- 7	-27	+115

$$S_{\text{totaal } AC} = \frac{(4)^2 + (-1)^2 + (39)^2 + \dots + (17)^2}{4} - \frac{(115)^2}{60}$$

$$= 1898,25 - 220,4167 = 1677,8333,$$

$$S_{AC} = 1677,8333 - (20,0333 + 894) = 763,8000.$$

De interactie-kwadratensom $B \times C$ (met $3 \times 4 = 12$ vrijheidsgraden) volgt uit de volgende tabel (die eenvoudig uit tabel 22.10 kan worden samengesteld, door per persoon alle 3 gehalten op één tijdstip te sommeren):

		Factor C					Totaal
		I	II	III	IV	V	
Factor B	10	- 3	-19	- 7	-29	-13	-71
	12	+10	- 9	+17	-21	- 2	- 5
	14	+44	+ 7	+39	+23	+ 1	+114
	16	+24	+12	+34	+20	-13	+77
Totaal		+75	- 9	+83	- 7	-27	+115

$$S_{\text{totaal } BC} = \frac{(-3)^2 + (-19)^2 + (-7)^2 + \dots + (-13)^2}{3} - \frac{(115)^2}{60}$$

$$= 2941,6667 - 220,4167 = 2721,25,$$

$$S_{BC} = 2721,25 - (1378,9833 + 894) = 448,2667.$$

De interactie-kwadratensom $A \times B \times C$ (met $2 \times 3 \times 4 = 24$ vrijheidsgraden) kan tenslotte het eenvoudigst worden verkregen, door het totaal voor interacties in tabel 22.11 met de som van de drie interacties van de eerste orde te verminderen:

$$S_{ABC} = 2441,5667 - (261,5667 + 763,8000 + 448,2667) = 967,9333.$$

De berekende kwadratensommen zijn met hun aantal vrijheidsgraden en hun variantie-schatting in tabel 22.12 opgenomen.

Tabel 22.12. Factorieel $3 \times 4 \times 5$ schema. Volledige uitkomsten van de analyse, toegepast op de gegevens in tabel 22.10

	S	V.g.	Var.
Factor <i>A</i> (dagen)	20,0333	2	10,02
Factor <i>B</i> (uren)	1378,9833	3	459,66
Factor <i>C</i> (personen)	894,0000	4	223,50
Interactie $A \times B$	261,5667	6	43,59
Interactie $A \times C$	763,8000	8	95,48
Interactie $B \times C$	448,2667	12	37,36
Interactie $A \times B \times C$	967,9333	24	40,33
Totaal	4734,5833	59	

De wijze waarop nu wordt getoetst is afhankelijk van het model, dat men kiest. Daar dit hoofdstuk bedoeld is om de mogelijkheden van de meer gecompliceerde proefopzetten te belichten, gaan wij hierop niet uitgebreid in. Het is echter duidelijk, dat bij dit schema bekend moet zijn (of moet worden aangenomen), dat er geen interactie $A \times B \times C$ aanwezig is, omdat slechts onder deze voorwaarde $s^2_{ABC} = 40,33$ een schatting van σ^2 is. Heeft men redenen om aan te nemen, dat deze interactie wel optreedt, dan dient men de proef met replicaties uit te voeren.

Zonder dieper op het model in te gaan kan men echter uit tabel 22.12 de volgende conclusies trekken:

1. Van de interacties is slechts $A \times C$ significant bij een 5% drempel, daar $F_0 = 95,48/40,33 = 2,37$ en $F_{0,95}(8, 24) = 2,36$. Bestudeert men de tabel op blz. 578 waaruit de interactie-kwadratensom $A \times C$ is berekend, dan blijkt inderdaad het verloop van de kaliumgehalten over de drie onderzoekdagen voor de proefpersonen verschillend te zijn. De gehalten van de personen I en V geven een stijgend verloop te zien, bij III en IV treedt een dalend verloop op en bij II zijn er geringe verschillen tussen de dagen.

2. Het verschil tussen de uren is zeer significant, daar $F_0 = 459,66/40,33 = 11,40$ en $F_{0,995}(3, 24) = 5,52$. Uit tabel 22.10 blijkt dat in het algemeen van 10 t/m 14 uur een stijgend verloop optreedt en dat daarna een zekere daling plaatsvindt.

3. Het verschil tussen de dagen is in ieder geval niet significant.

22.2.5. LATIJNSE EN GRIEKSLATIJNSE VIERKANTEN

Het zg. Latijnse vierkant is een bijzonder proefschema, dat kan worden gebruikt voor een proef met drie factoren zonder replicaties. Hierbij brengt men feitelijk in een tweevoudig schema met evenveel rijen als kolommen nog een derde factor onder. Hoe dit in zijn werk gaat blijkt uit het volgende voorbeeld. Het betreft een onderzoek naar het effect van vier verschillende extracten van de bijschildklier (*A*, *B*, *C* en *D*) op het serum-calcium gehalte bij honden. Men dient elk dier op vier tijdstippen (met flinke tussenruimten) achtereenvolgens

een dosis van elk extract toe en bepaalt de calciumgehalten 24 uur later. Bij deze proef kan men bv. gebruik maken van het Latijnse vierkant dat links in tabel 22.13 is gegeven. Men ziet, dat de 'Latijnse' letters welke overeenkomen met de vier preparaten zo over het vierkant zijn verdeeld, dat elk preparaat eenmaal bij elke hond en eenmaal op elk tijdstip voorkomt. Op deze wijze bereikt men, dat een bepaalde combinatie van de drie factoren precies één keer voorkomt (maar niet alle mogelijke combinaties komen voor: dit zijn er nl. 4^3 , terwijl het vierkant er slechts 4^2 bevat). Dit is een markant voordeel van het Latijnse vierkant boven het gebruikelijke drie-factoren experiment. In verschillende boeken en in FISHER en YATES (66) zijn vele Latijnse vierkanten (van 3×3 t/m 12×12) opgenomen.

Tabel 22.13. Latijns vierkant voor een proef met vier preparaten, vier tijdstippen van toediening en vier proefdieren

PROEFOPZET					UITKOMSTEN serum-calcium in mg%					
Honden	Tijdstippen				Honden	Tijdstippen				Som
	1	2	3	4		1	2	3	4	
1	A	B	D	C	1	12,8	17,0	16,0	15,0	60,8
2	D	C	A	B	2	14,8	13,3	14,8	15,4	58,3
3	B	A	C	D	3	14,0	14,5	13,0	15,0	56,5
4	C	D	B	A	4	12,7	15,4	14,8	14,0	56,9
Som					Som	54,3	60,2	58,6	59,4	232,5

UITKOMSTEN PER PREPARAAT

Honden	Preparaten			
	A	B	C	D
1	12,8	17,0	15,0	16,0
2	14,8	15,4	13,3	14,8
3	14,5	14,0	13,0	15,0
4	14,0	14,8	12,7	15,4
Som	56,1	61,2	54,0	61,2

De uitkomsten van de proef staan eveneens in tabel 22.13, eerst per hond/per tijdstip en vervolgens per hond/per preparaat. De berekeningen behoeven o.i. geen toelichting en verlopen op de volgende wijze:

$$S_{\text{totaal}} = (12,8)^2 + (17,0)^2 + \dots + (15,4)^2 - (232,5)^2/16 = 21,1944,$$

$$S_{\text{honden}} = \frac{(60,8)^2 + (58,3)^2 + (56,5)^2 + (56,9)^2}{4} - \frac{(232,5)^2}{16} = 2,8319,$$

$$S_{\text{tijden}} = \frac{(54,3)^2 + (60,2)^2 + (58,6)^2 + (59,4)^2}{4} - \frac{(232,5)^2}{16} = 5,1969,$$

$$S_{\text{prep}} = \frac{(56,1)^2 + (61,2)^2 + (54,0)^2 + (61,2)^2}{4} - \frac{(232,5)^2}{16} = 10,0069,$$

$$S_{\text{res}} = 21,1944 - (2,8319 + 5,1969 + 10,0069) = 3,1587.$$

Deze uitkomsten zijn in tabel 22.14 opgenomen.

Tabel 22.14. Latijns vierkant: uitkomsten van de analyse van de calciumgehalten in tabel 22.13

	S	V.g.	Var.schatting
Honden	2,8319	3	0,944
Tijdstippen	5,1969	3	1,732
Preparaten	10,0069	3	3,336
Residueel	3,1587	6	0,526
Totaal	21,1944	15	

Zoals gebruikelijk dient men nu te onderstellen, dat alle waarnemingen afkomstig zijn uit normaal verdeelde populaties met dezelfde spreiding. Een andere voorwaarde is, dat de drie bij de proef betrokken factoren geen interacties vertonen, omdat slechts onder deze voorwaarde de residuele variantie een schatting is van σ^2 . Indien hieraan voldaan is kan men de hypothese, dat de preparaten dezelfde invloed op het calciumgehalte van het serum uitoefenen toetsen, door de preparatenvariantie met de residuele te vergelijken. Men vindt dan $F_0(3, 6) = 6,34$ en deze uitkomst is significant bij een 5% drempel, daar $F_{0,95}(3, 6) = 4,76$. Wanneer men de factor 'tijdstip (van waarneming)' niet in de proef betreft, verkrijgt men de uitkomsten die links onder in tabel 22.13 staan. Past men hierop de tweevoudige analyse (volgens 17.3) toe, dan vindt men:

	S	V.g.	Var.schatting
Honden	2,8319	3	0,944
Preparaten	10,0069	3	3,336
Residueel	8,3556	9	0,928
Totaal	21,1944	15	

Voor de preparaten is nu $F_0(3, 9) = 3,336/0,928 = 3,59$, een waarde die niet significant is bij een 5% drempel, aangezien $F_{0,95}(3, 9) = 3,86$.

Een bezwaar van het Latijnse vierkant is, dat de residuele variantie het betrekkelijk kleine aantal van $(n-1)(n-2)$ vrijheidsgraden bezit. Dit bezwaar kan worden ondervangen door bij één proef een aantal Latijnse vierkanten te gebruiken, een procedure die o.m. besproken wordt door EDWARDS (55) en BENNETT en FRANKLIN (4).

Het Latijnse vierkant kan ook worden gecombineerd met een factoriële proefopzet. Onderstel bijvoorbeeld, dat men in een klimaatkamer een proef wil uitvoeren, waarbij per proefpersoon de huidtemperatuur op verschillende plaatsen (meetpunten) wordt gemeten

en dat hierbij de volgende drie factoren op twee niveau's worden gevarieerd:¹

- I — Inspanning bij het lopen op een tredmolen, — = laag, + = hoog.
- II — Temperatuur, — = laag, + = hoog.
- III — Windsnelheid, — = laag, + = hoog.

In de volgende tabel zijn de acht mogelijke combinaties van deze factoren opgenomen en met de letters A t/m H aangeduid (zij worden wel de experimentele condities genoemd):

		Factor I			
		—		+	
		Factor II			
		—	+	—	+
Factor III	+	A	B	C	D
	—	E	F	G	H

Onderstel nu, dat men de proef bij 8 personen uitvoert en dat steeds dezelfde 8 meetpunten worden gebruikt. Men kan dan de 8 experimentele condities in een 8×8 Latijns vierkant onderbrengen, bv. zoals dat in tabel 22.15 is aangegeven.

Tabel 22.15. Combinatie van een 2³ factoriële proefopzet en een Latijns vierkant

Personen	Meetpunten							
	1	2	3	4	5	6	7	8
1	A	B	C	D	E	F	G	H
2	B	C	A	E	F	D	H	G
3	C	A	D	G	H	E	F	B
4	D	F	G	C	A	H	B	E
5	E	H	B	F	G	C	A	D
6	F	D	H	A	B	G	E	C
7	G	E	F	H	C	B	D	A
8	H	G	E	B	D	A	C	F

Als men de proef op deze wijze opzet, neemt de analyse van de uitkomsten de in tabel 22.16 gegeven vorm aan. De uit te voeren berekeningen volgen direct uit het voorgaande.

¹ Wij laten verdere bijzonderheden omtrent deze factoren en hun niveau's achterwege.

Tabel 22.16. Analyse van het proefschema in tabel 22.15

	V.g.
Proefpersonen	7
Meetpunten	7
Experimentele condities:	
Factor I	1
Factor II	1
Factor III	1
Interactie I × II	1
Interactie I × III	1
Interactie II × III	1
Interactie I × II × III	1
Residueel	42
Totaal	63

Bij de methode van het *Grieks-Latijnse vierkant* gaat men nog een stap verder, door nog een vierde factor in een Latijns vierkant onder te brengen. De niveau's van deze factor worden met Griekse letters aangeduid en zij worden zo over het vierkant verdeeld, dat elk niveau eenmaal op elke rij, eenmaal op elke kolom en bovendien slechts eenmaal samen met elke 'Latijnse' letter voorkomt. Het resultaat is, dat voor elk paar van factoren elke combinatie der niveau's precies één keer voorkomt; tabel 22.17 geeft een voorbeeld van een 4 × 4 Grieks-Latijns vierkant. Een belangrijk bezwaar is echter, dat de residuele variantie slechts $(n-1)(n-3)$ vrijheidsgraden bezit, zodat men gewoonlijk met meer dan één vierkant zal moeten werken. MAXWELL (178) geeft hiervan een voorbeeld.

Tabel 22.17. Een 4 × 4 Grieks-Latijns vierkant

		Factor I				
		1	2	3	4	
Factor II	1	A α	B β	C γ	D δ	De niveau's van factor III zijn A, B, C en D.
	2	B γ	A δ	D α	C β	
	3	C δ	D γ	A β	B α	
	4	D β	C α	B δ	A γ	De niveau's van factor IV zijn α , β , γ en δ .

22.2.6. ENKELE OPMERKINGEN BETREFFENDE HET GEBRUIK VAN STATISTISCHE PROEFSHEMA'S

De proefschemas die in de voorgaande paragrafen zijn besproken vormen slechts een klein deel van het uitgebreide arsenaal, dat de onderzoeker tegenwoordig ter beschikking staat. Wij vertrouwen

echter, dat zij de lezer ervan hebben overtuigd, dat deze moderne proefopzetten een belangrijke bijdrage kunnen leveren tot de verhoging van de efficiency van een experiment. Zij leiden in het algemeen met aanzienlijk minder waarnemingen tot nauwkeuriger en beter gefundeerde conclusies dan de klassieke wijze van experimenteren, waarbij men steeds één factor tegelijk laat variëren.

De proefschema's zijn oorspronkelijk ontwikkeld ten behoeve van landbouwkundige experimenten. Aangezien men daarbij gewoonlijk slechts één waarneming per jaar kan verrichten, maar vele waarnemingen gelijktijdig onder verschillende omstandigheden kan doen, zijn hier de gecompliceerde proefopzetten op hun plaats, die bv. door COCHRAN en COX (52) en KEMPTHORNE (60) worden beschreven. Op vele andere gebieden vereisen de waarnemingen veel minder tijd en men maakt dan meestal gebruik van eenvoudiger proefopzetten, waarbij niet meer dan 3 of 4 factoren tegelijkertijd worden gevarieerd. In BENNETT en FRANKLIN (4) vindt men vele voorbeelden op chemisch gebied, door EDWARDS (55) worden psychologische toepassingen besproken en allerlei industriële problemen vindt men o.m. in BROWNLEE (133), DAVIES (141) en in de verschillende artikelen van HAMAKER (161, 162, 142). Op het medische terrein noemen wij bv. het gebruik van proefschema's bij de biologische ijking waarop wij in hoofdstuk 23 terugkomen.

Een ander voordeel van de statistische proefopzetten is, dat de onderzoeker gedwongen wordt zich een scherp beeld te vormen van het gestelde probleem en van de factoren die daarbij van belang (kunnen) zijn. Deze zijn gewoonlijk te talrijk om ze alle in één proef te verwerken, zodat men hieruit —mede door middel van literatuurstudie of door enkele voorbereidende proeven— een weloverwogen keuze dient te maken. Men moet daarbij steeds bedacht zijn op de mogelijke interacties tussen factoren en hiermee bij de opzet van de proef rekening houden. Wij zijn ervan overtuigd dat de onderzoeker, die de principes van de moderne experimenteermethoden kent, op deze wijze tot een eenvoudige maar doeltreffende factoriële proefopzet kan komen, waarbij de statistische analyse van de uitkomsten weinig moeilijkheden zal opleveren. WALLIS (230) merkt dan ook terecht op: 'In general, if a statistical investigation . . . is well planned and the data properly collected the interpretation will pretty well take care of itself. So-called 'high-powered', 'refined', or 'elaborate' statistical techniques are generally called for when the data are crude and inadequate — exactly the opposite, if I may be permitted an *obiter dictum*, of what crude and inadequate statisticians usually think'. Een ander belangrijk argument vòòr het gebruik van niet te gecompliceerde proefopzetten is nog, dat men de juistheid van de verschillende noodzakelijke onderstellingen beter kan nagaan. Men kan bv.

uitbijters, waarvoor een variantie-analyse zeer gevoelig is, bij een eenvoudiger proefopzet beter signaleren dan bij een complexe.

De gegevens in tabel 10.6 (blz. 230) leveren een fraai voorbeeld van de gevoeligheid van de variantie-analyse voor uitbijters. De waarneming van analyst *B* betreffende het tweede bloedmonster is hier een evidente uitbijter. De opgegeven waarde van 13,4 g% berust dan ook op een fout en zij moet 12,5 g% zijn. De uitkomst van de (verdelingsvrije) toets van FRIEDMAN verandert door deze wijziging niet en de exacte overschrijdingskans is en blijft 0,052. Past men echter op de oorspronkelijke en op de voor deze fout gecorrigeerde gegevens een variantie-analyse toe, dan zijn de resultaten:

	Vr. gr.	Oorspronkelijk		Gecorrigeerd	
		S	s ²	S	s ²
Tussen de analysten	2	0,3079	0,15395	0,1378	0,0689
Tussen de bloedmonsters	5	56,9829		58,5978	
Residueel	10	0,6653	0,06653	0,1155	0,01155
Totaal	17	57,9561		58,8511	
		$F(2, 10) = 2,31$ $0,10 < P_R < 0,20$		$F(2, 10) = 5,97$ $0,01 < P_R < 0,025$	

22.3. Speciale steekproeftechnieken bij onderzoeken

22.3.1. LAAGSGEWIJS ASELECTE STEEKPROEVEN

Deze steekproeftechniek kan worden gebruikt als men te maken heeft met een populatie van complexe structuur, die op grond van reeds aanwezige kennis omtrent haar samenstelling in een aantal subpopulaties (lagen of strata) kan worden verdeeld. Men verkrijgt dan een laagsgewijs aselecte (gelaagde of gelede) steekproef (Eng. 'stratified random sample') door uit elke laag aselect een steekproef te trekken. Het gebruik van zo'n laagsgewijs opgebouwde steekproef is van belang, als de criteria waarop de stratificatie berust zodanig kunnen worden gekozen, dat de gemiddelden en/of de spreidingen van de lagen duidelijke verschillen vertonen. Een laagsgewijs aselecte steekproef van een bepaalde omvang veronacht dan nauwkeuriger schattingen betreffende de populatie (bv. van het populatiegemiddelde of van de fractie elementen met een bepaald kenmerk in de populatie) dan een aselecte steekproef van dezelfde omvang. Als men te maken heeft met *k* lagen, die in de populatie resp. de omvang N_1, N_2, \dots, N_k bezitten ($\sum_{i=1}^k N_i = N =$ de populatie-omvang) en waarvan de spreidingen $\sigma_1, \sigma_2, \dots, \sigma_k$ zijn, kan men de omvang der steekproeven uit deze lagen het best in de verhouding $N_1\sigma_1 : N_2\sigma_2 : \dots : N_k\sigma_k$ kiezen. Zijn de spreidingen der lagen gelijk, dan is dus de gunstigste verhouding $N_1 : N_2 : \dots : N_k$.

Laagsgewijs aselechte steekproeven worden voornamelijk toegepast bij opinie-, budget- en marktonderzoeken. De criteria voor de stratificatie zijn dan bv. geslacht, leeftijd, welstand, woonplaats (of de grootte daarvan), gezinsgrootte, e.d. Bij medische onderzoeken betreffende complex gestructureerde (en dus heterogene) populaties is men gewoonlijk meer geïnteresseerd in het bestuderen van de verschillen tussen de lagen dan in het geven van schattingen betreffende de totale populatie. Voor verdere bijzonderheden betreffende deze steekproeftechniek en de formules die voor verschillende schattingen gelden verwijzen wij dan ook naar de literatuur [zie bv. COCHRAN (52) en YATES (64)].

22.3.2. STEEKPROEVEN BESTAANDE UIT GROEPEN

Bij deze steekproeftechniek verdeelt men de populatie in een groot aantal betrekkelijk kleine groepen van elementen (Eng. 'clusters') en aselechte een steekproef van k groepen. Men kan dan alle elementen in trekt een groep of een aselechte gekozen deel daarvan onderzoeken. Deze procedure is voordelig, wanneer de kosten voor het onderzoeken van één element laag, maar die voor het verkrijgen van dit element betrekkelijk hoog zijn, bv. door een sterke geografische spreiding van de tot de populatie behorende elementen. Beschouw bv. een onderzoek waarbij de populatie bestaat uit alle Nederlandse kinderen, die op 1 oktober 1959 een lagere school bezochten. Onderstel, dat het onderzoeken van elk kind ongeveer 5 minuten vergt. Trekt men nu uit deze populatie aselechte een steekproef van bv. 1000 kinderen, dan zullen deze kinderen door het gehele land verspreid zijn, zodat per kind een veelvoud van de onderzoektijd aan reis- en verblijfkosten moet worden uitgetrokken. Kiest men echter een aselechte steekproef van bv. 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 alle kinderen kan onderzoeken. Dit onderzoek van ongeveer 3500 kinderen zou wel eens minder tijdrovend en dus minder kostbaar kunnen zijn dan het eerstgenoemde en het zal wellicht meer informatie opleveren.¹

Men verkrijgt met deze techniek de gunstigste resultaten, als de groepen ongeveer even groot zijn en als de elementen in elke groep heterogeen zijn. Hoewel wij voor uitvoeriger gegevens weer naar de literatuur verwijzen [zie vooral COCHRAN (52)] geven wij wel enkele formules die gelden, als men elke getrokken groep volledig onderzoekt.

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

Wij gebruiken daarbij de volgende notatie:

M = het aantal groepen in de populatie,

m = het aantal groepen in de steekproef,

n_i = het aantal elementen in groep i van de steekproef,

\bar{x}_i = het gemiddelde van groep i van de steekproef,

$\bar{n} = \frac{\sum_{i=1}^m n_i}{m}$ = het gemiddelde aantal elementen per groep,

$n = \sum n_i$ = het totale aantal elementen in de steekproef.

Men kan een zuivere schatting van het populatiegemiddelde μ verkrijgen door te nemen

$$(22.2) \quad \bar{x} = \frac{\sum_{i=1}^m n_i \bar{x}_i}{\sum_{i=1}^m n_i} = \frac{\sum \sum x_i}{n}.$$

De variantie van \bar{x} is

$$(22.3) \quad s_{\bar{x}}^2 = \frac{M-m}{Mm} \cdot \frac{\sum_{i=1}^m n_i^2 (\bar{x}_i - \bar{x})^2}{(n-1) (\bar{n})^2}.$$

Wanneer $n_1 = n_2 = \dots = n_k$, dan kunnen deze formules worden vereenvoudigd tot

$$(22.4) \quad \bar{x} = \frac{\sum_{i=1}^m \bar{x}_i}{m}$$

en

$$(22.5) \quad s_{\bar{x}}^2 = \frac{M-m}{Mm} \cdot \frac{\sum_{i=1}^m (\bar{x}_i - \bar{x})^2}{n-1}.$$

Bij een dichotomie verkrijgt men een zuivere schatting p van de parameter P door in (22.2) of (22.4) \bar{x}_i te vervangen door p_i . De formules voor s_p^2 volgen dan uit (22.3) of (22.5) door in plaats van \bar{x}_i te nemen p_i en in plaats van \bar{x} te nemen p .

STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

23.1. Inleiding

In de geneeskunde maakt men vaak gebruik van middelen welke van dierlijke, plantaardige of minerale oorsprong zijn. Een schatting van de werkzaamheid van zo'n natuurlijk middel, zoals bv. een gift dat door middel van een of andere extractiemethode uit plantaardig materiaal is verkregen, kan dan gewoonlijk niet via een chemische analyse plaatsvinden. Hierbij komt nog, dat in vele gevallen de kennis van de precieze samenstelling van een middel niet voldoende is om uit te maken, welk effect het zal veroorzaken. In dergelijke gevallen dient men voor het bepalen van de activiteit gebruik te maken van een *biologische standaardisatiemethode* (of ijking).

FINNEY (153) geeft van zo'n methode de volgende kenschets: 'The typical bioassay involves a *stimulus* (for example, a vitamin, a drug a fungicide) applied to a *subject* (for example an animal, a piece of animal tissue, a bacterial culture). The intensity of the stimulus may be varied . . . so as to vary the dose given to the subject, and this dose can be measured (perhaps as a weight, a volume, or a concentration). Application of the stimulus is followed by a change in some measurable characteristic of the subject, the magnitude of the change being dependent upon the dose. A measurement of this characteristic (for example, a weight of the whole subject or of some particular organ, an analytical value such as blood sugar content or bone ash percentage, or even a simple record of occurrence or non-occurrence of a certain muscular contraction, recovery from the symptoms of a dietary deficiency, or death) is the *response* of the subject. The relationship between dose and response will not be exact, but will be obscured by random variations between replicate subjects, or, if the response is one that can be produced more than once in the same subject, in the responsiveness of one subject on different occasions. Nevertheless, the relationship can be used to indicate the potency of a dose from knowledge of the response it produces. . . . Biological assays are usually comparative, the estimate of potency being obtained relative to a *standard preparation* of the stimulus. This standard preparation may be merely a convenient working standard adopted in a particular laboratory as a scale of reference or it

may have a wider currency. The Permanent Commission on Biological Standardization of the League of Nations, has performed great services in the preparation and supply of standard preparations of many important therapeutic substances. For most assays of these substances, the standard preparation is likely to be a sample of either the international standard or a more readily available working standard whose potency relative to the international standard has previously been carefully evaluated.

The potency of the standard is expressed either in ordinary units of weight or volume or in arbitrary units, such as the international units adopted for insulin and penicillin. Any *test preparation* of the stimulus, having an unknown potency, is assayed by finding the mean response to a selected dose, and equating this those to that of the standard preparation shown by experiment to produce the same mean response; experimentation with several different doses of one or both preparations is almost always needed in order to accomplish this satisfactorily. The ratio of the two equally effective doses is an estimate of the potency of the test preparation relative to that of the standard.'

Men kan drie soorten biologische standaardisatie-methoden onderscheiden, nl.: 1. Directe standaardisatie, 2. Indirecte standaardisatie gebaseerd op kwantitatieve responsies, en 3. Indirecte standaardisatie gebaseerd op het wel-of-niet reageren.¹ Wij zullen per soort methode een korte beschouwing wijden aan de daarbij gebruikte statistische technieken.

23.2. Directe standaardisatie

Een directe standaardisatie kan worden uitgevoerd als men voor het testpreparaat en voor het standaardpreparaat precies de dosis kan meten, die juist voldoende is om bij een proefdier een bepaalde, ondubbelzinnig waarneembare reactie te veroorzaken.² Daar deze z.g. *individueel-effectieve dosis* van dier tot dier varieert dient men bij de ijking gebruik te maken van meerdere proefdieren. In het ideale geval, dat men beide preparaten successievelijk aan één dier kan geven, kan men met één groep dieren volstaan, maar vaak is dit niet mogelijk, bv. omdat het dier door de eerste proef reeds ongeschikt wordt voor een volgende of omdat de bestudeerde reactie de dood van het dier is. Wanneer men twee groepen proefdieren gebruikt dienen deze vanzelfsprekend aselekt te worden samengesteld. De ene groep wordt behandeld met het testpreparaat, de andere met de standaard. Per preparaat verkrijgt men dan de individueel-effectieve doses van de

¹ Niet reageren kan in dit verband ook inhouden, dat de reactie een bepaalde, tevoren gedefinieerde maat niet overschrijdt.

² Wij laten hier proeven met twee of meer testpreparaten buiten beschouwing.

23.2 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

dieren en het gemiddelde van deze doses is een schatting van de gemiddelde activiteit van het preparaat. De verhouding van de gemiddelden van het testpreparaat en het standaardpreparaat vormt een schatting van de *zg. activiteitsverhouding* (Eng. potency ratio) van deze preparaten. Daar de *logarithmen* van de individueel-effectieve doses gewoonlijk ongeveer normaal verdeeld blijken te zijn, wordt de statistische analyse van de gegevens op de grootheid $x = \log \text{ dosis}$ uitgevoerd.

In een artikel over het kwantitatieve verband tussen dosering en letaliteit bij proefdieren schrijft NASS (182) o.m. het volgende betreffende deze transformatie: 'De logarithme van de dosis, als maat voor de gevoeligheid van een willekeurig proefdier, (is) 'normaler' gedistribueerd dan de dosis zelf. Blijkens de ervaring doet zich dit verschijnsel bij de overgrote meerderheid der doseringsletaliteitproeven voor, een feit dat merkwaardig genoeg is om er even bij te blijven stilstaan. Populair gezegd: Een toename van de dosis van 1 op 2 is voor een dier 'erger' dan een toename van 2 op 3 en ongeveer even 'erg' als een toename van 2 op 4. Vervangt men hierin 'dosis' door 'prikkelsterkte', dan heeft men een populair versie van de Wet van WEBER-FECHNER. Men heeft dan ook niet nagelaten om verband te zoeken tussen de genoemde groep van verschijnselen en de genoemde wet, welke ontiaan van zijn psycho-fysische speculaties slechts de samenvatting is van een tweede groep van verschijnselen: Het juist waarneembare verschil van de logarithmen van de prikkelsterkte is voor een bepaald dier (of mens) 'meer constant' dan het juist waarneembare verschil van de prikkelsterkten zelf. Op het eerste gezicht is de inhoud van de beide groepen van verschijnselen wel zeer verschillend. De eerste zijn statistisch en hebben betrekking op de morbiditeit in populaties van dieren. De laatste zijn volstrekt onstatistisch en hebben betrekking op het perceptievermogen van een bepaald dier.

Er bestaat echter een derde groep van verschijnselen, die min of meer een schakel vormt tussen de beide vorige: De logarithmen van de drempelwaarden der sensorische gevoeligheid in een populatie van personen is 'normaler' gedistribueerd dan de drempelwaarden zelf. Deze derde groep blijft (evenals trouwens de Wet van W.F.) niet beperkt tot het perceptievermogen. Men kan zelfs zeggen dat *in den regel* in de biologie de logarithmen 'normaler' gedistribueerd zijn dan de 'gewone' waarden. Wellicht zou men tegen de laatste bewering willen aanvoeren dat vele 'gewone' maten dan toch maar heel behoorlijk normaal verdeeld zijn. Men bedenke evenwel dat het verschil tussen de arithmetische en de logarithmische schaal alleen merkbaar is, als de variatiebreedte groot is t.o.v. de gemiddelde waarde¹. . . . De m.i. beste samenvatting van alle bovengenoemde verschijnselen werd in 1878 door GALTON gegeven: In de biologie treden de logarithmische progressies meer op de voorgrond dan de arithmetische. Dat dit vooral opvalt bij de perceptie en de gevoeligheid tegenover giften, komt omdat daar de variatiebreedte relatief zeer groot is. De maximale perceptie van een individu is vele malen groter dan de minimale; en de drempelwaarde van de ongevoeligste individuen uit een populatie is enige malen groter dan die van de gevoeligste individuen.'

Een veel gebruikte dosis-schaal is die, waarbij elke volgende dosis verkregen wordt door de voorgaande met 2 te vermenigvuldigen, dus bv. 1, 2, 4, 8, Het verschil tussen de logarithmen van twee opeenvolgende doses is dan $c = 0,3010$. Als deze schaal te grof is kan men bv. met de factor $\sqrt{2}$ vermenig-

¹ Zie blz. 118.

vuldigen en dan is $c = 0,1505$, zoals bij de schaal: 4,25, 6, 8,5, 12, 17, 24, 34, 48, 68. Een bekende schaal met $c = 0,1761$ vormt de reeks 3,6, 5,3, 8, 12, 27, 40,5 en 60,8.

Wij volstaan met het geven van een eenvoudig voorbeeld van een directe ijking. Tabel 23.1 bevat de uitkomsten van een proef, waarbij twee strophantus preparaten, een standaard S en een te toetsen preparaat T , op ononderbroken wijze met constante snelheid intraveneus bij katten worden ingespoten tot op het ogenblik, dat het hart stilstaat.¹ De letale dosis, d.i. dus de totale hoeveelheid ingespoten stof per dier, is uitgedrukt in 0,01 cc per kg lichaamsgewicht.²

Tabel 23.1. Letale doses strophantus bij katten

	Strophantus S (Standaard)		Strophantus T	
	d_s	$x = \log d_s$	d_T	$x = \log d_T$
	1,55	0,190	2,42	0,384
	1,58	0,199	1,85	0,267
	1,71	0,233	2,00	0,301
	1,44	0,158	2,27	0,356
	1,24	0,093	1,70	0,230
	1,89	0,276	1,47	0,167
	2,34	0,369	2,20	0,342
Som	11,75	1,518	13,91	2,047
Gemiddelde	1,68	0,217	1,99	0,292
Kwadratensom		0,0467		0,0352

Uit: FINNEY (193), gegevens van BURN c.s.

Bij de statistische analyse maken wij gebruik van de logarithmen x der individueel-effectieve doses. Men vindt dan dat $\bar{x}_S = 0,217$ en dat $\bar{x}_T = 0,292$ en de activiteitsverhouding van preparaat T t.o.v. de standaard S kan nu gevonden worden door eerst te berekenen

$$(23.1) \quad \begin{aligned} M &= \bar{x}_T - \bar{x}_S \\ &= 0,292 - 0,217 = 0,075 \end{aligned}$$

en door vervolgens te nemen

$$(23.2) \quad \begin{aligned} av &= \text{antilog } M \\ &= \text{antilog } 0,075 = 1,19. \end{aligned}$$

Als wordt aangenomen dat voor de populatie, waaruit de proefdieren stammen, de verdelingen van de logarithmen der individueel-effectieve doses (x) voor het testpreparaat en voor het standaardpreparaat onge-

¹ Het spreekt vanzelf, dat hierbij aan farmacologische eisen — betreffende het totaal te injecteren volume en de tijdsduur van de proef — dient te worden voldaan.

² Het lichaamsgewicht kan ook (en soms beter!) worden uitgeschakeld door een zg. covariantie-analyse toe te passen. Deze techniek laten wij echter buiten beschouwing.

23.2 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

veer normaal zijn, resp. met gemiddelden μ_T en μ_S , maar met dezelfde variantie σ^2 , dan volgt de grootheid $\underline{M} = \bar{x}_T - \bar{x}_S$ eveneens een bij benadering normale verdeling. Deze heeft als gemiddelde $\mu_T - \mu_S$ en als variantie $\sigma_M^2 = \sigma^2 \left[\frac{1}{n_T} + \frac{1}{n_S} \right]$. Uit de waargenomen x -waarden kan men dan op de gebruikelijke wijze een zuivere schatting van σ^2 verkrijgen door te nemen

$$s^2 = \frac{(n_T-1) s_T^2 + (n_S-1) s_S^2}{n_1 + n_2 - 2},$$

zodat

$$(23.3) \quad s_M^2 = s^2 \left[\frac{1}{n_T} + \frac{1}{n_S} \right].$$

Daar de grootheid $t = \underline{M}/s_M$ volgens formule (13.32) bij benadering een STUDENT-verdeling met $n_T + n_S - 2$ vrijheidsgraden volgt, kan men een $100(1-\alpha)\%$ betrouwbaarheidsinterval van $\mu_T - \mu_S$ verkrijgen met

$$(23.4) \quad M + t_{\frac{1}{2}\alpha} s_M < \mu_T - \mu_S < M + t_{1-\frac{1}{2}\alpha} s_M,$$

waarin $t_{\frac{1}{2}\alpha} = -t_{1-\frac{1}{2}\alpha}$ met $\nu = n_T + n_S - 2$ uit tabel C volgt. Door de antilogarithmen van de verkregen grenswaarden te nemen verkrijgt men een $100(1-\alpha)\%$ betrouwbaarheidsinterval van de werkelijke activiteitsverhouding, AV .

Bij het voorbeeld in tabel 23.1 verschillen de waargenomen varianties niet significant [$F_0 = 0,0467/0,0352 = 1,33$, $F_{0,95}(6, 6) = 4,28$], zodat de populatievarianties gelijk mogen worden ondersteld. Dan is

$$s^2 = \frac{0,0467 + 0,0352}{6 + 6} = 0,006825,$$

zodat uit (23.3) volgt

$$s_M^2 = 0,006825 \left(\frac{1}{7} + \frac{1}{7} \right) = 0,00195 = (0,0442)^2.$$

Onderstelt men tevens normaliteit, dan volgt uit (23.4) als 95% betrouwbaarheidsinterval van $\mu_T - \mu_S$:

$$\begin{aligned} 0,075 - (2,179) (0,0442) &< \mu_T - \mu_S < 0,075 + (2,179) (0,0442), \\ -0,021 &< \mu_T - \mu_S < 0,171, \\ 0,979 - 1 &< \mu_T - \mu_S < 0,171. \end{aligned}$$

Daar antilog $(0,979 - 1) = 0,95$ en antilog $(0,171) = 1,48$ vindt men dus als 95% betrouwbaarheidsinterval van de ware activiteitsverhouding:

$$0,95 < AV < 1,48.$$

23.3. Indirecte standaardisatie met kwantitatieve responsies¹

23.3.1. INLEIDING

In vele gevallen kan de directe methode niet worden gebruikt, omdat het praktisch onmogelijk is om per proefdier de effectieve dosis vast te stellen: zo kan men bv. niet nagaan, welke dosis van een insecticide voor één bepaald insect dodelijk is. Men dient dan over te gaan tot een methode van indirecte standaardisatie en wij bespreken eerst de situatie, waarin de responsie kwantitatief is (bv. een uterusgewicht, een asgehalte, een bloedsuikergehalte, e.d.).

Bij deze indirecte standaardisatie maakt men per preparaat gebruik van een aselechte steekproef van proefdieren (uit een bepaalde populatie), die — eveneens aselechte — in k groepen van de omvang $n_1, n_2, \dots, n_j, \dots, n_k$ wordt gesplitst. Gewoonlijk zal men, als de responsie kwantitatief is, alle groepen dezelfde omvang n geven (als de responsie echter bestaat uit het al of niet reageren van de dieren, het geval dat in de volgende paragraaf besproken wordt, is dit niet altijd zo). Deze groepen worden dan aan k opklimmende doseringen blootgesteld, zodat men bij de dosis j over n_j waarden van de responsie y beschikt en de regressie van de responsie op de dosis kan onderzoeken. Hierbij kunnen zich in principe twee situaties voordoen:

1. '*Parallel line assay*' — Er is een — gewoonlijk lineaire — regressie van de responsie y (of een daaruit door een of andere transformatie volgende grootte) op $x = \log$ dosis. Indien geen storende factoren optreden, zullen in dit geval de regressielijnen van y op x voor het testpreparaat en het standaardpreparaat evenwijdig zijn, zodat de activiteitsverhouding van deze preparaten kan worden geschat uit de afstand van deze lijnen, gemeten evenwijdig aan de X -as.

2. '*Slope ratio assay*' — In het bijzonder bij microbiologische standaardisatie komt het voor, dat de responsie een lineaire functie is van de dosis. Verschillen dan twee preparaten in activiteit, dan vindt men voor de regressie van de responsie y op de dosis d twee regressielijnen met een verschillende helling, die convergeren in $x = 0$. De activiteitsverhouding der preparaten volgt dan uit het quotiënt van de richtingscoëfficiënten (b) van deze lijnen.

FINNEY geeft in zijn boek '*Statistical method in biological assay*' een volledige systematische uiteenzetting van de verschillende proefopzetten, die bij de kwantitatieve indirecte standaardisatie kunnen worden gebruikt en van de daarbij behorende methoden van statistische analyse (ook laat hij aan de hand van voorbeelden zien, welke nadelen aan een niet goed doordachte proefopzet verbonden zijn). Na een inleiding (hoofdstuk 3) behandelt hij de '*parallel line assays*' in de

¹ JERNE en WOOD (223) analyseren kritisch alle schakels van 'the long chain of reasoning and inference from the performance of a biological assay to the statement of its results'. Men verzuieme niet dit artikel te lezen!

23.3 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

hoofdstukken 4 t/m 6, de 'slope ratio assays' in de hoofdstukken 7 en 8 en verschillende meer complexe proefopzetten in de hoofdstukken 9 t/m 14. Deze twaalf hoofdstukken beslaan reeds meer dan 400 bladzijden en wij dienen ons in deze inleiding dan ook te beperken tot het geven van enkele voorbeelden, die bedoeld zijn om de lezer een indruk te geven van de wijze waarop de statistische analyse verloopt bij een eenvoudige, maar goed opgezette ijkingsproef.

23.3.2. EEN 'PARALLEL LINE ASSAY'

Het voorbeeld dat wij bespreken betreft een zg. symmetrische $2k$ -punts proef, waarbij een testpreparaat van testosteroone-propionaat (T) wordt vergeleken met een standaardpreparaat (S) en wel bij $k = 3$ identieke doseringen met $n = 5$ responsies per dosis. De responsie is de toeneming in de groei van de kam (lengte + breedte) van hanen in mm. en de waargenomen uitkomsten zijn in tabel 23.2 opgenomen.

Tabel 23.2. Lengttoeneming (lengte + breedte) van de hanekam na injectie van testosteroone-propionaat

Dosis (d) x	Standaardpreparaat			Testpreparaat		
	20 μ g -1	40 μ g 0	80 μ g +1	20 μ g -1	40 μ g 0	80 μ g +1
Responsie = y	7	12	18	6	12	17
	6	11	15	8	13	16
	6	10	14	6	9	18
	4	7	15	7	12	19
	5	10	14	5	10	15
Som	28	50	76	32	56	85
Gemiddelde	5,6	10,0	15,2	6,4	11,2	17,0
Kwadratensom	5,2	14,0	10,8	5,2	10,8	10,0

De regressie van de responsie op de logaritmen van de dosis blijkt ongeveer rechtlijnig te zijn (zie figuur 23.1). Om de berekeningen te vereenvoudigen kan men bij deze gegevens het best logaritmen met grondtal 2 gebruiken en deze met $\log_2 40$ verminderen, dus nemen

$$x = \log_2 \text{dosis} - \log_2 40 = \log_2 \text{dosis} - 5,3219.$$

De waarden van x worden dan resp. -1 , 0 en $+1$.

Bij de analyse van de uitkomsten dient in de eerste plaats te worden ondersteld dat de responsies (per preparaat/per dosis) ongeveer normaal verdeeld zijn met dezelfde variantie. De normaliteitsonderstelling lijkt wel gerechtvaardigd te zijn; de waarden $x_{ij} - \bar{x}_j$ vertonen geen duidelijke afwijkingen van de normaliteit. De toets van HARTLEY laat zien dat de verschillen tussen de varianties ver van significant

zijn, daar $F_{\max} = 14,0/5,2 = 2,69$ en $F_{0,95} (k = 6, n = 5) = 29,5$. Voor een valide berekening van de activiteitsverhouding — op de wijze die wij even verder bespreken — dient nog aan drie andere voorwaarden voldaan te zijn, t.w.: *a.* Voor elk preparaat dient een uitgesproken regressie van y (responsie) op x (log dosis) aanwezig te zijn, *b.* De beide regressies moeten rechtlijnig zijn, *c.* De twee regressielijnen moeten evenwijdig zijn. Door middel van een variantie-analyse kan men onderzoeken of aan deze voorwaarden voldaan is en wij zullen de uitvoering van deze analyse eerst beknopt bespreken. De berekeningen verlopen als volgt:

$$\Sigma y = 327, \quad \Sigma y^2 = (7)^2 + (12)^2 + (18)^2 + (6)^2 + \dots + (15)^2 = 4145, \\ S_{\text{totaal}} = 4145 - (327)^2/30 = 580,7.$$

De kwadratensom tussen de zes groepen is

$$S_{\text{tussen}} = S_t = \frac{(28)^2 + (50)^2 + \dots + (85)^2}{5} - \frac{(327)^2}{30} = 524,7,$$

zodat

$$S_{\text{binnen}} = S_b = 580,7 - 524,7 = 56,0,$$

een uitkomst die eveneens verkregen wordt als men de zes afzonderlijke kwadratensommen optelt (zie tabel 23.2). De eerste stap van de analyse levert dus op

$$S_t = 524,7 [2k - 1 = 5 \text{ vrijheidsgraden}], \\ S_b = 56,0 [2k(n-1) = 24 \text{ vrijheidsgraden}].$$

Vervolgens kan S_t in verschillende componenten worden gesplitst. In de eerste plaats is er de kwadratensom tussen de twee preparaten (met $2 - 1 = 1$ vrijheidsgraad):

$$S_{\text{prep}} = S_p = \frac{(28 + 50 + 76)^2 + (32 + 56 + 85)^2}{15} - \frac{(327)^2}{30} = 12,0333.$$

Ten tweede kan men de kwadratensom voor de regressie (voor de gecombineerde gegevens van de twee preparaten) van y op x berekenen en wel uit de volgende gegevens (zie weer tabel 23.2):

x	-1	0	+1
Σy	28 + 32 = 60	50 + 56 = 106	76 + 85 = 161
Σxy	-60	0	+161

Hieruit volgt dat $\Sigma xy = -60 + 161 = 101$ en daar $\Sigma x = 10(-1) + 10(0) + 10(1) = 0$ en $\Sigma x^2 = 10(-1)^2 + 10(0)^2 + 10(1)^2 = 20$, is $\Sigma (x - \bar{x})^2 = 20$ en is

$$S_{\text{regressie}} = S_{\text{reg}} = \frac{(101)^2}{20} = 510,05.$$

23.3 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

Wanneer er per preparaat k groepen zijn, bezit deze kwadratensom $k - 2$ vrijheidsgraden. Bij het voorbeeld is $k = 3$, zodat S_{rgr} één vrijheidsgraad heeft.

Op overeenkomstige wijze vindt men voor de regressie van y op x bij het standaardpreparaat, dat $\sum xy = (-1) 28 + (0) 50 + (1) 76 = 76 - 28 = 48$ en dat $\sum (x_S - \bar{x}_S)^2 = 10$, zodat

$$S_{\text{regressie (standaard)}} = S_{rgr(S)} = \frac{(48)^2}{10} = 230,4.$$

De lezer kan zelf vaststellen, dat voor het testpreparaat geldt

$$S_{\text{regressie (test)}} = S_{rgr(T)} = \frac{(53)^2}{10} = 280,9.$$

Wij vermelden nu zonder hierop nader in te gaan, dat de kwadraten-som voor de evenwijdigheid van de twee regressielijnen (met één vrijheidsgraad) kan worden gevonden door te nemen

$$S_{\text{evenwijdigheid}} = S_e = [S_{rgr(S)} + S_{rgr(T)}] - S_{rgr},$$

zodat men vindt

$$S_e = (230,4 + 280,9) - 510,05 = 1,25.$$

Van de $2k - 1 = 5$ vrijheidsgraden van S_t zijn nu achtereenvolgens afgesplitst: $2 - 1 = 1$ vrijheidsgraad voor S_p (preparaten), $k - 2 = 1$ vrijheidsgraad voor S_{rgr} (gecombineerde regressie) en 1 vrijheidsgraad voor S_e (evenwijdigheid). De resterende $k - 1 = 2$ vrijheidsgraden betreffen de kwadratensom voor de rechtlijnigheid van de regressie die kan worden berekend door te nemen

$$S_{\text{rechtlijnigheid}} = S_r = S_t - (S_p + S_{rgr} + S_e),$$

zodat

$$S_r = 524,7 - (12,0333 + 510,05 + 1,25) = 1,3667.$$

De lezer ziet, dat de berekeningen — mede door de gunstige transformatie van de doses — betrekkelijk eenvoudig zijn. De uitkomsten zijn in tabel 23.3 opgenomen, tezamen met de aantallen vrijheidsgraden die in het algemeen gelden bij een symmetrische $2k$ -punts proef met n responsies per dosis.

Men kan nu de volgende hypothesen toetsen:

a. De hypothese dat de waargenomen regressie van y op x op toeval berust. De toetsingsgrootheid is $F = s_{rgr}^2 / s_e^2$ met $\nu_1 = k - 2$ en $\nu_2 = 2k(n - 1)$. Het is duidelijk, dat de regressie op z 'n minst significant dient te zijn, indien men de responsie-dosis relatie wil gebruiken voor het schatten van de activiteitsverhouding der preparaten. Bij een goede ijkingsproef behoort zelfs de regressie-kwadratensom relatief (zeer) groot te zijn, zoals bij het voorbeeld het geval is.

Tabel 23.3. Uitkomsten van de analyse van de gegevens in tabel 23.2 (symmetrische $2k$ -puntsproef, $k = 3$, $n = 5$)

	S	Vrijheidsgraden		Var.schatting
Preparaten	12,0333	1	1	12,0333
Regressie	510,0500	$k-2$	1	510,0500
Evenwijdigheid	1,2500	1	1	1,2500
Rechtlijnigheid	1,3667	$k-1$	2	0,6833
Tussen de groepen	524,7000	$2k-1$	5	
Binnen de groepen	56,0000	$2k(n-1)$	24	2,3333
Totaal	580,7000	$2kn-1$	29	

b. De hypothese dat de regressie van y op x rechtlijnig is. De toetsingsgroottheid is $F = s^2_r/s_b^2$ met $\nu_1 = 1$ en $\nu_2 = 2k(n-1)$. Bij het voorbeeld is $F_0 = 0,6833/2,3333 = 0,293$. Daar F_0 kleiner is dan 1, wordt de hypothese dat de regressie lineair is dus niet verworpen.

Wanneer men tot het verwerpen van deze hypothese moet overgaan, kan de analyse niet zonder meer worden voortgezet. Uit een inspectie van het spreidingsdiagram kan dan blijken, van welke aard de afwijkingen van de rechtlijnigheid zijn en men kan dan nagaan, of zij door middel van een transformatie op de responsie kunnen worden opgeheven.

c. De hypothese dat de regressielijnen van y op x voor het standaardpreparaat en het testpreparaat evenwijdig zijn. De toetsingsgroottheid is $F = s_a^2/s_b^2$ met $\nu_1 = k-1$ en $\nu_2 = 2k(n-1)$. Bij het voorbeeld is F_0 kleiner dan 1, zodat het verantwoord is de regressielijnen evenwijdig te onderstellen.

De evenwijdigheid van de regressielijnen is een zo fundamentele voorwaarde voor de geldigheid van de schatting van de activiteitsverhouding, dat men er goed aan zal doen bij deze toetsing een betrekkelijk hoge onbetrouwbaarheidsdrempel (bv. $\alpha = 0,10$) te kiezen.

Na deze voorbereidende analyse, waarbij geen evidente afwijkingen van het model te voorschijn zijn gekomen, kunnen wij overgaan tot het schatten van de activiteitsverhouding. Daartoe dient eerst de regressiecoëfficiënt van y op x voor de gecombineerde gegevens van de twee preparaten te worden berekend. Wij beschikken reeds over $\Sigma(x-\bar{x})^2 = 20$ en $\Sigma(x-\bar{x})(y-\bar{y}) = 101$, zodat $b_{yx} = 101/20 = 5,05$. De regressievergelijking voor het standaardpreparaat luidt dus (bedenk dat $\bar{x} = 0$)

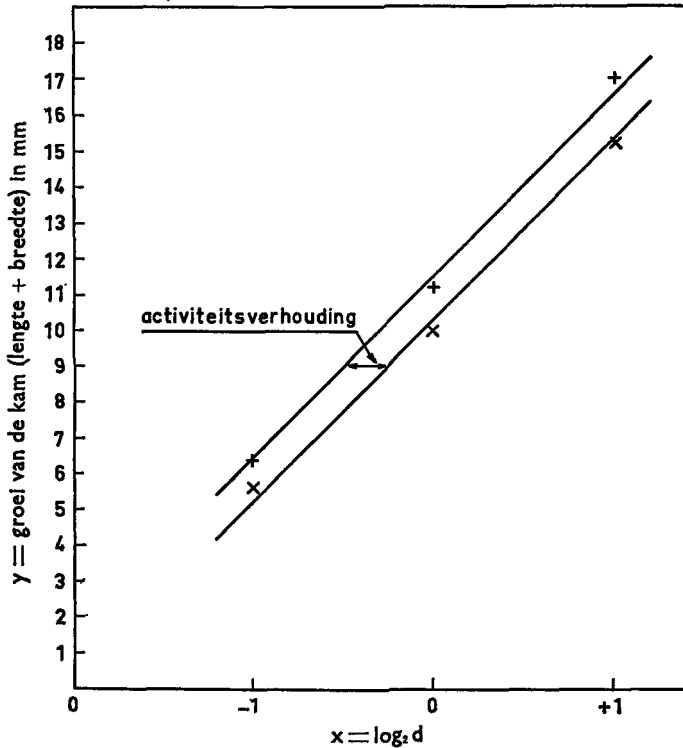
$$\begin{aligned} \hat{y}_S &= \bar{y}_S + b_{yx}(x - \bar{x}) \\ &= 10,27 + 5,05x \end{aligned}$$

en voor het testpreparaat geldt

$$\begin{aligned} \hat{y}_T &= \bar{y}_T + b_{yx}(x - \bar{x}) \\ &= 11,53 + 5,05x. \end{aligned}$$

23.3 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

Deze vergelijkingen zelf zijn voor de verdere berekeningen niet nodig, maar zij dienen voor het tekenen van de twee regressielijnen in het spreidingsdiagram, dat in figuur 23.1 is gegeven.



Figuur 23.1. Regressie van de responsie op de logaritmie van de dosis voor de gegevens in tabel 23.2:

x = gemiddelde responsie bij de standaard,
+ = gemiddelde responsie bij het testpreparaat.

De activiteitsverhouding wordt in deze figuur gerepresenteerd door de horizontale afstand tussen de twee lijnen, positief gerekend als het testpreparaat een grotere en negatief, als het een kleinere activiteit bezit dan het standaardpreparaat. Dit verschil is in het algemeen

$$(23.5.1) \quad M = \bar{x}_S - \bar{x}_T - \frac{\bar{y}_S - \bar{y}_T}{b_{yx}}$$

en bij de als voorbeeld gebruikte 2k-punts proef met $\bar{x}_S = \bar{x}_T$

$$(23.5.2) \quad M = \frac{\bar{y}_T - \bar{y}_S}{b_{yx}} \\ = \frac{11,53 - 10,27}{5,05} = 0,2495.$$

De activiteitsverhouding volgt dan, evenals in 23.2, door te nemen

$$(23.6) \quad av = \text{antilog } M.$$

Wij dienen echter te bedenken, dat bij het voorbeeld logaritmen met grondtal 2 zijn gebruikt. Deze moeilijkheid kan worden opgelost door M met $\log_{10} 2$ te vermenigvuldigen (zie blz. 325), waarna geldt dat

$$\begin{aligned} av &= \text{antilog}_{10} (M) (\log_{10} 2) = \text{antilog}_{10} (0,2495) (0,3010) \\ &= \text{antilog}_{10} 0,07510, \end{aligned}$$

zodat $av = 1,19$. Hieruit volgt, dat $1\mu\text{g}$ van het testpreparaat equivalent is aan $1,19\mu\text{g}$ van het standaardpreparaat.

Voor de variantie van \underline{M} geldt in het algemeen bij benadering

$$(23.7.1) \quad s_M^2 = \frac{s_b^2}{b^2} \left[\frac{1}{n_S} + \frac{1}{n_T} + \frac{[M - (\bar{x}_S - \bar{x}_T)]^2}{\Sigma (x - \bar{x})^2} \right],$$

waarin n_S = het totale aantal waarnemingen bij de standaard en n_T = het totale aantal waarnemingen bij het testpreparaat en waarin M volgt uit (23.5.1). Bij de als voorbeeld gegeven symmetrische $2k$ -punts proef is $n_S = n_T = kn$ en is $\bar{x}_S = \bar{x}_T$, zodat daarbij (23.7.1) kan worden vereenvoudigd tot

$$(23.7.2) \quad s_M^2 = \frac{s_b^2}{b^2} \left[\frac{2}{kn} + \frac{M^2}{\Sigma (x - \bar{x})^2} \right],$$

waarin M volgt uit (23.5.2). Met deze formule vindt men:

$$s_M^2 = \frac{2,3333}{(5,05)^2} \left[\frac{2}{15} + \frac{(0,2495)^2}{20} \right] = \frac{2,3333 \times 0,13446}{25,5025} = 0,0124839 = (0,112)^2.$$

Men kan dus een benaderend $100(1-\alpha)\%$ betrouwbaarheidsinterval van $\log(AV)$ verkrijgen met

$$(23.8) \quad M + t_{\frac{1}{2}\alpha} s_M < \log AV < M + t_{1-\frac{1}{2}\alpha} s_M,$$

waarin $t_{\frac{1}{2}\alpha} = -t_{1-\frac{1}{2}\alpha}$, met $2k(n-1)$ vrijheidsgraden in tabel C kan worden opgezocht. Voor het voorbeeld vindt men dan als 95% betrouwbaarheidsinterval

$$\begin{aligned} 0,2495 - (2,064) (0,112) &< \log_2 AV < 0,2495 + (2,064) (0,112), \\ 0,01833 &< \log_2 AV < 0,4807. \end{aligned}$$

Ook deze grenswaarden kunnen in logaritmen met grondtal 10 worden omgezet door ze met $\log_{10} 2 = 0,3010$ te vermenigvuldigen en daarna vindt men door het nemen van de antilogaritmen als benaderende 95% betrouwbaarheidsgrenzen van de activiteitsverhouding $1,01$ en $1,40\mu\text{g}$.

23.3 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

Tenslotte enkele opmerkingen omtrent de opzet van een 'parallel line assay', waarbij wij ons wederom beperken tot de $2k$ -punts proef. Uit formule (23.7.2) blijkt, dat men een nauwkeuriger schatting van de activiteitsverhouding (een kleinere waarde van s_M) verkrijgt naarmate:

1. Het aantal waarnemingen ($2kn$) groter is. Daar men (zie punt 2) gewoonlijk $k = 3, 4$ of 5 kiest, betekent dit dat men n zo groot mogelijk dient te maken.

2. $\Sigma(x - \bar{x})^2$ groter is. Men kan bij een gegeven aantal waarnemingen deze kwadratensom zo groot mogelijk maken, door $k = 2$ te kiezen met één lage en één hoge dosis, maar deze proefopzet laat geen toets op de rechtlijnigheid van de regressie toe, zodat men gewoonlijk $k = 3, 4$ of 5 kiest. Een groter aantal doses is slechts aan te bevelen, wanneer men een voorbereidend onderzoek uitvoert teneinde de vorm van de regressie te bestuderen.

3. s_b kleiner is. De waarde van s_b hangt af van de homogeniteit van de proefdieren, de zorg waarmee de proef wordt uitgevoerd en de keuze van een goede proefopzet.

4. b groter is. De waarde van b representeert de gemiddelde toename in responsie per eenheid toename van de log dosis en is dus een maatstaf voor de gevoeligheid van de dieren voor veranderingen in de dosering. Vergroting van deze gevoeligheid leidt dus tot verbetering van de ijking.

5. M kleiner is, d.i. dus (zie formule 23.5.2) naarmate \bar{y}_T en \bar{y}_S minder verschillen. Is $\bar{y}_T = \bar{y}_S$ dan is $M = 0$.

23.3.3. EEN 'SLOPE RATIO ASSAY'

Wij bespreken een voorbeeld van een zg. symmetrische ($2k + 1$)-punts 'slope ratio assay', waarvan de uitkomsten in tabel 23.4 zijn opgenomen. Het betreft een onderzoek naar het riboflavinegehalte van mout met *Lactobacillus helveticus* als test-organisme, waarbij de responsie werd gemeten als titer in ml $0,1 N NaOH$. Daar van elk preparaat $k = 2$ doses zijn gekozen (met één blanco proef) hebben wij hier te doen met een 5-punts proef.

Evenals bij een 'parallel line assay' dient men eerst de juistheid van de verschillende te maken onderstellingen te onderzoeken. In de eerste plaats dient men na te gaan, of de responsies per preparaat/per dosis ongeveer normaal verdeeld zijn met dezelfde variantie. De normaliteitsonderstelling is bij dit kleine aantal waarnemingen niet bevredigend te verifiëren, maar de varianties verschillen niet significant (de kwadratensommen zijn resp. 0,0819, 0,0950, 0,1825, 0,1700 en 0,0119, zodat $F_{\max} = 15,3$, terwijl $F_{0,95}(k = 5, n = 4) = 50,7$). Verder dient voldaan te zijn aan de volgende drie voorwaarden: a. De regressie van y op x dient voor beide preparaten zeer

Tabel 23.4. Symmetrische 5-punts proef voor de bepaling van het riboflavinegehalte van mout:

Standaard preparaat: 1 eenheid = 0,1µg riboflavine,

Test preparaat: 1 eenheid = 0,025 gram mout.

Dosis	0	Standaard		Testpreparaat		
		1	2	1	2	
	1,90	4,85	8,35	4,00	6,05	$k = \text{aantal doses} = 2$
	2,25	5,00	8,20	4,40	6,20	$n = \text{aantal buisjes} = 4$
	2,00	5,25	7,95	4,50	6,10	$m = n(2k+1) = 20$
	2,20	4,90	7,80	4,10	6,10	
Σy_i	8,35	20,00	32,30	17,00	24,45	$\Sigma y = 102,10$
Σxy		20,00	64,60	17,00	48,90	$T = 150,50$
		$T_S = 84,60$		$T_T = 65,90$		

Gegevens van Wood, Analyst 71 (1946), 1.

significant te zijn, *b*. De beide regressies moeten rechtlijnig zijn, *c*. Het snijpunt van de twee regressielijnen moet in overeenstemming zijn met het gemiddelde van de blanco waarnemingen. Men kan ook hier door middel van een variantie-analyse onderzoeken, of aan deze voorwaarden voldaan is (het onderzoek naar de rechtlijnigheid kan alleen worden uitgevoerd, als $k > 2$ is en kan dus bij het voorbeeld niet plaatsvinden). Wij geven de uitvoering van deze analyse niet, maar een inspectie van figuur 23.2, waarin de waargenomen punten (*x*, *y*) zijn getekend laat zien, dat de voorwaarden (*a*) en (*c*) bij dit voorbeeld geen moeilijkheden opleveren.

Wij bespreken nu in 't kort de wijze waarop men bij een $(2k + 1)$ -punts proef de activiteitsverhouding kan berekenen. In het algemene geval, waarbij de proef is uitgevoerd met *h* blanco waarnemingen, f_S waarnemingen per dosis van het standaardpreparaat en f_T waarnemingen per dosis van het testpreparaat, dus met in totaal $h + k(f_S + f_T) = m$ waarnemingen, kan men de gemiddelde responsie bij $x = 0$ schatten door te berekenen

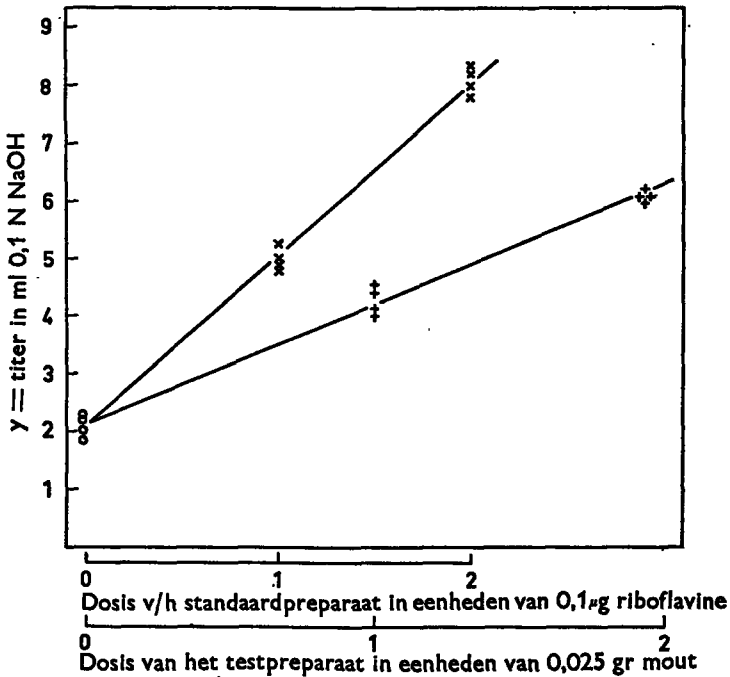
$$(23.9.1) \quad a = \frac{2(2k + 1) \Sigma y - 6T}{m(k - 1) + 3h(k + 1)}$$

Als men een symmetrische proef uitvoert, zoals in het voorbeeld, dan is $h = f_S = f_T = n$, zodat $m = n(2k + 1)$ en

$$(23.9.2) \quad a = \frac{2(2k + 1) \Sigma y - 6T}{n(2k + 1)(k - 1) + 3n(k + 1)}$$

De betekenis van Σy en *T* blijkt uit tabel 23.4.

23.3 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE



Figuur 23.2. Regressie van de responsie op de dosis voor de gegevens in tabel 23.4:
 × = responsies bij de standaard,
 + = responsies bij het testpreparaat,
 0 = blanco's.

De richtingscoëfficiënt van de regressielijn van de responsie (y) op de dosis (x) voor het standaardpreparaat volgt uit

$$(23.10) \quad b_S = \frac{3}{2k + 1} \left[\frac{2T_S}{f_S k (k + 1)} - a \right]$$

(zie voor T_S tabel 23.4). De regressiecoëfficiënt b_T voor het testpreparaat kan worden berekend door in (23.10) f_S te vervangen door f_T en T_S door T_T . Bij de symmetrische opzet is $f_S = f_T = n$.

Voor het voorbeeld in tabel 23.4 vindt men met (23.9.2)

$$a = \frac{2(2 \times 2 + 1)(102,10) - 6(150,50)}{4(2 \times 2 + 1)(2 - 1) + 3(4)(2 + 1)} = \frac{118}{56} = 2,10714$$

en met (23.10)

$$b_S = \frac{3}{5} \left[\frac{2 \times 84,60}{4 \times 2 \times 3} - 2,1071 \right] = 2,9657$$

en

$$b_T = \frac{3}{5} \left[\frac{2 \times 65,90}{4 \times 2 \times 3} - 2,1071 \right] = 2,0307.$$

De regressievergelijking van y op x voor het standaardpreparaat luidt dus $\hat{y}_S = 2,1071 + 2,9657 x_S$ en die voor het testpreparaat $\hat{y}_T = 2,1071 + 2,0307 x_T$. Met deze vergelijkingen zijn de regressielijnen in figuur 23.2 getekend.

De activiteit van het testpreparaat t.o.v. het standaardpreparaat is nu, uitgedrukt in de gecodeerde doses

$$(23.11) \quad M = b_T/b_S$$

en deze activiteitsverhouding kan in de oorspronkelijke eenheden worden uitgedrukt door te nemen

$$(23.12) \quad av = \frac{c_S}{c_T} \cdot M = \frac{c_S b_T}{c_T b_S},$$

waarin $c_S =$ één eenheid van het standaardpreparaat en $c_T =$ één eenheid van het testpreparaat. Met deze formule vindt men als activiteitsverhouding bij het voorbeeld

$$av = \frac{0,1}{0,025} \cdot \frac{2,0307}{2,9657} = 4 (0,6847) = 2,74.$$

Op grond van de waargenomen activiteitsverhouding kan men een $100(1-\alpha)\%$ betrouwbaarheidsinterval van de werkelijke activiteitsverhouding verkrijgen door te nemen

$$(23.13) \quad \frac{c_S}{c_T} [M + t_{1-\alpha} s_M] < AV < \frac{c_S}{c_T} [M + t_{1-\alpha} s_M].$$

Wij volstaan hier met de vermelding, dat voor het symmetrische 5-punts schema van het voorbeeld, met als doses 1 en 2 eenheden, bij benadering geldt

$$(23.14) \quad s_M^2 = \frac{2s^2 (8 - 9M + 8M^2)}{7mb_S^2}$$

met

$$s^2 = \frac{\sum (n_j - 1) s_j^2}{\sum (n_j - 1)} = \frac{(n - 1) \sum s_j^2}{k(n - 1)} = \frac{\sum s_j^2}{k} \quad [v = k(n - 1)].$$

Daar voor het voorbeeld

$$s^2 = \frac{0,0819 + 0,0950 + 0,1825 + 0,1700 + 0,0119}{5 \times 3} = 0,0361,$$

vindt men met (23.14)

$$s_M^2 = \frac{2(0,0361) [8 - 9(0,6847) + 8(0,6847)^2]}{7(20) (2,9657)^2} = 0,0003277 = (0,0181)^2.$$

23.4 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

De 95% grenzen van de activiteitsverhouding zijn dus bij benadering

$$[t_{0,975} (\nu = 15) = 2,131, c_S/c_T = 0,1/0,025 = 4]:$$

$$4[0,6847 - (2,131) (0,0181)] < AV < 4[0,6847 + (2,131) (0,0181)], \\ 2,58 < AV < 2,89.$$

Met een betrouwbaarheid van 0,95 kan dus worden beweerd, dat het riboflavinegehalte van de mout tussen 2,58 en 2,89 μg per gram ligt.

23.4. Indirecte standaardisatie gebaseerd op het al dan niet reageren: ED_{50} bepalingen¹

23.4.1. INLEIDING

Bij sommige standaardisatie-proeven kan men van elk proefdier slechts vaststellen, of het al dan niet op een bepaalde wijze op de toegediende dosis reageert. Voert men dan een indirecte standaardisatie uit op de wijze, die in 23.3 beschreven is, zodat men (voor de standaard en het testpreparaat) k groepen proefdieren van de omvang n_1, n_2, \dots, n_k aan doseringen blootstelt, die logaritmisch met een constante opklimmen, en bepaalt men per dosis x_j het aantal dieren r_j , dat reageert, dan verkrijgt men per preparaat k fracties $p_j = r_j/n_j$. Uit deze fracties kan het verloop van de reacties bij opklimmende doses worden afgeleid en een schatting worden verkregen van de zg. *mediaan-effectieve dosis* of ED_{50} , d.i. de dosis waarbij 50% van de proefdieren nog juist reageert (als de waargenomen reactie de dood van het dier is, spreekt men wel van de mediaan-letale dosis of LD_{50}). De verhouding van de mediaan-effectieve doses van het testpreparaat en het standaardpreparaat verschafte dan een schatting van de activiteitsverhouding der preparaten.

Mén dient bij de analyse van het verloop van de fractie reacties bij opklimmende doses rekening te houden met het feit, dat de variantie van een schatting p_j (van de parameter P_j , de fractie reacties bij de j -de dosis in de populatie waaruit de proefdieren stammen) gelijk is aan $\sigma_p^2 = P_j(1-P_j)/n_j$, en dus niet alleen afhangt van de omvang van de groep proefdieren, maar tevens van de waarde van de parameter P_j — dus indirect van de dosis. Men kan dientengevolge niet rechtstreeks de regressievergelijking van p op x , berekend volgens de methode der kleinste kwadraten, gebruiken, daar men dan de waarde van de grootte $\sum n_j (p_j - \hat{p})^2$ minimaal zou maken, zonder rekening te houden met het feit, dat de variabiliteit van p groter is in de buurt van 0,5 dan bij lage of hoge waarden van P . Deze moeilijkheid kan worden opgeheven door gebruik te maken van de zg. *probit-analyse*, van de hierop gebaseerde grafische methode van LITCHFIELD en WILCOXON, of door toepassing van de bgsin transformatie. Deze methoden vereisen echter

¹ Bij alle berekeningen in deze paragraaf onderstellen wij weer, dat zij met $x = \log$ dosis worden uitgevoerd, zodat een schatting $m = \log ED_{50}$ wordt verkregen en $ED_{50} = \text{antilog } m$ is.

vrij veel rekenwerk, resp. het gebruik van speciale tabellen of nomogrammen, zodat wij er slechts een korte beschouwing aan wijden, die bedoeld is de lezer ter zake te oriënteren. Verder zijn er een groot aantal betrekkelijk eenvoudige methoden, waarvan wij de methode van SPEARMAN en KÄRBER en de methode der voortschrijdende gemiddelden van THOMPSON bespreken.¹ Tenslotte volgt dan nog de sequente methode van DIXON en MOOD.

23.4.2. DE PROBIT-ANALYSE

Deze methode berust op twee onderstellingen, t.w. (a) dat voor elk proefdier een bepaalde individuele dosis bestaat, waarop het nog juist reageert en (b) dat voor een homogene populatie van proefdieren de logaritmische x van deze individueel-effectieve dosis (ongeveer) normaal verdeeld is. Dan geldt, dat $m = \log ED_{50} =$ de mediaan = het gemiddelde van deze verdeling. De cumulatieve verdeling van x is dan (bij benadering) een cumulatieve normale verdeling, d.i. een S-vormige kromme of sigmoïde, die in een rechte lijn kan worden omgezet door de percentages p_j te vervangen door zg. *probits*. Dit zijn waarden van de standaardnormaal verdeelde grootte \underline{T} , waarvoor geldt dat $P_L(T_j) = p_j$, en die om negatieve waarden te vermijden vermeerderd worden met 5, zodat dus $\text{probit}(p_j) = T_j + 5$. Er zijn verschillende tabellen vervaardigd, waarmee men fracties of percentages in probits kan omzetten. Beschikt men nu over een reeks probits (y), die berusten op k waargenomen fracties, dan kan hieruit na toepassing van zekere correcties en langs iteratieve weg de regressievergelijking van y op x (= log dosis) worden berekend met een zg. gewogen kleinste kwadraten methode. Men vindt dan een regressievergelijking van de vorm $\hat{y} = \bar{y} + b(x - \bar{x})$ en daar een probitwaarde van 5 overeenkomt met 50% reacties geldt, dat $m = \log ED_{50} = (5 - \bar{y})/b + \bar{x}$.

De volledige berekeningsmethode is vrij bewerkelijk en omvat tevens de toetsing van de rechtlijnigheid van de regressie en bij het vergelijken van twee preparaten de toetsing van de evenwijdigheid der beide regressielijnen.² De methode wordt zeer diepgaand beschreven door FINNEY (152) en behandeld door FINNEY (153, hoofdstuk 18), BLISS (130), EMMENS (151), BURN (135) en VAN STRIK (199). Deze publicaties bevatten de benodigde tabellen, evenals FISHER en YATES (66) en de BIOMETRIKA TABLES (74).

¹ FINNEY (153) behandelt in hoofdstuk 20 nog verschillende andere eenvoudige methoden, maar op verschillende gronden beveelt hij slechts deze twee voor gebruik naast de probit-analyse aan.

² Verschillende statistici zijn dan ook van mening, dat men gewoonlijk beter een eenvoudiger methode voor de ED_{50} bepaling kan toepassen (bv. die van SPEARMAN-KÄRBER). Deze blijkt vrijwel dezelfde ED_{50} met slechts iets wijdere betrouwbaarheidsgrenzen op te leveren als de volledige probit-analyse (met een voldoende aantal iteraties), terwijl laatstgenoemde methode gemiddeld ongeveer twintig maal zoveel tijd kost. Zie ook van DER WAERDEN (32) en BERKSON (214).

23.4 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

23.4.3. DE GRAFISCHE METHODE VAN LITCHFIELD EN WILCOXON

Deze auteurs hebben een reeks vernuftige nomogrammen samengesteld, met behulp waarvan de probit-analyse met een minimum aan rekenwerk kan worden uitgevoerd. Op het eerste gezicht lijkt deze methode ook nog niet zo eenvoudig, maar als men eenmaal met de gebruiksaanwijzing bekend is, kan men er snel en voldoende nauwkeurig mee werken. Men raadplege daartoe de oorspronkelijke publicatie (174).

23.4.4. DE REGRESSIE-ANALYSE NA TOEPASSING VAN DE BGSIN TRANSFORMATIE

Bij deze methode wordt de fractie reacties p_j per log dosis x_j in de waarde $y_j = 2 \text{ bgsin } \sqrt{p_j}$ omgezet (zie 21.5.4), waarna de regressielijn van y op x wordt berekend. Als schatting van $m = \log ED_{50}$ neemt men dan het getal, dat volgens deze lijn behoort bij $p = 0,5$, d.i. $y = 1,5708$, zodat $m = (1,5708 - \bar{y})/b + \bar{x}$. Door het gebruik van deze transformatie wordt — zoals wij reeds hebben gezien — bereikt, dat de variantie van \underline{y} vrijwel onafhankelijk wordt van P en bij benadering gesteld kan worden op $1/n$. Wij volstaan met een eenvoudig voorbeeld waaruit blijkt hoe men de berekeningen kan uitvoeren en verwijzen voor een uitvoeriger behandeling naar CLARINGBOLD c.s. (217).

Tabel 23.5. Bepaling van de ED_{50} via de bgsin transformatie. De gegevens betreffen de toxiciteit van rotenone bij *Macrosiphoniella sanborni*

Conc. mg/l	Log conc. x_j	Aantal ins. n_j	Aantal gestast r_j	$p_j = \frac{r_j}{n_j}$	y_j	$n_j x_j$	$n_j y_j$	$n_j x_j^2$	$n_j x_j y_j$
10,2	1,01	50	44	0,880	2,4341	50,50	121,7050	51,0050	122,922050
7,7	0,89	49	42	0,857	2,3660	43,61	115,9340	38,8129	103,181260
5,1	0,71	46	24	0,522	1,6148	32,66	74,2808	23,1886	52,739368
3,8	0,58	48	16	0,333	1,2303	27,84	59,0544	16,1472	34,251552
2,6	0,41	50	6	0,120	0,7075	20,50	35,3750	8,4050	14,503750
Totaal		243				175,11	406,3492	137,5587	327,597980
		Σn				Σnx	Σny	Σnx^2	Σnxy

Uit J. T. MARTIN, Ann. Appl. Biol., 29 (1942), 69—81, tabel 9.

Het voorbeeld betreft een proef waarbij zes groepen van ongeveer 50 exemplaren van *Macrosiphoniella sanborni* werden bespoten met zes verschillende concentraties van rotenone in een alcohol-saponine oplossing. De insecten werden 24 uur na de bespuiting onderzocht, waarbij het aantal 'aangetaste' dieren (dood, stervend of niet meer in staat zich voort te bewegen) werd geregistreerd. In een contrôle-

groep van 49 exemplaren, die alleen met het oplosmiddel werd bespoten, werden geen aangetaste insecten waargenomen, zodat mag worden aangenomen dat gedurende de proef geen (of een te verwaarlozen kleine) natuurlijke sterfte is opgetreden. De uitkomsten van de proef zijn links in tabel 23.5 gegeven.

De waarden van $y = 2bgsin \sqrt{p}$ kunnen uit tabel AA worden afgelezen. De coëfficiënten van de regressievergelijking van y op x kunnen worden berekend met de volgende formules, waarin de aantallen n_j als gewichten zijn genomen:

$$(23.15) \quad a = \frac{\sum nx^2 \cdot \sum ny - \sum nx \cdot \sum nxy}{\sum n \cdot \sum nx^2 - (\sum nx)^2}$$

en

$$(23.16) \quad b = \frac{\sum n \cdot \sum nxy - \sum nx \cdot \sum ny}{\sum n \cdot \sum nx^2 - (\sum nx)^2}$$

Zoals reeds is opgemerkt is dan

$$(23.17) \quad m = (1,5708 - \bar{y})/b + \bar{x} = (1,5708 - a)/b,$$

zodat $ED_{50} = \text{antilog } m$.

Voor het voorbeeld zijn de benodigde berekeningen in tabel 23.5 uitgevoerd en hierbij vindt men

$$a = \frac{(137,5587)(406,3492) - (175,11)(327,59798)}{243(137,5587) - (175,11)^2} = -0,5316$$

en

$$b = \frac{(243)(327,59798) - (175,11)(406,3492)}{243(137,5587) - (175,11)^2} = 3,0582,$$

zodat

$$m = \{1,5708 - (-0,5316)\}/3,0582 = 0,6875.$$

De ED_{50} wordt dus geschat op antilog $0,6875 = 4,87$ mg/l. Deze uitkomst is vrijwel gelijk aan die van de probit-analyse, $m = 0,686$ en $ED_{50} = 4,85$ mg/ml.

23.4.5. DE METHODE VAN SPEARMAN EN KÄRBER

Deze methode kan worden toegepast, als men een niet te klein aantal doses gebruikt, die betrekkelijk dicht bij elkaar liggen en die zodanig gekozen zijn, dat voor de laagste log dosis, x_1 , geldt dat $p_1 = 0$ en voor de hoogste log dosis, x_k , dat $p_k = 1$. Dan is $p_{j+1} - p_j$ een schatting van de fractie proefdieren, waarvan de individueel-effectieve dosis tussen x_j en x_{j+1} ligt. Als vaststaat, dat doses kleiner dan x_1 geen reactie opleveren en dat doses groter dan x_k steeds tot een reactie leiden en als

ondersteld wordt, dat de verdeling van de logaritmen der individueel-effectieve doses symmetrisch is, kan men $m = \log ED_{50}$ schatten met

$$(23.18) \quad m = \sum_{j=1}^h (\hat{p}_{j+1} - \hat{p}_j) \left(\frac{x_j + x_{j+1}}{2} \right).$$

Als de logaritmen van de doses met een constante opklimmen, zodat $x_{j+1} - x_j = d$ voor elke j , dan kan men voor (23.18) schrijven

$$(23.19) \quad m = x_h + \frac{1}{2}d - d \sum_{j=1}^h \hat{p}_j$$

en als tevens het aantal dieren per dosis constant en gelijk aan n is, kan men gebruik maken van

$$(23.20) \quad m = x_h + \frac{1}{2}d - \frac{d}{n} \sum_{j=1}^h r_j.$$

De variantie van m is bij benadering gelijk aan

$$(23.21) \quad s_m^2 = \sum_{j=1}^h \frac{\hat{p}_j q_j}{4(n_j - 1)} (x_{j+1} - x_{j-1})^2.$$

Wanneer $x_{j+1} - x_j$ constant is geldt

$$(23.22) \quad s_m^2 = d^2 \sum_{j=1}^h \left(\frac{\hat{p}_j q_j}{n_j - 1} \right)$$

en wanneer men bij elke dosis n dieren gebruikt kan men deze formule verder vereenvoudigen tot

$$(23.23) \quad s_m^2 = \frac{d^2}{n^2(n-1)} \sum_{j=1}^h [r_j(n-r_j)].$$

Strikt genomen dient σ_m^2 met een component te worden verhoogd, die een functie is van d/σ (waarin σ de spreiding van de verdeling van de logaritmen der individueel-effectieve doses is). Deze component kan echter zonder bezwaar worden verwaarloosd, als d/σ zeer klein is en dit kan worden bereikt door d klein te maken t.o.v. σ , d.i. door zoveel mogelijk doseringen over het gehele te onderzoeken dosistraject te kiezen. Men behoeft hierbij niet met constante n te werken. VAN DER WAERDEN (32) merkt op: 'Die Genauigkeit wird sogar erhöht, wenn die Anzahlen in der Mitte grösser gewählt werden als an den Enden der Reihe. Das starke Heranziehen der starken Dosen, bei denen fast alle Tiere sterben, und der schwachen Dosen, bei denen fast keines stirbt, ist Verschwendung von Zeit und Tiermaterial.

Men kan de methode ook nog wel gebruiken, als de fractie \hat{p}_1 zo dicht bij 0 ligt, dat het verantwoord is om aan te nemen, dat bij uitbreiding

Tabel 23.6. Bepaling van de ED_{50} volgens SPEARMAN-KÄRBER bij muizen voor de beschermende werking van luminal (phenylaethylbarbituurzuur) t.o.v. electroshock, toediening intraperitoneaal, 1 uur voor electroshock. Gegevens: VAN PROOSDIJ-HARTZEMA (188)

1	2	3	4	5	6	7	8	9	10	11	12	13	14
Dosis in mg/kg	Log dosis x_j	Aantal dieren n_j	Beschermde dieren r_j	$p_j = r_j/n_j$	q_j	$x_j + x_{j+1}$	$p_{j+1} - p_j$	$(7) \times (8)$	$x_{j+1} - x_{j-1}$	$(10)^2$	$p_j q_j$	$4(n_j - 1)$	$\frac{(11) \times (12)}{(13)}$
5	0,699	10	0	0,000	1,000								
10	1,000	10	0	0,000	1,000	2,301	0,033	0,075933	0,602	0,362404	0	36	0
20	1,301	30	1	0,033	0,967	2,903	0,167	0,484801	0,602	0,362404	0,031911	116	0,00009970
40	1,602	20	4	0,200	0,800	3,350	0,450	1,507500	0,447	0,199809	0,160000	76	0,00042065
56	1,748	20	13	0,650	0,350	3,651	0,217	0,792267	0,301	0,090601	0,227500	76	0,00027121
80	1,903	30	26	0,867	0,133	3,952	0,083	0,328016	0,301	0,090601	0,115311	116	0,00009006
112	2,049	20	19	0,950	0,050	4,253	0,050	0,212650	0,301	0,090601	0,047500	76	0,00005663
160	2,204	10	10	1,000	0,000								
$2m = 3,401167$							$m = 1,70058$				$s_m^2 = 0,00093825$		

23.4 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

van de dosisreeks met een lagere dosis x_0 als resultaat $p = 0$ zal verschijnen, resp. als de fractie p_k zo dicht bij 1 ligt, dat de naasthogere dosis in $p_{k+1} = 1$ zal resulteren. Deze uitkomsten worden dan aan de waargenomen reeks p 's toegevoegd.

De SPEARMAN-KÄRBER methode is zeer geschikt, als men weinig weet omtrent de ED_{50} en dus een lange dosisreeks moet kiezen teneinde er zeker van te zijn dat deze de ED_{50} insluit. BROSS (132) heeft gevonden, dat de methode in zo'n geval (waarbij men betrekkelijk weinig dieren per dosis gebruikt) waarschijnlijk te verkiezen is boven de probit-analyse. Weet men ongeveer, waar de ED_{50} ligt, dan kan men efficiënter werken door een reeks doses met een kleinere spreidingsbreedte en grotere onderlinge afstand in de buurt van de ED_{50} te kiezen en de analyse uit te voeren volgens de methode van THOMPSON of met de probit-analyse. Hierbij dient wel te worden opgemerkt, dat de probit-analyse een zeer belangrijk voordeel boven de methoden van SPEARMAN-KÄRBER en THOMPSON heeft, n.l. dat het tevens mogelijk is de validiteit van de analyse te onderzoeken (toetsing op rechtlijnigheid, en bij twee preparaten de toetsing op de evenwijdigheid van de regressielijnen).

In tabel 23.6 is de SPEARMAN-KÄRBER methode op een voorbeeld toegepast. Dit is gekozen, omdat de aantallen proefdieren verschillend zijn en omdat de ED_{50} , bepaald volgens de methode van LITCHFIELD en WILCOXON, ter beschikking stond. De berekeningen, die nodig zijn voor het toepassen van de formules (23.18) en (23.21) zijn in deze tabel uitgevoerd en zij leveren $m = 1,70058$ op, zodat $ED_{50} = \text{antilog } 1,70058 = 50,2$. Deze uitkomst stemt goed overeen met de ED_{50} volgens LITCHFIELD-WILCOXON, die 51 bedraagt. Verder is $s_m = \sqrt{0,00093825} = 0,03063$, zodat het 95% betrouwbaarheidsinterval van de ware ED_{50} bij benadering gelijk is aan

$$\begin{aligned} \text{antilog } [1,70058 - 2(0,03063)] &< ED_{50} < \text{antilog } [1,70058 + 2(0,03063)], \\ \text{antilog } (1,63932) &< ED_{50} < \text{antilog } (1,76184), \\ 43,6 &< ED_{50} < 57,8. \end{aligned}$$

Dit interval is iets kleiner dan dat volgens LITCHFIELD-WILCOXON, waarvoor geldt: $42,1 < ED_{50} < 61,7$.

Wij geven in tabel 23.7 nog een tweetal fictieve voorbeelden om te demonstreren, dat bij constante opklimming van de log doses de berekeningen aanzienlijk vereenvoudigd kunnen worden en dat een nog verdere vereenvoudiging optreedt, als tevens alle groepen even groot worden gekozen. De fracties reagerende dieren zijn bij A en B identiek gekozen, zodat dezelfde ED_{50} wordt verkregen. De variantie van m is bij A (met grotere aantallen dieren in het midden) zo weinig lager dan bij B (constante n), dat het verschil praktisch niet van betekenis is.

Tabel 23.7. Bepaling van de ED_{50} volgens SPEARMAN en KÄRBER, met *A*. Log doses equidistant, verschillende waarden van *n*, *B*. Log doses equidistant en *n* constant (fictieve gegevens)

Dosis mg/kg d_j	Log dosis x_j	<i>A</i>						<i>B</i>		
		Aantal dieren n_j	Aantal reacties r_j	p_j	q_j	$p_j q_j$	$\frac{p_j q_j}{n_j - 1}$	Aantal dieren	Aantal reacties	$r_j(n - r_j)$
3,15	0,50	10	0	0,00	1,00	0,0000	0,000000	20	0	0
3,55	0,55	10	1	0,10	0,90	0,0900	0,010000	20	2	36
4,00	0,60	20	3	0,15	0,85	0,1275	0,006711	20	3	51
4,45	0,65	40	16	0,40	0,60	0,2400	0,006154	20	8	96
5,00	0,70	40	26	0,65	0,35	0,2275	0,005833	20	13	91
5,60	0,75	20	15	0,75	0,25	0,1875	0,009868	20	15	75
6,30	0,80	10	9	0,90	0,10	0,0900	0,010000	20	18	36
7,10	0,85	10	10	1,00	0,00	0,0000	0,000000	20	20	0
		160		3,95			0,048566	160	79	385

Proef *A*:

$$m = 0,85 + 0,5 (0,05) - 0,05 (3,95) = 0,6775, \quad ED_{50} = 4,76 \text{ mg/kg}$$

$$s_m^2 = (0,05)^2 (0,048566) = 0,00012142, \quad s_m = 0,01102.$$

Proef *B*:

$$m = 0,85 + 0,5 (0,05) - \frac{0,05 (79)}{20} = 0,6775, \quad ED_{50} = 4,76 \text{ mg/kg}$$

$$s_m^2 = \frac{(0,05)^2 (385)}{(20)^2 (19)} = 0,00012664, \quad s_m = 0,01125.$$

23.4.6. DE METHODE DER VOORTSCHRIJDENDE GEMIDDELDEN VAN THOMPSON

THOMPSON heeft een methode aangegeven, waarbij voor het bepalen van *m* gebruik wordt gemaakt van voortschrijdende gemiddelden.

Bij het bepalen van het voortschrijdend gemiddelde van een reeks getallen gaat men als volgt te werk. Men kiest eerst de *lengte* van het voortschrijdend gemiddelde, d.i. het aantal termen waarop het wordt gebaseerd. Bedraagt deze lengte bv. drie, dan bepaalt men eerst het gemiddelde van de eerste drie termen van de reeks, vervolgens bepaalt men het gemiddelde van de termen met de rangnummers twee t/m vier, dan dat van de termen met de rangnummers drie t/m vijf, enz. Voor ieder volgend gemiddelde laat men dus het groepje termen, waarvan het gemiddelde berekend wordt, één term in de reeks 'voortschrijden'.¹

¹ Zie voor een uitvoeriger beschrijving van voortschrijdende gemiddelden bv. DE WOLFF (38) of CROXTON en COWDEN (6).

23.4 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

Bij de methode van THOMPSON wordt een voortschrijdend gemiddelde met een lengte K verkregen, door voor iedere reeks van K opeenvolgende doses, te beginnen met p_1 , te berekenen

$$(23.24) \quad p^* = (p_j + p_{j+1} + \dots + p_{j+K-1})/K$$

en deze waarde te verbinden aan de dosis

$$(23.25) \quad x = (x_j + x_{j+1} + \dots + x_{j+K-1})/K.$$

Men kan dan $m = \log ED_{50}$ bepalen door lineair te interpoleren tussen de twee opeenvolgende waarden van p^* aan weerszijden van 0,50. De methode kan alleen worden gebruikt voor het schatten van de ED_{50} en hierbij dient te worden aangenomen, dat de logarithmen van de individueel-effectieve doses symmetrisch verdeeld zijn. Gewoonlijk kiest men K zo groot als mogelijk is.

Wij beschouwen slechts de toepassing van deze methode bij een volledig symmetrische proef, zoals die in tabel 23.7, voorbeeld B. Kiest men bij dit voorbeeld $K = 4$, dan dient men eerst met behulp van de formules (23.25) en (23.24) de volgende hulptabel samen te stellen:

	Begindosis x_j	x volgens (23.25)	p^* volgens (23.24)
	0,50	0,575	0,1625
	0,55	0,625	0,3250
$x_m =$	0,60	0,675	0,4875
$x_{m+1} =$	0,65	0,725	0,6750 ← 0,50

Uit deze tabel kan m rechtstreeks worden bepaald door als volgt lineair te interpoleren:

$$m = 0,675 + \frac{0,5000 - 0,4875}{0,6750 - 0,4875} \cdot 0,05 = 0,6783.$$

Men kan ook gebruik maken van de formule

$$(23.26) \quad m = x_m + \frac{1}{2}d(K+1) - df,$$

met

$$(23.27) \quad f = \frac{p_{m+1} + p_{m+2} + \dots + p_{m+K} - \frac{1}{2}K}{p_{m+K} - p_m},$$

waarin p_m de met x_m (zie de hulptabel) corresponderende dosis is.

De berekening verloopt dan op de volgende wijze:

$$f = \frac{0,40 + 0,65 + 0,75 + 0,90 - 2}{0,90 - 0,15} = \frac{0,70}{0,75} = \frac{14}{15},$$

$$m = 0,60 + (0,5)(0,05)(5) - (0,05)\frac{14}{15} = 0,6783.$$

De variantie van m is bij benadering gelijk aan

$$(23.28) \quad s_m^2 = \frac{d^2 \{f^2 p_m q_m + p_{m+1} q_{m+1} + \dots + p_{m+K-1} q_{m+K-1} + (1-f)^2 p_{m+K} q_{m+K}\}}{(n-1) (p_{m+K} - p_m)^2}$$

In het algemeen neemt deze variantie af, naarmate K groter wordt gekozen. Voor het voorbeeld vindt men

$$s_m^2 = \frac{(0,05)^2 \left\{ \left(\frac{14}{15} \right)^2 (0,15) (0,85) + (0,40) (0,60) + \dots + \left(\frac{1}{15} \right)^2 (0,90) (0,10) \right\}}{19 (0,90 - 0,15)^2}$$

= 0,00017929, zodat $s_m = 0,01339$.

Gebruikt men voortschrijdende gemiddelden van de lengte $K = 6$, dan vindt men $m = 0,6778$ en $s_m = 0,01244$. Deze uitkomsten verschillen slechts weinig van de voorgaande en zij stemmen fraai overeen met die van de SPEARMAN-KÄRBER methode.

In tegenstelling tot de methode van SPEARMAN en KÄRBER is het bij deze methode niet noodzakelijk dat er doses met $p = 0$ en $p = 1$ voorkomen. Zij kan dus worden gebruikt als men reeds ongeveer de ligging van de ED_{50} kent en zodoende de proefdieren over een beperkter dosistraject heeft verdeeld.

23.4.7. DE SEQUENTE METHODE VAN DIXON EN MOOD

De tot dusver behandelde methoden voor de ED_{50} bepaling bezitten het bezwaar, dat men gewoonlijk een vrij brede dosisreeks moet kiezen wanneer men er zeker van wil zijn, dat deze de ED_{50} insluit. Men verkrijgt dan een aantal lage doses (met p dicht bij 0) en/of een aantal hoge doses (met p in de nabijheid van 1) die omtrent de ED_{50} naar verhouding weinig informatie over de ED_{50} geven.

Door DIXON en MOOD (143) is een sequente methode ontworpen waarmee bereikt kan worden, dat het merendeel der proefdieren een dosis in de nabijheid van de ED_{50} krijgt. Hierdoor bereikt men dat voor een ED_{50} bepaling met een gegeven nauwkeurigheid 30 à 40 procent minder waarnemingen vereist zijn dan bij de niet-sequente methoden. De sequente methode kan worden gebruikt onder de volgende voorwaarden: 1. Men moet de proefdieren één voor één behandelen. 2. Men moet de uitkomst van elk dier kennen, voordat men het volgende dier kan behandelen. 3. Het moet mogelijk zijn om opeenvolgend verschillende doses toe te dienen. 4. De verdeling van de logaritmen der individueel-effectieve doses dient (ongeveer) normaal te zijn en men moet de spreiding σ van deze verdeling bij benadering kennen.

De door DIXON en MOOD aangegeven procedure verloopt op de volgende wijze:

23.4 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

1. Kies een dosis d_0 die zo dicht mogelijk bij de te verwachten ED_{50} ligt. Stel — zoals gebruikelijk — $\log d_0 = x_0$. Kies dan vervolgens een reeks doses kleiner dan d_0 met als logaritmen x_{-1}, x_{-2}, x_{-3} , enz. en een reeks doses groter dan d_0 met als logaritmen x_1, x_2, x_3 , enz. Tussen elke twee opeenvolgende waarden van de x -reeks dient een constant verschil d aanwezig te zijn. Voor d kiese men een waarde die ongeveer gelijk is aan σ , maar die in ieder geval kleiner is dan 2σ .

2. Stel het aantal proefdieren N vast, dat men wil behandelen.

3. Geef de dosis x_0 aan het eerste proefdier. De reactie van dit dier op deze dosis bepaalt de dosis, die aan het volgende dier zal worden gegeven. Reageert dit dier positief, dan krijgt het tweede dier de dosis x_{-1} (dus één stap *lager*). Treedt geen reactie op, dan krijgt het de dosis x_1 (dus één stap *hoger*). Ga op deze wijze door totdat alle N dieren zijn behandeld.

Tabel 23.8. Uitkomsten van een sequente proef ter bepaling van de ED_{50} volgens de methode van DIXON en MOOD

Dosis d	Log dosis x	Uitkomsten ($N = 40$)	Aantal	
			+	0
3,55	$x_3 = 0,55$			
3,15	$x_2 = 0,50$	+	1	
2,82	$x_1 = 0,45$	+ + + 0 + +	4	1
2,51	$x_0 = 0,40$	0 + 0 + + + 0 + + 0 + +	11	4
2,24	$x_{-1} = 0,35$	0 + 0 0 0 0 + 0 0 0 + 0 0 +	4	10
2,00	$x_{-2} = 0,30$	0 + 0 0 0 0 0 0 0 +	1	3
1,78	$x_{-3} = 0,25$	0		1
			21	19 = n

In tabel 23.8 zijn de uitkomsten van een sequente proef met $N = 40$ opgenomen. De dieren waarbij een positieve reactie is waargenomen zijn met een + en de niet reagerende dieren met een 0 aangeduid. Men gaat nu verder als volgt te werk:

4. Bepaal voor de reagerende en voor de niet reagerende dieren de frequentieverdelingen van de doses (zie de laatste twee kolommen van de tabel). De totalen van deze verdelingen zullen gewoonlijk ongeveer gelijk zijn. Er wordt nu verder gewerkt met de verdeling die het kleinste aantal (= n) waarnemingen bevat. Bij het voorbeeld is dit de verdeling der niet reagerende dieren, zodat $n = 19$.

5. Ken aan de x -waarden (log doses) van deze verdeling van laag naar hoog de getallen $0, 1, 2, \dots, k$ toe. Als de bij deze doses behorende frequenties resp. $f_0, f_1, f_2, \dots, f_k$ zijn ($\sum f_i = n$), dan berekent men eerst

$$(23.29) \quad A = \sum_{i=0}^k i f_i$$

en dan is

$$(23.30) \quad m = \log ED_{50} = x' + d \left(\frac{A}{n} \pm \frac{1}{2} \right),$$

waarin x' de log dosis is, waaraan het getal 0 werd toegekend. Het plus-teken wordt gebruikt, als de berekeningen zijn uitgevoerd met de niet reagerende dieren, het minteken als zij berusten op de dieren waarbij de reactie wel is waargenomen.

Bij het voorbeeld verlopen de berekeningen als volgt:

Code i	f_i	if_i	i^2f_i
4	1	4	16
3	4	12	36
2	10	20	40
1	3	3	3
$x' = 0,25 = 0$	1	0	0
	$n = 19$	$A = 39$	$B = 95$

$$m = 0,25 + 0,05 \left(\frac{39}{19} + \frac{1}{2} \right) = 0,3776,$$

$$ED_{50} = \text{antilog } 0,3776 = 2,39.$$

Een schatting van σ kan worden verkregen door eerst te berekenen

$$(23.31) \quad B = \sum_{i=0}^h i^2 f_i$$

en vervolgens

$$(23.32) \quad s = 1,620d \left(\frac{nB - A^2}{n^2} + 0,029 \right).$$

Men kan tenslotte s_m berekenen met

$$(23.33) \quad s_m = \frac{Gs}{\sqrt{n}},$$

zodat men bij benadering een $100(1-\alpha)\%$ betrouwbaarheidsinterval van $\log ED_{50}$ verkrijgt door te nemen

$$(23.34) \quad m + T_{\frac{1}{2}\alpha} s_m < \log ED_{50} < m + T_{1-\frac{1}{2}\alpha} s_m,$$

waaruit dan weer het overeenkomstige betrouwbaarheidsinterval van de ED_{50} volgt door de antilogarithmen van de verkregen grenswaarden te nemen. Indien d/σ kleiner is dan 2,5, is G in (23.33) vrijwel alleen afhankelijk van de waarde, die d/σ bij de betrokken proef bezit. Uit

23.4 STATISTISCHE METHODEN BIJ DE BIOLOGISCHE STANDAARDISATIE

een nomogram van DIXON en MOOD (dat ook te vinden is in DIXON en MASSEY) blijkt, dat in dit gebied bij benadering geldt:

d/σ	G
0,5	0,885
0,75	0,92
1,00	0,99
1,5	1,08
2,0	1,15
2,5	1,23

De waarden van G blijken dus bij deze zeer uiteenlopende waarden van d/σ betrekkelijk weinig van 1 af te wijken.

Past men het voorgaande op het voorbeeld toe, dan vindt men

$$s = (1,620) (0,05) \left(\frac{19 \times 95 - (39)^2}{(19)^2} + 0,029 \right) = 0,0661.$$

Daar $d = 0,05$, is $d/s = 0,76$. Wij blijven aan de veilige kant, als wij $d/\sigma = 1$ en dus $G \cong 1$ stellen en vinden dan

$$s_m = \frac{(1) (0,0661)}{\sqrt{19}} = 0,0152.$$

De 95% betrouwbaarheidsgrenzen van de 'ware' ED_{50} zijn dus bij benadering antilog $[0,3776 \pm 1,96 (0,0152)]$, d.i. 2,23 en 2,56.

DIXON en MOOD menen, dat hun methode pas bruikbaar is als men tenminste 40 waarnemingen verricht. Dit houdt in dat de totale duur van een sequente proef op z 'n minst 40 maal zo groot is als de gemiddelde tijd die nodig is om te kunnen vaststellen of de reactie al dan niet optreedt. De proef kan hierdoor een zo lange duur krijgen, dat zij in de praktijk niet goed is uit te voeren. BROWNLEE c.s. (134) hebben echter gedemonstreerd, dat men zonder bezwaar gebruik kan maken van reeksen van minstens 10 waarnemingen, mits elke reeks begint met een dosis, die niet verder dan $2d$ van m verwijderd is. Teneinde aan laatstgenoemde voorwaarde te voldoen hebben zij een modificatie in de door DIXON en MOOD ontworpen methode aangebracht. Men kan dan met een aantal korte series werken, die zo nodig gelijktijdig kunnen worden ingezet en waarvan de uitkomsten later kunnen worden gecombineerd. Een nadeel van hun methode is echter, dat geen weg wordt aangegeven voor het bepalen van betrouwbaarheidsgrenzen. Men zou dit dan echter kunnen doen langs de door DIXON en MOOD aangegeven weg. Het is ook mogelijk om op de sequent verkregen uitkomsten de probit-analyse (of de grafische methode van LITCHFIELD en WILCOXON) toe te passen.

De methode van BROWNLEE wordt als volgt uitgevoerd:

1. Als punt 1 van DIXON en MOOD (zie blz. 614).
2. Kies het totaal aantal proefdieren n waarop men de ED_{50} bepaling wil baseren.

3. Evenals bij DIXON en MOOD geeft men het eerste dier de dosis x_0 ; bij het optreden van de reactie krijgt het volgende dier de dosis x_{-1} , bij het uitblijven van de reactie de dosis x_1 . Ga op deze wijze door, totdat $n-1$ dieren behandeld zijn *na de dieren, waarbij de eerste reeks uitkomsten met hetzelfde teken is verkregen*. Het totale benodigde aantal dieren (N) varieert dus: zijn de reacties van de eerste twee dieren verschillend, dan is $N = n$, anders is steeds $N > n$.

4. Bereken $m = C/n$. Hierin is C de som van de logaritmen van de doses, welke men aan de laatste $n-1$ behandelde dieren heeft toegevend, vermeerderd met de logaritme van de dosis, die aan het hierop volgende dier *zou zijn gegeven* (dit laatste dier behoeft dus niet te worden behandeld).

5. Bereken $ED_{50} = \text{antilog } m$.

Tabel 23.9. Uitkomsten van een sequente proef ter bepaling van de ED_{50} volgens de methode van BROWNLEE c.s.

Dosis d	Log dosis x	Uitkomsten						f	fx			
4,45	$x_0 = 0,65$	+										
4,00	$x_{-1} = 0,60$	+										
3,55	$x_{-2} = 0,55$	+										
3,15	$x_{-3} = 0,50$	+										
2,82	$x_{-4} = 0,45$	+	+				+	2	0,90			
2,51	$x_{-5} = 0,40$		0	+	+	+	0	+	+	+	9	3,60
2,24	$x_{-6} = 0,35$			0	0	+	0	0	0	0	8	2,80
2,00	$x_{-7} = 0,30$					0					1	0,30
		←----- $n-1 = 19$ ----->						$n=20$	$C = 7,60$			

$$m = C/n = 7,60/20 = 0,3800, \quad ED_{50} = \text{antilog } 0,3800 = 2,40$$

In tabel 23.9 is een voorbeeld gegeven, waarbij $n = 20$ is gekozen. Het is duidelijk, dat men bij deze proef een hogere ED_{50} heeft verwacht en zodoende de dosis d_0 (en dus x_0) te hoog heeft gekozen. Deze 'fout' wordt geredresseerd, doordat bij de berekening de eerste vijf uitkomsten niet in aanmerking worden genomen. De met * aangeduide uitkomst volgt zonder meer uit de voorgaande en berust dus niet op een proef.

Wanneer ter bepaling van een ED_{50} k series proeven, ieder van de lengte n , zijn uitgevoerd en men bij iedere serie van dezelfde x_0 is uitgegaan, dan kan men eerst per serie $m_i = C_i/n$ berekenen en vervolgens de ED_{50} vinden door de antilogaritme te nemen van $\sum_{i=1}^k m_i/k$.

ENKELE BIJZONDERE TECHNIEKEN

In dit hoofdstuk bespreken wij enkele speciale technieken die niet bij de voorgaande hoofdstukken konden worden ondergebracht.

24.1. De toets van Yates

24.1.1. TOEPASSING OP EEN $2 \times r$ TABEL

In tabel 9.11 (blz. 205) zijn de resultaten vergeleken van twee groepen parkinsonpatiënten, waarvan de eerste alleen met disipal werd behandeld en de tweede bestaat uit patiënten, die voordien reeds andere middelen kregen. Het resultaat van de behandeling, de mate van verbetering, is uitgedrukt in vier kwalitatieve categorieën ($++$, $+$, \pm en 0). De toegepaste χ^2 -toets levert $X_0^2 = 5,16$ (met $\nu = 3$) op, zodat $0,2 > P > 0,1$. Deze toets leidt dus tot de uitspraak dat de uitkomsten van de beide groepen bij een 5% drempel niet significant verschillen.

Men kan echter het gebruikte criterium, de mate van verbetering, ook zien als een *kwantitatieve* grootheid waarvan de 'waarden' in vier 'klassen' zijn ingedeeld. Deze klassen missen een scherpe kwantitatieve aanduiding en het staat dus bv. niet vast, dat zij dezelfde breedte bezitten. Wanneer het echter mogelijk is, deze klassen *arbitrair* een reeks waarden toe te kennen, zoals bv. $++ = 3$, $+$ = 2, $\pm = 1$ en $0 = 0$, dan kan men een door YATES (232) aangegeven toets toepassen, die vnl. gevoelig is voor verschillen in niveau van de verdelingen van beide groepen. Wij zullen deze toets en zijn relatie tot de toets van WILCOXON beknopt bespreken.

COCHRAN (80) merkt omtrent het toekennen van de klassewaarden ('scores') het volgende op: 'Such scores are to some extent subjective and arbitrary, and some scientists may feel that the assignment of scores is slightly unscrupulous, or at least they are uncomfortable about it. Actually, *any* set of scores gives a valid test, provided that they are constructed without consulting the results of the experiment. If the set of scores is poor, in that it badly distorts a numerical scale that really does underlie the ordered classification, the test will not be sensitive. The scores should therefore embody the best insight available about the way in which the classification was constructed and used.'

Onderstel dat men aan de categorieën een reeks waarden z_i ($i = 1, 2, \dots, r$) kan toekennen en dat men de bij de z_i behorende frequenties

van de eerste steekproef (groep) aangeeft met f_{i1} ($\sum_{i=1}^r f_{i1} = n_1$), die van de tweede steekproef (groep) met f_{i2} ($\sum_{i=1}^r f_{i2} = n_2$) en de totale frequenties ($f_{i1} + f_{i2}$) met f_i ($\sum_{i=1}^r f_i = n = n_1 + n_2$). Deze notatie wordt verduidelijkt in tabel 24.1, waarin de gegevens van het disipal-voorbeeld nogmaals zijn opgenomen met als waarden van z resp. 3, 2, 1 en 0. Men berekent nu achtereenvolgens:

$$(24.1.1) \quad \bar{z}_1 = \sum_{i=1}^r f_{i1} z_i / n_1,$$

$$(24.1.2) \quad \bar{z}_2 = \sum_{i=1}^r f_{i2} z_i / n_2,$$

$$(24.2) \quad s_z^2 = \frac{\sum_{i=1}^r f_i (z_i - \bar{z})^2}{n - 1} = \frac{\sum_{i=1}^r f_i z_i^2 - (\sum_{i=1}^r f_i z_i)^2 / n}{n - 1},$$

en

$$(24.3) \quad s_{z_1 - z_2}^2 = s_z^2 \left[\frac{1}{n_1} + \frac{1}{n_2} \right] = \frac{n \sum f_i z_i^2 - (\sum f_i z_i)^2}{n_1 n_2 (n - 1)}.$$

Onder de hypothese dat beide groepen uit dezelfde populatie afkomstig zijn is voor niet te kleine n_1 en n_2 de grootheid

$$(24.4) \quad T = \frac{\bar{z}_1 - \bar{z}_2}{s_{z_1 - z_2}}$$

bij benadering standaardnormaal verdeeld. Bij tweezijdige toetsing kan men dus bij een drempelwaarde α tot het verwerpen van H_0 overgaan, indien $|T_0| \geq T_{1-\alpha}$ is.

Tabel 24.1. De toets van YATES voor een $2 \times r$ tabel (gegevens uit tabel 9.11)

Resultaat	Waarden z_i	Frequenties		Totaal f_i	$f_{i1} z_i$	$f_{i2} z_i$	$f_i z_i$	$f_i z_i^2$
		Groep 1 f_{i1}	Groep 2 f_{i2}					
++	3	5	17	22	15	51	66	198
+	2	14	54	68	28	108	136	272
±	1	8	57	65	8	57	65	65
0	0	2	32	34	0	0	0	0
Totaal		29 n_1	160 n_2	189 n	51	216	267	535

$$\bar{z}_1 = 51/29 = 1,7586, \quad \bar{z}_2 = 216/160 = 1,3500, \quad \bar{z}_1 - \bar{z}_2 = 0,4086,$$

$$\text{Var}(\bar{z}_1 - \bar{z}_2) = \frac{(189)(535) - (267)^2}{(29)(160)(188)} = \frac{29826}{872320} = 0,03419158,$$

$$\sqrt{\text{Var}(\bar{z}_1 - \bar{z}_2)} = 0,1849, \quad T_0 = \frac{0,4086}{0,1849} = 2,21, \quad P_D < 0,05$$

In tabel 24.1 is de toets op de disipal-gegevens toegepast. Hoewel de uitkomst met enige reserve moet worden beschouwd, mede omdat het aantal patiënten in groep 1 betrekkelijk klein is, blijkt T_0 zoveel groter te zijn dan $T_{0,975} = 1,96$, dat H_0 bij een 5% drempel kan worden verworpen ten gunste van de alternatieve hypothese, dat 'gemiddeld' bij groep 1 een grotere verbetering optreedt dan bij groep 2. De toets van YATES bezit dus kennelijk t.o.v. het niveauverschil tussen de resultaten van de beide groepen een groter onderscheidingsvermogen dan de χ^2 -toets op de $2 \times r$ tabel, die ook gevoelig is voor andere verschillen.

Men kan deze toets ook toepassen als de waarnemingen een kwantitatieve grootheid x betreffen. Men verkrijgt dan een klassieke toets betreffende de gemiddelden van twee aselechte steekproeven, die in principe verschilt van de toets van STUDENT, omdat de populatievariantie niet uit s_1^2 en s_2^2 maar uit de gecombineerde verdeling van de twee steekproeven wordt geschat.

Voor ongeveer normaal verdeelde populaties is o.i. de toets van STUDENT te verkiezen boven de toets van YATES. Bij de STUDENT-toets neemt men immers (zie 17.2.3 op blz. 438)

$$f^2(n-2) = F(1, n-2) = \frac{s^2_{\text{tussen}}}{s^2_{\text{binnen}}} = \frac{S_{\text{tussen}}/1}{S_{\text{binnen}}/(n-2)}$$

en bij de toets van YATES

$$T^2 = \frac{s^2_{\text{tussen}}}{s^2_{\text{totaal}}} = \frac{S_{\text{tussen}}/1}{(S_{\text{tussen}} + S_{\text{binnen}})/(n-1)}$$

zodat tussen deze toetsen de volgende relatie bestaat:

$$T = t \sqrt{\frac{n-1}{n-2+f^2}}$$

Men ziet dus dat $T = t$ indien $f^2 = 1$, dat $T < t$ indien $f^2 > 1$ en dat $T > t$ indien $f^2 < 1$. Overigens bepaalt men de kans van een waargenomen waarde van T via de standaardnormale verdeling en de kans van een waargenomen waarde van t via de STUDENT-verdeling met $n-2$ vrijheidsgraden, zodat bij een betrekkelijk kleine omvang der steekproeven t_0 groter moet zijn dan T_0 om bij een drempelwaarde α tot significantie te leiden. Wanneer echter de steekproeven groot zijn, dan is $n-1 \cong n-2+f^2$ en dan kan de STUDENT-verdeling goed worden benaderd door de standaardnormale. De twee toetsen leveren dan vrijwel dezelfde uitkomst op.

Voor niet normaal verdeelde populaties bestaat de mogelijkheid, dat de benadering volgens YATES beter is dan die volgens STUDENT.

24.1.2. VERGELIJKING MET DE TOETS VAN WILCOXON

Voor een $2 \times r$ tabel kan men de hypothese, dat de twee groepen uit dezelfde populatie stammen, ook toetsen met de toets van WILCOXON. De normale benadering van deze toets (met correctie voor gelijken door middel van formule (10.15), maar zonder continuïteitscorrectie) blijkt dan identiek te zijn aan een toets van YATES, waarbij als waarden

van z de *gemiddelde rangnummers* van de groepen gelijk worden gekozen.¹ In tabel 24.2 wordt deze identiteit gedemonstreerd met behulp van het reeds in tabel 24.1 gebruikte voorbeeld.

Tabel 24.2. De toets van WILCOXON en de toets van YATES (met als waarden van z de gemiddelde rangnummers)

Resultaat	Frequenties		f_i	F_i	Gem. rang-nrs.* z_i	$f_{1i}z_i$	$f_{2i}z_i$	$f_i z_i$	$f_i z_i^3$	$f_i^3 = z_i^3$
	Gr. 1 f_{1i}	Gr. 2 f_{2i}								
++	5	17	22	189	178,5	892,5	3034,5	3927	700969,5	10648
+	14	54	68	167	133,5	1869,0	7209,0	9078	1211913,0	314432
±	8	57	65	99	67,0	536,0	3819,0	4355	291785,0	274625
0	2	32	34	34	17,5	35,0	560,0	595	10412,5	39304
Totaal	29 n_1	160 n_2	189 n	—	—	3332,5 S_1	14622,5 S_2	$\frac{17955}{2}$ $\frac{n(n+1)}{2}$	2215080,0	639009 D

Toets van Wilcoxon, normale benadering met correctie voor gelijken

$$n^3 = (189)^3 = 6751269, \quad D = \sum f_i^3 = \sum f_i^3 = 639009, \quad n^3 - D = 6112260,$$

$$(10.15) \quad \sigma^2 = \frac{(29)(160)(6112260)}{(3)(189)(188)} = 266059,57, \quad \sigma = 515,81,$$

$$(10.12) \quad T_0 = \frac{2(3332,5) - 29(190)}{515,81} = \frac{1155}{515,81} = 2,2392.$$

Toets van Yates

$$\bar{z}_1 = 3332,5/29 = 114,9138, \quad \bar{z}_2 = 14622,5/160 = 91,3906, \quad \bar{z}_1 - \bar{z}_2 = 23,5232,$$

$$\text{Var}(\bar{z}_1 - \bar{z}_2) = \frac{(189)(2215080) - (17955)^2}{(29)(160)(188)} = 110,35869,$$

$$\sqrt{\text{Var}(\bar{z}_1 - \bar{z}_2)} = 10,5052, \quad T_0 = \frac{23,5232}{10,5052} = 2,2392.$$

*) Deze kunnen eenvoudig worden bepaald uit de somfrequentieverdeling (F_i). Het gemiddelde rangnummer van de categorie 0 is $(1 + 34)/2$, dat van de categorie \pm is $(35 + 99)/2$, enz.

De normale benadering van de toets van WILCOXON kan dus worden beschouwd als een speciaal geval van de toets van YATES, die op zijn beurt voor grote steekproeven bij benadering dezelfde uitkomst oplevert als de toets van STUDENT.

24.1.3. TOEPASSING OP EEN $k \times r$ TABEL

Beschikt men over een $k \times r$ tabel en kan men aan de r categorieën opeenvolgend de waarden z_1, z_2, \dots, z_r toekennen, dan verloopt de

¹ Het bewijs is eenvoudig te leveren, maar wij laten dit in verband met de plaatsruimte achterwege.

door YATES aangegeven toets op de volgende wijze: Men berekent de gemiddelden

$$(24.5) \quad \bar{z}_j = \sum_{i=1}^r f_{ij} z_i / n_j = U_j / n_j \quad (j = 1, 2, \dots, k),$$

uit de verdeling van de totaalfrequenties $f_i = \sum_{j=1}^k f_{ij}$

$$(24.6) \quad \bar{z} = \sum_{i=1}^r f_i z_i / n = \sum_{j=1}^k U_j / n,$$

met $n = \sum n_j$, en

$$(24.7) \quad s_z^2 = \frac{\sum f_i z_i^2 - (\sum f_i z_i)^2 / n}{n - 1} = \frac{S}{n - 1}.$$

Stelt men nu dat

$$(24.8) \quad s_{z_j}^2 = s_z^2 / n_j = \frac{S}{n_j (n - 1)},$$

dan kan men als toetsingsgrootheid kiezen

$$(24.9.1) \quad X^2 = \sum_{j=1}^k \frac{(\bar{z}_j - \bar{z})^2}{s_{z_j}^2} = \frac{\sum_{j=1}^k n_j (\bar{z}_j - \bar{z})^2}{s_z^2}.$$

Onder de hypothese H_0 dat de k steekproeven afkomstig zijn uit populaties met dezelfde verdeling, volgt de grootheid X^2 bij benadering een χ^2 -verdeling met $k-1$ vrijheidsgraden. De benadering is beter, naarmate de steekproeven groter zijn. De nulhypothese wordt dus bij een onbetrouwbaarheidsdrempel α verworpen wanneer $X_0^2 \geq \chi_{1-\alpha}^2$ is. Voor de berekening kan men beter gebruik maken van de formule

$$(24.9.2) \quad X^2 = \frac{(n - 1) [\sum U_j^2 / n_j - (\sum U_j)^2 / n]}{S}.$$

De betekenis van de symbolen in deze formule blijkt uit de voorgaande formules.

In tabel 24.3 is een voorbeeld van de toepassing van deze toets gegeven. Het betreft de resultaten van drie groepen lijdende aan peri-arthritis humero-scapularis, uitgedrukt in de categorieën: +++ = nagenoeg geen klachten, ++ = goed verbeterd, + = iets verbeterd en 0 = onverbeterd. Groep A kreeg een serie van drie schijnbestralingen, groep B ontving drie bestralingen à 150 r huiddosis en groep C drie bestralingen à 250 r huiddosis. Voorzover dit uit de tekst kan worden opgemaakt zijn de groepen aselekt samengesteld (men heeft echter een ongunstige proefopzet gekozen: hoe zou deze verbeterd kunnen worden?). De patiënten werden qua verbetering beoordeeld door een rheumatoloog die onkundig was van de gegeven behandeling.

A priori dient te worden gezegd dat de uitkomst van de toetsing

met reserve gehanteerd zal moeten worden, daar de omvang der groepen tamelijk klein is (een proefopzet met even grote groepen zou in ieder geval beter zijn geweest). Aangezien het aannemelijk is, dat het verschil tussen de categorieën 0 en + kleiner is dan de overige verschillen zijn aan de grootheid z resp. de waarden 5, 3, 1 en 0 toegekend.

Tabel 24.3. De toets van YATES op een 3×4 tabel*

Verbe- tering	z_i	Groepen			f_i	$f_{11}z_i$	$f_{12}z_i$	$f_{13}z_i$	f_{i2}	$f_{i2}z_i^2$
		A	B	C						
		f_{i1}	f_{i2}	f_{i3}						
+++	5	—	5	4	9	—	25	20	45	225
++	3	4	20	8	32	12	60	24	96	288
+	1	4	6	—	10	4	6	—	10	10
0	0	5	9	3	17	—	—	—	—	—
Totaal		13	40	15	68	16	91	44	151	523
		n_1	n_2	n_3	n	U_1	U_2	U_3	Σf_{i2}	$\Sigma f_{i2}z_i^2$

$$X_0^2 = \frac{67 [(16)^2/13 + (91)^2/40 + (44)^2/15 - (151)^2/68]}{523 - (151)^2/68} = 7,31$$

* Gegevens uit N.T.v.G., 99 (1955), 463—464. Zie de tekst.

De waargenomen $X_0^2 = 7,31$ kan beschouwd worden als een waarde van χ^2 met 2 vrijheidsgraden en bezit bij benadering een overschrijdingskans van 0,03. De hypothese dat de waargenomen verschillen op toeval berusten kan dus bij een 5% drempel worden verworpen.

24.1.4. VERGELIJKING MET DE TOETS VAN KRUSKAL EN WALLIS

De toets van KRUSKAL en WALLIS, toegepast op een $k \times r$ tabel, blijkt identiek te zijn aan de toets van YATES met als waarden van z de gemiddelde rangnummers van de r groepen gelijken (totaalfrequenties). Beide toetsen leveren dan dezelfde X_0^2 op. Bij de eerstgenoemde toets kan men in plaats van formule (10.20) beter gebruik maken van

$$(24.10) \quad X^2 = \frac{(n-1) [\Sigma S_j^2/n_j - n(n+1)^2/4]}{(n^3 - D)/12},$$

waarin S_j de som van de rangnummers in de j -de steekproef is, terwijl $D = \Sigma f_i^3$. De gegevens in tabel 24.3 bezitten als gemiddelde rangnummers $+++ = 64$, $++ = 43,5$, $+ = 22,5$ en $0 = 9$, zodat de rangsommen van de groepen A, B en C resp. zijn $S_A = 309$, $S_B = 1406$ en $S_C = 631$. Verder geldt dat $D = (9)^3 + (32)^3 + (10)^3 + (17)^3 = 39410$, zodat formule (24.10) oplevert

$$X_0^2 = \frac{67 [(309)^2/13 + (1406)^2/40 + (631)^2/15 - (68)(69)^2/4]}{[(68)^3 - 39410]/12} = 6,94.$$

Deze uitkomst is iets lager dan die van de toets van YATES met de arbitrair gekozen waarden van z .

24.1.5. DE 2×2 TABEL

Voor een 2×2 tabel geldt het volgende: T^2 [normale benadering van de toets van FISHER met formule (9.10)] = $\frac{n-1}{n} X^2$ [χ^2 -benadering van de toets van FISHER met formule (9.12)] = T^2 [toets van YATES volgens formule (24.4), met $z_1 = 0$ en $z_2 = 1$] = T^2 [normale benadering van de toets van WILCOXON met correctie voor gelijken].

Met deze methoden toetst men de hypothese $H_0: P_1 = P_2 = P$, d.w.z. de twee aselecte steekproeven stammen uit dezelfde dichotome populatie, waarin de fractie elementen met kenmerk A gelijk is aan P . Noemt men de fractie elementen met kenmerk A in de eerste steekproef p_1 en in de tweede steekproef p_2 , dan kan men ook berekenen

$$(24.11) \quad T_0 = \frac{p_1 - p_2}{\sqrt{\frac{p_0 q_0}{n_1} + \frac{p_0 q_0}{n_2}}} = \frac{p_1 - p_2}{\sqrt{p_0 q_0}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}},$$

waarin $p_0 = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$ en $q_0 = 1 - p_0$. Onder H_0 volgt de grootheid T voor niet te kleine n_1 en n_2 bij benadering een standaardnormale verdeling. Deze toets is equivalent aan de χ^2 -benadering van de toets van FISHER zonder continuïteitscorrectie, daar $T_0^2 = X_0^2$. De continuïteitscorrectie kan eenvoudig worden aangebracht door de absolute waarde van $p_1 - p_2$ met $2/n$ te verminderen.

24.2. Verdere analyse van een $k \times r$ tabel

Wanneer men op een $k \times r$ tabel de χ^2 -toets toepast en deze leidt tot het verwerpen van de hypothese, dat de k steekproeven uit populaties met dezelfde verdeling afkomstig zijn, dan zal men zich vanzelfsprekend afvragen welke verschillen er nu tussen de k verdelingen bestaan. Gewoonlijk zal men deze vraag slechts kunnen beantwoorden door de waargenomen verdelingen te inspecteren en/of door bepaalde verdelingen met elkaar te vergelijken (waartoe dan gewoonlijk weer een χ^2 -toets zal worden gebruikt). Het bezwaar tegen deze procedure is, dat men min of meer subjectief te werk gaat, resp. dat men een aantal toetsingen verricht, die niet onderling onafhankelijk zijn. Toetst men steeds met dezelfde onbetrouwbaarheidsdrempel α , dan kan de onbetrouwbaarheid van de uitspraken die men met het gehele complex van toegepaste toetsen verkrijgt wel aanzienlijk groter worden dan α . Men verkrijgt dan feitelijk een detectieprobleem en in principe zal men de

juistheid van de getrokken conclusies door middel van één of meer nieuwe experimenten dienen te verifiëren. Soms zijn echter de verschillen tussen de verdelingen onderling zo duidelijk en zo sterk significant, dat men toch wel tot verantwoorde conclusies omtrent verschillen tussen de verdelingen onderling kan komen. Wij zullen hiervan enkele voorbeelden geven.

Voorbeelden

24.1. Uit KIMBALL (231). Door GAULDEN is een onderzoek verricht betreffende de invloed van twee soorten straling op de kerndelingssnelheid van de neuroblasten van sprinkhanen (*Chortophaga viridifasciata*). Bij deze proeven werden embryo's uit eenzelfde legsel in drie groepen verdeeld. Eén groep diende als contrôle, de twee andere werden aan fysisch equivalente doses Röntgen- en Betastralen blootgesteld. Na de bestraling werden ongeveer gelijke aantallen cellen uit de drie groepen gedurende een periode van 209 minuten (de gemiddelde duur van de kerndelingscyclus) geobserveerd. Van elke groep werd het aantal cellen geteld waarbij het middenstadium van het delingsproces werd bereikt of gepasseerd. Daar de bestraling de kerndelingssnelheid pleegt te verlagen is dit aantal een indicator voor de mate van beschadiging tengevolge van de ontvangen straling.

Tabel 24.4. De uitkomsten van een bestralingsonderzoek

	Groepen			Totaal
	Röntgenstralen	Betastralen	Contrôle	
+	5	2	11	18
-	17	17	9	43
	22	19	20	61

+ = middenstadium van de mitose bereikt of gepasseerd,
 - = middenstadium niet bereikt.

	Proefgroepen		Totaal
	Röntgen	Beta	
+	5	2	7
-	17	17	34
	22	19	41

	Proef	Contrôle	Totaal
	+	7	11
-	34	9	43
	41	20	61

De uitkomsten van één dezer proeven zijn in tabel 24.4 opgenomen. De χ^2 -toets levert, toegepast op deze 3×2 tabel, $X_0^2 = 10,026$ op. Daar de grootheid \underline{X}^2 onder de nulhypothese bij benadering een χ^2 -verdeling met 2 vrijheidsgraden volgt, bezit deze X_0^2 een overschrijdingskans

die tussen 0,005 en 0,01 ligt, zodat H_0 bij een 1% drempel kan worden verworpen.

De verdere analyse ligt hier voor de hand: men zal eerst willen nagaan, of de twee proefgroepen verschillen. De betreffende 2×2 tabel bevat slechts 7 'plussen', zodat men op de plusregel vrij lage verwachtingen (resp. 3,76 en 3,24) krijgt. Past men toch de χ^2 -benadering toe dan vindt men (zonder continuïteitscorrectie) slechts $X_0^2 = 1,07$ en deze waarde is zo laag, dat men zonder bezwaar kan concluderen, dat het verschil tussen de proefgroepen niet significant is. De 2×2 tabel die men verkrijgt door de proefgroepen gezamenlijk met de contrôle-groep te vergelijken wijst direct uit, dat de proef- en de contrôle-uitkomsten significant verschillen ($X_0^2 = 7,57$, met continuïteitscorrectie).

Hoewel de drie toegepaste toetsen niet onderling onafhankelijk zijn leveren zij een zo duidelijk beeld op, dat de volgende uitspraak gerechtvaardigd is: Röntgen- en Betastralen vertragen beide de kerndelingssnelheid, maar een verschillend effect van deze twee soorten straling is niet aan te tonen. Hierbij dient volledigheidshalve te worden gezegd, dat GAULDEN de proef verschillende malen herhaalde en dat elke afzonderlijke proef tot deze conclusie leidde.

LANCASTER en IRWIN hebben een methode ontwikkeld, waarmee men een X_0^2 van een $k \times r$ tabel kan splitsen in $(k-1)(r-1)$ onderling onafhankelijke componenten, die dus elk 1 vrijheidsgraad bezitten. Voor verdere bijzonderheden verwijzen wij naar het artikel van KIMBALL (231), waarin tevens een aantal berekeningsformules zijn uitgewerkt.

24.2. In tabel 24.5 zijn de (fictieve) uitkomsten opgenomen van een onderzoek, waarbij de resultaten van drie verschillende behandelingen A , B en C worden vergeleken bij drie aselekt samengestelde groepen patiënten. Voor deze 3×3 tabel vindt men bij toepassing van de χ^2 -toets, dat $X_0^2(4) = 37,63$ met $P \ll 0,001$. Beschouwt men de sommen van de waarden $(f_{ij} - e_{ij})^2/e_{ij}$ per behandeling dan blijken deze alle van dezelfde orde van grootte te zijn. Hieruit is op te maken, dat de verdeling bij elke behandeling duidelijk van de totaalverdeling der

Tabel 24.5. Fictieve resultaten van drie verschillende behandelingen

Resultaat	f_{ij}			Totaal	e_{ij}			$f_{ij} - e_{ij}$			$(f_{ij} - e_{ij})^2/e_{ij}$		
	A	B	C		A	B	C	A	B	C	A	B	C
++	21	7	5	33	11	11	11	+10	-4	-6	9,09	1,45	3,27
+	12	8	25	45	15	15	15	-3	-7	+10	0,60	3,27	6,67
0	7	25	10	42	14	14	14	-7	+11	-4	3,50	8,64	1,14
Totaal	40	40	40	120	40	40	40	0	0	0	13,19	13,36	11,08

$$X_0^2(4) = 13,19 + 13,36 + 11,08 = 37,63.$$

behandelingen afwijkt, zodat de verdelingen ook onderling duidelijk moeten verschillen. Gezien de zeer hoge waarde van X_0^2 kan men hier, zonder verdere toetsingen uit te voeren, concluderen dat behandeling *A* betere resultaten oplevert dan behandeling *C* en dat deze behandeling te prefereren is boven *B*.

24.3. Het derde voorbeeld betreft de — eveneens gefingeerde — uitkomsten van een onderzoek naar het voorkomen van diphterietoxine in het speeksel van vier groepen diphteriepatiënten in tabel 24.6. De uitslagen van de toxineproef zijn aangeduid met de categorieën: ++ = duidelijk positief, + = zwak positief, 0 = negatief.

Tabel 24.6. Fictieve resultaten van een toxineproef bij vier groepen patiënten

Tox. proef	f_{ij}				Tot.	$e_{i.}$	$f_{ij} - e_{ij}$				$(f_{ij} - e_{ij})^2/e_{ij}$			
	A	B	C	D			A	B	C	D	A	B	C	D
++	5	7	18	4	34	8,5	-3,5	-1,5	+9,5	-4,5	1,44	0,26	10,62	2,38
+	24	20	22	22	88	22,0	+2,0	-2,0	0	0	0,18	0,18	0	0
0	21	23	10	24	78	19,5	+1,5	+3,5	-9,5	+4,5	0,12	0,63	4,63	1,04
Tot.	50	50	50	50	200	50,0	0	0	0	0	1,74	1,07	15,25	3,42

$$X_0^2(6) = 1,74 + 1,07 + 15,25 + 3,42 = 21,48.$$

De berekende $X_0^2(6) = 21,48$ bezit een overschrijdingskans tussen 0,001 en 0,005. De betrekkelijk kleine sommen van de waarden $(f_{ij} - e_{ij})^2/e_{ij}$ van de groepen *A*, *B* en *D* wijzen er reeds op, dat de verdelingen van deze groepen weinig uiteenlopen. Vergelijkt men deze groepen onderling dan vindt men $X_0^2(4) = 1,44$ met $0,80 < P < 0,90$. De grootste bijdrage tot de totale X_0^2 wordt geleverd door groep *C* met een som van 15,25, zodat men zonder verder te toetsen kan concluderen, dat tussen de groepen *A*, *B* en *D* geen verschillen in de uitkomsten van de toxineproef aan te tonen zijn, maar dat bij groep *C* een verschuiving naar de zijde der duidelijk positieve reacties optreedt.

24.4. Wanneer de toepassing van de χ^2 -toets op een $k \times r$ tabel tot de uitspraak leidt dat de nulhypothese niet kan worden verworpen, dan zijn twee mogelijkheden te onderscheiden: 1. De verdelingen van de k populaties verschillen niet of althans betrekkelijk weinig, 2. Deze verdelingen vertonen onderling wel verschillen, maar de toegepaste toets bezit bij de gegeven omvang der steekproeven of groepen en bij de gekozen onbetrouwbaarheidsdrempel een te gering onderscheidingsvermogen. Gaat men bij een niet significante X_0^2 toch de verdelingen onderling vergelijken, dan neemt men als het ware aan, dat men met (2) te doen heeft. Dit behoeft echter niet het geval te zijn en men loopt

dan bij deze procedure het risico, dat men door *selectie* toevallige verschillen voor werkelijke aanziet.

Beschouw bijvoorbeeld de 5 verdelingen in tabel 24.7. Wij veronderstellen dat zij de uitkomsten bevatten van een onderzoek, waarbij vijf verschillende behandelingen worden vergeleken en dat men a priori geen redenen heeft om aan te nemen, dat één of meer dezer behandelingen duidelijk van de andere afwijken. De te toetsen hypothese luidt dus, dat de waargenomen verschillen toevallig zijn, met als alternatieve hypothese dat wezenlijke verschillen optreden. Men kiest $\alpha = 0,05$. Wegens de gelijke omvang der steekproeven zijn de verwachte frequenties per rij constant en zij bedragen voor de categorieën +, 0 en - resp. 8,4, 5,0 en 6,6. Voor de gehele 5×3 tabel vindt men dan dat $X_0^2 = 10,20$ met $0,2 < P < 0,3$, zodat de nulhypothese niet wordt verworpen.

Tabel 24.7. Een 5×3 tabel

Oorspronkelijke gegevens						Na combinatie		
Categorie	f_{ij}					Totaal	I + II + IV + V	III
	I	II	III	IV	V			
+	10	11	5	7	9	42	37	5
0	6	4	3	6	6	25	22	3
-	4	5	12	7	5	33	21	12
Tot.	20	20	20	20	20	100	80	20

$$X_0^2(8) = 10,20$$

$$X_0^2(2) = 8,24$$

Bij inspectie blijkt echter dat de verdeling van groep III afwijkt van de verdelingen van de vier andere groepen, die onderling vrij goed overeenstemmen. Vergelijkt men op grond van deze bevinding de gecombineerde verdeling van deze groepen met die van groep III door middel van de χ^2 -toets, dan vindt men $X_0^2(2) = 8,24$ met $0,01 < P < 0,02$. Deze significante uitkomst berust echter op een bewuste selectie en daar de juiste toets (op de volledige 5×3 tabel) niet tot het verwerpen van de nulhypothese leidt, kan men de gesignaleerde afwijkende verdeling hoogstens als indicatie gebruiken ten behoeve van een nieuw onderzoek (vergelijk met 9.2.4, blz. 184).

Het behoeft geen betoog dat in werkelijkheid de verschillen of de overeenkomsten tussen de verdelingen vaak minder eenvoudig te duiden zullen zijn dan bij de gegeven voorbeelden. Men dient dan steeds in het oog te houden dat men bij de uitvoering van een aantal onderling niet onafhankelijke toetsingen, alle met dezelfde drempelwaarde α , in het algemeen te veel significante uitkomsten zal verkrijgen. Neem bv. het geval dat men beschikt over de gemiddelden van 5 aselechte steek-

proeven uit pseudo-normaal verdeelde populaties en dat men met $\alpha = 0,05$ de hypothese $H_0: \mu_1 = \mu_2 = \dots = \mu_5$ wil toetsen. Men zou dan deze gemiddelden op $C_2^5 = 10$ manieren twee aan twee kunnen vergelijken door middel van de toets van STUDENT, d.w.z. tienmaal de hypothese $H_0': \mu_i = \mu_j$ kunnen toetsen. Kiest men bij elke toetsing $\alpha = 0,05$ en is de hypothese H_0 (en dus elke H_0') juist, dan is de kans dat men tenminste éénmaal H_0' verworpt (en op grond daarvan tot het verwerpen van H_0 overgaat) gelijk aan $1 - (1 - 0,05)^{10} = 0,40$, dus veel groter dan 0,05! Men kan zich hiertegen wapenen door bij de 'detail'-toetsen een lagere drempelwaarde te gebruiken dan bij de 'overall'-toets van H_0 .

24.3. Toets voor de hypothese $\mu_A - \mu_B = \mu_C - \mu_D$

Beschouw vier (ongeveer) normaal verdeelde populaties A, B, C en D , die dezelfde spreiding σ bezitten. Men wil op grond van vier aselechte steekproeven uit deze populaties, resp. van de omvang n_A, n_B, n_C en n_D , de hypothese $H_0: \mu_A - \mu_B = \mu_C - \mu_D$ toetsen.

Onder H_0 blijkt de grootheid

$$(24.12) \quad t = \frac{(\bar{x}_A - \bar{x}_B) - (\bar{x}_C - \bar{x}_D)}{s \sqrt{\frac{1}{n_A} + \frac{1}{n_B} + \frac{1}{n_C} + \frac{1}{n_D}}}$$

(bij benadering) een STUDENT-verdeling te volgen met

$$n_A + n_B + n_C + n_D - 4$$

vrijheidsgraden. Men kan s^2 berekenen door volgens (13.27) te nemen

$$(24.13) \quad s^2 = \frac{\Sigma x^2 - \left\{ \frac{(\Sigma x_A)^2}{n_A} + \frac{(\Sigma x_B)^2}{n_B} + \frac{(\Sigma x_C)^2}{n_C} + \frac{(\Sigma x_D)^2}{n_D} \right\}}{n_A + n_B + n_C + n_D - 4},$$

waarin Σx^2 gelijk is aan de som van de kwadraten van alle waarnemingen. Wanneer alleen de spreidingen (of varianties) van de vier steekproeven bekend zijn, dient men gebruik te maken van de formule

$$(24.14) \quad s^2 = \frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2 + (n_C - 1)s_C^2 + (n_D - 1)s_D^2}{n_A + n_B + n_C + n_D - 4}.$$

Voor het speciale geval dat $n_A = n_B = n_C = n_D = n$ kan men de waarde van t berekenen met

$$(24.15) \quad t_0 = \frac{(\Sigma x_A + \Sigma x_D) - (\Sigma x_B + \Sigma x_C)}{2s\sqrt{n}} = \frac{\{(\bar{x}_A + \bar{x}_D) - (\bar{x}_B + \bar{x}_C)\}\sqrt{n}}{2s},$$

terwijl verder geldt dat

$$(24.16) \quad s^2 = \frac{n \cdot \Sigma x^2 - [(\Sigma x_A)^2 + (\Sigma x_B)^2 + (\Sigma x_C)^2 + (\Sigma x_D)^2]}{4n(n-1)} = \\ = \frac{s_A^2 + s_B^2 + s_C^2 + s_D^2}{4}.$$

Voorbeeld

24.5. Bij een onderzoek van kinderen van 7, 8 en 9 jaar in een plaats S werd o.m. de alkalische phosphatase (in BODANSKY-eenheden) bepaald van *a*) 30 kinderen uit boerengezinnen, en *b*) 30 kinderen uit gezinnen van middenstanders. De uitkomsten waren: $\bar{x}_A = 6,69$, $\bar{x}_B = 7,97$, $s_A^2 = 3,5091$ en $s_B^2 = 4,7883$. Ongeveer tegelijkertijd werd — eveneens bij kinderen van 7, 8 en 9 jaar — de alkalische phosphatase bepaald in een plaats O en wel bij *c*) 60 kinderen uit boerengezinnen, en *d*) 60 kinderen uit middenstandsgesinnen. Dit onderzoek leverde op:

$$\bar{x}_C = 7,91, \quad \bar{x}_D = 8,51, \quad s_C^2 = 4,5863, \quad s_D^2 = 3,6268.$$

In plaats S bedraagt het verschil tussen de gemiddelden $7,97 - 6,69 = 1,28$ eenheden en in plaats O is dit verschil $8,51 - 7,91 = 0,60$ eenheden. De verschillende groepen kunnen als aselechte steekproeven uit hun populaties worden beschouwd. Gevraagd wordt met $\alpha = 0,05$ de hypothese $H_0: \mu_B - \mu_A = \mu_D - \mu_C$ te toetsen.

De waargenomen varianties verschillen niet significant, zodat de gelijkheidsonderstelling van de populatievarianties aanvaardbaar is. De verdelingen van de alkalische phosphatases zijn alle scheef naar rechts, maar zij blijken geen duidelijke uitbijters te vertonen en de aantallen waarnemingen zijn vrij groot. De toepassing van de zojuist beschreven toets lijkt dus wel gerechtvaardigd en men vindt dan

$$s^2 = \frac{29(3,5091 + 4,7883) + 59(4,5863 + 3,6268)}{2(29) + 2(59)} = 4,1204, \\ s = 2,03$$

en

$$t_0 = \frac{1,28 - 0,60}{2,03 \sqrt{2/30 + 2/60}} = \frac{0,68 \sqrt{10}}{2,03} = 1,06,$$

met $P_D(1,06) = 0,29$. De uitspraak luidt dus dat H_0 niet kan worden verworpen.

De toetsing van H_0 kan ook door middel van een variantie-analyse worden uitgevoerd. Het volgende voorbeeld, waarbij steekproeven van dezelfde omvang zijn gekozen, laat zien hoe deze toetsing verloopt.

Voorbeeld

24.6. De fictieve gegevens in tabel 24.8 betreffen bloeddrukwaarden van ratten. Bij een in de zomer verricht onderzoek zijn de bloeddrukken (in mm Hg) gemeten van 20 ratten, waarbij 10 ratten een leeftijd van 4–6 maanden en 10 ratten een leeftijd van 1½–2 jaar bezaten. Bij een ander onderzoek in de winter werden eveneens bij 10 jonge en bij 10 oude ratten bloeddrukmetingen verricht. De waargenomen gemiddelde bloeddrukken volgen uit de tabel:

	Zomer	Winter
oude ratten	102,4	98,0
jonge ratten	93,0	94,2
verschil	9,4	3,8

Men wil bij een 5% drempel de hypothese H_0 toetsen, dat het gemiddelde verschil (oud-jong) in de zomer en in de winter hetzelfde is, met als alternatieve hypothese H_1 , dat in de zomer een groter verschil optreedt.

Vanzelfsprekend dient voor de geldigheid van de variantie-analyse te worden ondersteld, dat de vier populatieverdelingen van de bloeddrukwaarden ongeveer normaal verdeeld zijn met dezelfde spreiding.

Tabel 24.8. Bloeddrukwaarden (in mm Hg, verminderd met 75) van vier groepen ratten (fictieve gegevens)

Zomer				Winter			
4–6 maanden		1½–2 jaar		4–6 maanden		1½–2 jaar	
x_A	$y_A = \log x_A$	x_B	$y_B = \log x_B$	x_C	$y_C = \log x_C$	x_D	$y_D = \log x_D$
11	1,04	41	1,61	10	1,00	19	1,28
20	1,30	32	1,51	15	1,18	24	1,38
19	1,28	17	1,23	33	1,52	13	1,11
10	1,00	16	1,20	22	1,34	28	1,45
17	1,23	27	1,43	16	1,20	22	1,34
22	1,34	39	1,59	17	1,23	22	1,34
30	1,48	27	1,43	20	1,30	17	1,23
12	1,08	25	1,40	14	1,15	37	1,57
31	1,49	22	1,34	19	1,28	33	1,52
8	0,90	28	1,45	26	1,42	15	1,18
Som 180	12,14	274	14,19	192	12,62	230	13,40
Σy_j^2	15,1054		20,3011		16,1186		18,1492
$(\Sigma y_j)^2/10$	14,73796		20,13561		15,92644		17,9560
S_j	0,36744		0,16549		0,19216		0,1932

$\Sigma y = 12,14 + 14,19 + 12,62 + 13,40 = 52,35$
 $\Sigma y^2 = 15,1054 + 20,3011 + 16,1186 + 18,1492 = 69,6743$

Nu zijn gewoonlijk bloeddrukken enigszins scheef naar rechts verdeeld, zodat men de analyse beter kan uitvoeren op de logaritmen van de bloeddrukwaarden. De vier steekproefvarianties blijken dan ver van significant te verschillen, zodat ook de gelijkheidsonderstelling van de vier populatievarianties gerechtvaardigd lijkt.

De verdere berekeningen verlopen als volgt:

$$S_{\text{totaal}} = 69,6743 - (52,35)^2/40 \dots\dots\dots = 1,1612375$$

$$S_{\text{tussen}} = \frac{(12,24)^2 + (14,19)^2 + \dots + (13,40)^2}{10} - \frac{(52,35)^2}{40} = 0,2429475$$

$$S_{\text{binnen}} = 0,36744 + 0,16549 + 0,19216 + 0,19320 \dots = 0,9182900$$

De kwadratensom tussen de groepen bezit 3 vrijheidsgraden. Met behulp van de volgende tabel kan zij in drie componenten met elk 1 vrijheidsgraad worden gesplitst:

		Seizoen		
		Zomer	Winter	
Leeftijd	Jong	12,14	12,62	24,76
	Oud	14,19	13,40	27,59
		26,33	26,02	52,35

$$S_{\text{seizoen}} = \frac{(26,33)^2 + (26,02)^2}{20} - \frac{(52,35)^2}{40} = 0,0024025$$

$$S_{\text{leeftijd}} = \frac{(24,76)^2 + (27,59)^2}{20} - \frac{(52,35)^2}{40} = 0,2002225$$

$$\underline{0,2026250}$$

$$S_{\text{interactie}} = (\text{seizoen} \times \text{leeftijd}) \dots\dots\dots = 0,0403225$$

$$S_{\text{tussen}} \dots\dots\dots = 0,2429475$$

De volledige uitkomsten van de variantie-analyse zijn in tabel 24.9 verzameld. De hypothese $H_0: \mu_B - \mu_A = \mu_D - \mu_C$ houdt in, dat er tussen het seizoen en de leeftijd *geen interactie* aanwezig is en deze hypothese kan worden getoetst via de grootheid

$$F(1, 36) = s^2_{\text{interactie}}/s^2_{\text{binnen}}$$

Bij het voorbeeld vindt men dan dat $F_0 = 0,0403225/0,0255081 = 1,5808$. Daar $F_{0,95}(1, 36)$ ongeveer gelijk is aan 4,11 kan H_0 dus niet worden verworpen. (Het is duidelijk dat het verschil tussen de leeftijden significant is, daar $F_0 = 0,2002225/0,0255081 = 7,85$.)

Tabel 24.9. Uitkomsten van de variantie-analyse op de gegevens in tabel 24.8

	S	V.g.	Var.
Seizoenen	0,0024025	1	0,0024025
Leeftijden	0,2002225	1	0,2002225
Interactie S × L	0,0403225	1	0,0403225
Tussen de groepen	0,2429475	3	
Binnen de groepen	0,9182900	36	0,0255081
Totaal	1,1612375	39	

Past men de t -toets toe, dan volgt uit (24.16) dat

$$s^2 = \frac{0,91829}{36} = 0,0255081, \quad s = 0,15971,$$

zodat volgens (24.15)

$$t_0 = \frac{(14,19 + 12,62) - (12,14 + 13,40)}{2 (0,15971) \sqrt{10}} = 1,2573.$$

Zoals men ziet geldt, dat $F_0(1, 36) = 1,5808 = t_0^2(36) = (1,2573)^2$.

BIJLAGE I — TABELLENVERZAMELING (vervolg)

T.	Fractielen van χ^2 -verdelingen.	636
U.	De z-transformatie van FISHER: $z = \frac{1}{2} \log_e(1 + r)/(1 - r) = 1,1513 [\log_{10}(1 + r) - \log_{10}(1 - r)]$. . .	638
V.	De normaliteitstoetsen van FISHER Linker en rechter kritieke waarden van de grootheden g_1 en g_2 bij één- zijdige toetsing en $\alpha = 0,05$ en $0,01$	640
W.	Betrouwbaarheidsgrenzen van het <i>percentage</i> elementen met kenmerk <i>A</i> in de populatie, indien x van de n elementen in een aselechte steek- proef dit kenmerk bezitten.	641
X.	95% betrouwbaarheidsgrenzen van het <i>percentage</i> elementen met kenmerk <i>A</i> in de populatie, indien a <i>procent</i> van de elementen in een aselechte steekproef van de omvang n dit kenmerk bezit.	642
Y.	De uitbijtertoets van DIXON Rechter kritieke waarden van de grootheid y voor $\alpha = 0,10, 0,05$ en $0,01$	643
Z.	Rechter kritieke waarden van de grootheid \underline{y} voor $\alpha = 0,05$ en $0,01$	644
AA.	De bgsin transformatie 1. Waarden van de grootheid $\underline{y} = 2 \text{ bgsin } \sqrt{p}$ 646 2. Waarden van $y(0)$ bij $p = 0$ en van $y(1)$ bij $p = 1$ voor $10 < n < 50$	648
BB.	De vierkantswortel transformatie van FREEMAN en TUKEY Waarden van $y = \sqrt{x} + \sqrt{x+1}$ en van $y^2 = (\sqrt{x} + \sqrt{x+1})^2$. .	648
CC.	Lotingsgetallen (aselechte permutaties)	649

TABEL T. FRACTIELEN VAN χ^2 -VERDELINGEN

ν	$\chi^2_{0,005}$	$\chi^2_{0,01}$	$\chi^2_{0,025}$	$\chi^2_{0,05}$	$\chi^2_{0,10}$	$\chi^2_{0,25}$	$\chi^2_{0,50}$
1	0,04393	0,03157	0,03982	0,02393	0,01579	0,1015	0,4549
2	0,01003	0,02010	0,05064	0,10259	0,2107	0,5754	1,3863
3	0,0717	0,1148	0,2158	0,3518	0,5844	1,2125	2,3660
4	0,2070	0,2971	0,4844	0,7107	1,0636	1,92255	3,3567
5	0,4117	0,5543	0,8312	1,1455	1,6103	2,6746	4,3515
6	0,6757	0,8721	1,2373	1,6354	2,2041	3,4546	5,3481
7	0,9893	1,2390	1,6899	2,16735	2,8331	4,25485	6,3458
8	1,3444	1,6465	2,1797	2,7326	3,4895	5,0706	7,3441
9	1,7349	2,0879	2,7004	3,3251	4,1682	5,8988	8,3428
10	2,1559	2,5582	3,2470	3,9403	4,8652	6,7372	9,3418
11	2,6032	3,0535	3,8158	4,5748	5,5778	7,5841	10,3410
12	3,0738	3,5706	4,4038	5,2260	6,3038	8,4384	11,3403
13	3,5650	4,1069	5,0087	5,8919	7,0415	9,2991	12,3398
14	4,0747	4,6604	5,6287	6,5706	7,7895	10,1653	13,3393
15	4,6009	5,2294	6,2621	7,2609	8,5468	11,0365	14,3389
16	5,1422	5,8122	6,9077	7,9616	9,3122	11,9122	15,3385
17	5,6972	6,4078	7,5642	8,6718	10,0852	12,7919	16,3381
18	6,2648	7,0149	8,2308	9,3905	10,8649	13,6753	17,3379
19	6,8440	7,6327	8,9066	10,1170	11,6509	14,5620	18,3376
20	7,4339	8,2604	9,5908	10,8508	12,4426	15,4518	19,3374
21	8,0337	8,8972	10,2829	11,5913	13,2396	16,3444	20,3372
22	8,6427	9,5425	10,9823	12,3380	14,0415	17,2396	21,3370
23	9,2604	10,1957	11,6885	13,0905	14,8479	18,1373	22,3369
24	9,8862	10,8564	12,4001	13,8484	15,6587	19,0372	23,3367
25	10,5197	11,5240	13,1197	14,6114	16,4734	19,9393	24,3366
26	11,1603	12,1981	13,8439	15,3791	17,2919	20,8434	25,3364
27	11,8076	12,8786	14,5733	16,1513	18,1138	21,7494	26,3363
28	12,4613	13,5648	15,3079	16,9279	18,9392	22,6572	27,3363
29	13,1211	14,2565	16,0471	17,7083	19,7677	23,5666	28,3362
30	13,7867	14,9535	16,7908	18,4926	20,5992	24,4776	29,3360
40	20,7065	22,1643	24,4331	26,5093	29,0505	33,6603	39,3354
50	27,9907	29,7067	32,3574	34,7642	37,6886	42,9421	49,3349
60	35,5346	37,4848	40,4817	43,1879	46,4589	52,2938	59,3347
70	43,2752	45,4418	48,7576	51,7393	55,3290	61,6983	69,3344
80	51,1720	53,5400	57,1532	60,3915	64,2778	71,1445	79,3343
90	59,1963	61,7541	65,6466	69,1260	73,2912	80,6247	89,3342
100	67,3276	70,0648	74,2219	77,9295	82,3581	90,1332	99,3341

Uit: C. M. THOMPSON (203).

TABEL T (vervolg)

ν	$\chi^2_{0,75}$	$\chi^2_{0,90}$	$\chi^2_{0,95}$	$\chi^2_{0,975}$	$\chi^2_{0,99}$	$\chi^2_{0,995}$
1	1,3233	2,7055	3,8415	5,0239	6,6349	7,8794
2	2,7726	4,6052	5,9915	7,3778	9,2103	10,5966
3	4,10835	6,2514	7,8147	9,3484	11,3449	12,8381
4	5,3853	7,7794	9,4877	11,1433	13,2767	14,8602
5	6,6257	9,23635	11,0705	12,8325	15,0863	16,7496
6	7,8408	10,6446	12,5916	14,4494	16,8119	18,5476
7	9,0372	12,0170	14,0671	16,0128	18,4753	20,2777
8	10,2188	13,3616	15,5073	17,5346	20,0902	21,9550
9	11,3887	14,6837	16,9190	19,0228	21,6660	23,5893
10	12,5489	15,9871	18,3070	20,4831	23,2093	25,1882
11	13,7007	17,2750	19,6751	21,9200	24,7250	26,7569
12	14,8454	18,5494	21,0261	23,3367	26,2170	28,2995
13	15,9839	19,8119	22,3621	24,7356	27,6883	29,8194
14	17,1170	21,0642	23,6848	26,1190	29,1413	31,3193
15	18,2451	22,3072	24,9958	27,4884	30,5779	32,8013
16	19,3688	23,5418	26,2962	28,8454	31,9999	34,2672
17	20,4887	24,7690	27,5871	30,1910	33,4087	35,7185
18	21,6049	25,9894	28,8693	31,5264	34,8053	37,1564
19	22,7178	27,2036	30,1435	32,8523	36,1908	38,5822
20	23,8277	28,4120	31,4104	34,1696	37,5662	39,9968
21	24,9348	29,6151	32,6705	35,4789	38,9321	41,4010
22	26,0393	30,8133	33,9244	36,7807	40,2894	42,7956
23	27,1413	32,0069	35,1725	38,0757	41,6384	44,1813
24	28,2412	33,1963	36,4151	39,3641	42,9798	45,5585
25	29,3389	34,3816	37,6525	40,6465	44,3141	46,9278
26	30,4345	35,5631	38,8852	41,9232	45,6417	48,2899
27	31,5284	36,7412	40,1133	43,1944	46,9630	49,6449
28	32,6205	37,9159	41,3372	44,4607	48,2782	50,9933
29	33,7109	39,0875	42,5569	45,7222	49,5879	52,3356
30	34,7998	40,2560	43,7729	46,9792	50,8922	53,6720
40	45,6160	51,8050	55,7585	59,3417	63,6907	66,7659
50	56,3336	63,1671	67,5048	71,4202	76,1539	79,4900
60	66,9814	74,3970	79,0819	83,2976	88,3794	91,9517
70	77,5766	85,5271	90,5312	95,0231	100,425	104,215
80	88,1303	96,5782	101,879	106,629	112,329	116,321
90	98,6499	107,565	113,145	118,136	124,116	128,299
100	109,141	118,498	124,342	129,561	135,807	140,169

Voor $30 < \nu < 100$ kan lineair worden geïnterpoleerd. Men verkrijgt dan een uitkomst met 4 betrouwbare cijfers.

TABEL U. DE z -TRANSFORMATIE VAN FISHER:

$$z = \frac{1}{2} \log_e \frac{1+r}{1-r} = 1,1513 [\log_{10}(1+r) - \log_{10}(1-r)]$$

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	0,120561	0,121596	0,122611	0,123626	0,124642	0,125657	0,126673	0,127690	0,128708	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	0,171687	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

TABEL U (vervolg)

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	0,677866	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	0,922333	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	1,082414	1,085284	1,088189	1,091133	1,094082	1,097041
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	1,191368	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	1,533873	1,539285	1,545261	1,551302	1,557411	1,563590	1,569839	1,576190	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	1,665884	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	2,043850	2,059519	2,075648
0,97	2,092296	2,109501	2,127300	2,145738	2,164861	2,184723	2,205389	2,226922	2,249400	2,272913
0,98	2,297561	2,323460	2,350746	2,379577	2,410142	2,442663	2,477141	2,514717	2,554990	2,598747
0,99	2,646653	2,695958	2,758727	2,825744	2,903070	2,994482	3,106304	3,250396	3,453379	3,800203
1,00	∞									

TABEL V. DE NORMALITEITSTOETSEN VAN FISHER
 Linker en rechter kritieke waarden van de grootheden g_1 en g_2 bij éézijdige
 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 (4).

De linker kritieke waarden van g_1 zijn numeriek gelijk aan de rechter kritieke waarden, maar zij bezitten een tegengesteld teken.

TABEL W.

Betrouwbaarheidsgrenzen van het *percentage* elementen met kenmerk *A* in de populatie, indien x van de n elementen in een aselechte 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 — 7,6	0,82 — 5,7	0,54 — 3,8	0,32 — 2,3	0,16 — 1,2
6	2,2 — 12,6	1,5 — 8,5	1,1 — 6,4	0,74 — 4,3	0,44 — 2,6	0,22 — 1,3
7	2,9 — 13,9	1,9 — 9,4	1,4 — 7,1	0,94 — 4,7	0,56 — 2,9	0,28 — 1,4
8	3,5 — 15,2	2,3 — 10,2	1,7 — 7,7	1,2 — 5,2	0,69 — 3,1	0,35 — 1,6
9	4,2 — 16,4	2,8 — 11,1	2,1 — 8,4	1,4 — 5,6	0,83 — 3,4	0,41 — 1,7
10	4,9 — 17,6	3,2 — 11,9	2,4 — 9,0	1,6 — 6,0	0,96 — 3,6	0,48 — 1,8

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	0,13 — 2,5	0,067 — 1,3
5	1,1 — 13,5	0,73 — 9,2	0,54 — 6,9	0,36 — 4,6	0,22 — 2,8	0,11 — 1,4
6	1,6 — 14,9	1,0 — 10,1	0,78 — 7,6	0,52 — 5,1	0,31 — 3,1	0,15 — 1,6
7	2,1 — 16,3	1,4 — 11,0	1,0 — 8,4	0,68 — 5,6	0,41 — 3,4	0,20 — 1,7
8	2,6 — 17,6	1,7 — 12,0	1,3 — 9,0	0,86 — 6,1	0,52 — 3,7	0,26 — 1,8
9	3,2 — 18,9	2,1 — 12,9	1,6 — 9,7	1,1 — 6,5	0,63 — 4,0	0,31 — 2,0
10	3,8 — 20,2	2,5 — 13,7	1,9 — 10,4	1,2 — 7,0	0,75 — 4,2	0,37 — 2,1

Uit MAINLAND (45).

TABEL X.

95% betrouwbaarheidsgrenzen van het percentage elementen met kenmerk A in de populatie, indien a procent van de elementen in een aselechte steekproef van de omvang n dit kenmerk bezit

a	n			
	100	200	500	1000
0	*	*	*	*
1	*	*	*	*
2	*	*	*	*
3	*	*	1 — 5	2 — 5
4	*	*	2 — 7	2 — 6
5	*	*	3 — 8	3 — 7
6	*	3 — 11	4 — 9	4 — 8
7	*	3 — 12	4 — 10	5 — 9
8	*	4 — 13	5 — 11	6 — 10
9	*	5 — 15	6 — 12	7 — 11
10	*	6 — 16	7 — 14	8 — 13
15	8 — 24	10 — 21	12 — 19	12 — 18
20	12 — 30	14 — 27	16 — 24	17 — 23
25	16 — 35	19 — 32	21 — 30	22 — 38
30	21 — 40	23 — 37	26 — 35	27 — 33
35	25 — 46	28 — 43	30 — 40	32 — 39
40	30 — 51	33 — 48	35 — 45	36 — 44
45	35 — 56	38 — 53	40 — 50	41 — 49
50	39 — 61	42 — 58	45 — 55	46 — 54

* = zie tabel W.

Uit Rümke (193). Voor $n = 100$ zijn de intervallen exact berekend. Voor $n > 100$ en $an > 2000$ werd de normale en voor $an < 2000$ de Poissonbenadering toegepast.

TABEL Y. DE UITBIJERTOETS VAN DIXON
 Rechter kritieke waarden van de grootheid r

Toetsings- grootheid	n	Onbetrouwbaarheidsdrempel α			
		0,10	0,05	0,01	
r_{10}	3	0,886	0,941	0,988	
	4	0,679	0,765	0,889	
	5	0,557	0,642	0,780	
	6	0,482	0,560	0,698	
	7	0,434	0,507	0,637	
	8	0,399	0,468	0,590	
	9	0,370	0,437	0,555	
	10	0,349	0,412	0,527	
	r_{22}	11	0,578	0,637	0,745
		12	0,543	0,600	0,704
13		0,515	0,570	0,670	
14		0,492	0,546	0,641	
15		0,472	0,525	0,616	
16		0,454	0,507	0,595	
17		0,438	0,490	0,577	
18		0,424	0,475	0,561	
19		0,412	0,462	0,547	
20		0,401	0,450	0,535	
21		0,391	0,440	0,524	
22		0,382	0,430	0,514	
23		0,374	0,421	0,505	
24		0,367	0,413	0,497	
25		0,360	0,406	0,489	
26		0,354	0,399	0,482	
27		0,348	0,393	0,475	
28		0,342	0,387	0,469	
29		0,337	0,381	0,463	
30		0,332	0,376	0,457	

Uit DIXON (145).

TABEL Z. RECHTER KRITIEKE WAARDEN VAN DE GROOTHEID w

$w = \bar{R} \sqrt{n}/s$, waarin $\bar{R} = \bar{x}_{\max} - \bar{x}_{\min}$ (in een reeks van h gemiddelden) en $s =$ een van \bar{R} onafhankelijke schatting van σ met ν vrijheidsgraden

$\alpha = 0,05$

$\nu \backslash h$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	18,0	26,7	32,8	37,2	40,5	43,1	45,4	47,3	49,1	50,6	51,9	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	5,88	6,83	7,51	8,04	8,47	8,85	9,18	9,46	9,72	9,95	10,16	10,35	10,52	10,69	10,84	10,98	11,12	11,24
4	3,93	5,00	5,76	6,31	6,73	7,06	7,35	7,60	7,83	8,03	8,21	8,37	8,52	8,67	8,80	8,92	9,03	9,14	9,24
5	3,61	4,54	5,18	5,64	5,99	6,28	6,52	6,74	6,93	7,10	7,25	7,39	7,52	7,64	7,75	7,86	7,95	8,04	8,13
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	7,04	7,14	7,24	7,34	7,43	7,51	7,59
7	3,34	4,16	4,68	5,06	5,35	5,59	5,80	5,99	6,15	6,29	6,42	6,54	6,65	6,75	6,84	6,93	7,01	7,08	7,16
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	6,65
10	3,15	3,88	4,33	4,66	4,91	5,12	5,30	5,46	5,60	5,72	5,83	5,93	6,03	6,12	6,20	6,27	6,34	6,41	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	6,27	6,33
12	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,40	5,51	5,61	5,71	5,80	5,88	5,95	6,02	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	6,06	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	5,56	5,64	5,72	5,79	5,86	5,92	5,98	6,03
15	3,01	3,67	4,08	4,37	4,59	4,78	4,94	5,08	5,20	5,31	5,40	5,49	5,57	5,65	5,72	5,79	5,85	5,91	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	5,73	5,79	5,84	5,90
17	2,98	3,62	4,02	4,31	4,52	4,70	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	4,83	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	4,64	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	5,50	5,56	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	5,55	5,59
30	2,89	3,48	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	4,90	4,98	5,05	5,11	5,17	5,22	5,27	5,32	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	5,15	5,20	5,24
120	2,80	3,36	3,69	3,92	4,10	4,24	4,36	4,47	4,56	4,64	4,71	4,78	4,84	4,90	4,95	5,00	5,04	5,09	5,13
∞	2,77	3,32	3,63	3,86	4,03	4,17	4,29	4,39	4,47	4,55	4,62	4,68	4,74	4,80	4,84	4,89	4,93	4,97	5,01

TABEL Z (vervolg)

$\alpha = 0,01$

$\gamma \backslash k$	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	291	295	298
2	14,0	18,9	22,3	24,7	26,6	28,2	29,5	30,7	31,7	32,6	33,4	34,2	34,8	35,5	36,0	36,5	37,0	37,5	38,0
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	8,08	9,17	9,97	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	5,62	6,83	7,65	8,26	8,73	9,12	9,46	9,76	10,02	10,25	10,46	10,65	10,83	10,99	11,14	11,29	11,42	11,54	11,66
6	5,24	6,32	7,03	7,56	7,97	8,31	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	4,94	5,89	6,52	6,98	7,35	7,65	7,91	8,14	8,34	8,52	8,68	8,82	8,96	9,08	9,20	9,31	9,42	9,52	9,61
8	4,74	5,63	6,20	6,63	6,96	7,24	7,47	7,68	7,86	8,03	8,18	8,31	8,44	8,55	8,66	8,76	8,86	8,95	9,03
9	4,60	5,42	5,96	6,35	6,66	6,91	7,13	7,33	7,50	7,65	7,79	7,91	8,02	8,13	8,23	8,33	8,41	8,50	8,58
10	4,48	5,26	5,77	6,14	6,43	6,67	6,88	7,06	7,22	7,36	7,49	7,60	7,71	7,81	7,91	7,99	8,08	8,15	8,23
11	4,39	5,14	5,62	5,98	6,25	6,47	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	7,60	7,67	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	7,35	7,42	7,49	7,55
14	4,21	4,89	5,32	5,64	5,88	6,08	6,26	6,41	6,54	6,66	6,77	6,87	6,96	7,05	7,13	7,20	7,27	7,34	7,40
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	6,85	6,93	7,00	7,07	7,14	7,20	7,26
16	4,13	4,78	5,19	5,49	5,72	5,91	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	4,73	5,14	5,43	5,66	5,85	6,01	6,15	6,27	6,38	6,48	6,57	6,66	6,73	6,81	6,87	6,94	7,00	7,05
18	4,07	4,70	5,09	5,38	5,60	5,79	5,95	6,08	6,20	6,31	6,41	6,50	6,58	6,65	6,73	6,79	6,85	6,91	6,97
19	4,05	4,66	5,05	5,34	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	4,63	5,02	5,30	5,51	5,69	5,84	5,97	6,09	6,19	6,28	6,37	6,45	6,52	6,59	6,66	6,71	6,77	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	6,57	6,61
30	3,89	4,45	4,80	5,05	5,24	5,40	5,53	5,65	5,76	5,85	5,93	6,01	6,08	6,14	6,20	6,26	6,31	6,36	6,41
40	3,82	4,36	4,70	4,93	5,11	5,26	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	5,78	5,83	5,88	5,93	5,98	6,01
120	3,70	4,20	4,50	4,71	4,87	5,00	5,12	5,21	5,30	5,38	5,44	5,50	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	5,53	5,57	5,61	5,64

TABEL AA. DE BGSIN TRANSFORMATIE

1. Waarden van de grootheid $y = 2bgsin \sqrt{p}$

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 AA. 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 (67).

Voor $0,501 < p < 0,999$ kan men nemen: $y = 3,1416 - 2 \operatorname{bgsin} \sqrt{1-p}$.
 Voorbeeld: $p = 0,881$, $y = 3,1416 - 2 \operatorname{bgsin} \sqrt{0,119} = 3,1416 - 0,7044 = 2,4372$.

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	$y(0)$	$y(1)$
10	0,3176	2,8240	20	0,2241	2,9175	30	0,1828	2,9588
11	0,3027	2,8338	21	0,2187	2,9229	31	0,1798	2,9617
12	0,2897	2,8519	22	0,2136	2,9280	32	0,1770	2,9646
13	0,2782	2,8683	23	0,2089	2,9327	33	0,1743	2,9673
14	0,2681	2,8735	24	0,2045	2,9371	34	0,1717	2,9699
15	0,2589	2,8827	25	0,2003	2,9413	35	0,1692	2,9724
16	0,2507	2,8909	26	0,1964	2,9452	36	0,1669	2,9747
17	0,2431	2,8985	27	0,1927	2,9488	37	0,1646	2,9770
18	0,2363	2,9053	28	0,1893	2,9523	38	0,1624	2,9792
19	0,2299	2,9117	29	0,1860	2,9556	39	0,1603	2,9813
						40	0,1583	2,9833
						41	0,1563	2,9853
						42	0,1545	2,9871
						43	0,1526	2,9889
						44	0,1509	2,9907
						45	0,1492	2,9924
						46	0,1476	2,9940
						47	0,1460	2,9956
						48	0,1445	2,9971
						49	0,1430	2,9986
						50	0,1415	3,0001

Uit: EISENHART (8a).

TABEL BB. DE VIERKANTSWORTEL TRANSFORMATIE VAN FREEMAN EN TURKEY

$$\text{Waarden van } y = \sqrt{x + \sqrt{x + 1}} \text{ en van } y^2 = (\sqrt{x + 1})^2$$

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	146,9264
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	165,9944
17	8,37	70,0669	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 CC. LOTINGSGETALLEN

(Aselecte permutaties)

1 - 1 t/m 10

Serie 3 6 10 2 8	Serie 8 10 3 4 7	Serie 3 1 5 7 4
1 9 1 4 5 7	2 6 5 1 9 2	3 8 10 9 2 6
Serie 9 3 4 10 6	Serie 7 3 5 10 9	Serie 9 10 8 1 3
4 2 7 1 5 8	5 2 8 4 1 6	6 2 7 4 6 5

2 - 1 t/m 20

Serie 06 16 09 07 19	Serie 17 01 08 13 07	Serie 08 06 19 13 03
1 03 18 12 11 01	2 10 04 05 09 14	3 05 11 02 01 18
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
Serie 20 09 10 14 11	Serie 06 04 16 15 08	Serie 05 03 09 16 06
4 05 19 03 06 04	5 19 13 12 17 03	6 02 17 08 15 13
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

Serie 42 16 32 40 17	Serie 37 21 38 30 19	Serie 22 38 27 02 35
1 36 03 23 41 10	2 02 43 01 04 20	3 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
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

Serie 81 74 67 95 70	Serie 03 16 15 40 71	Serie 84 05 56 07 91
1 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 (28)

LITERATUUROPGAVE

(vervolg)

DEEL II-C

118. ALCHIAN, A. A. - Analysis Procedures, Army Air Forces Aviation Psychology Program, Report 18, Washington, 1947.
119. ARMITAGE, P. en I. ALLEN - Methods of estimating the LD50 in quantal response data, *J. of Hygiene*, 48 (1950), 298-322.
120. ASPIN, A. A. - An examination and further development of a formula arising in the problem of comparing two mean values, *Biometrika*, 35 (1948), 88-96.
121. ASPIN, A. A. - Tables for use in comparisons whose accuracy involves two variances, separately estimated, *Biometrika*, 36 (1949), 290-296.
122. BARTLETT, M. S. - The effect of non-normality on the *t* distribution, *Proc. Cambr. phil. Soc.*, 31 (1935), 223-231.
123. BARTLETT, M. S. - Properties of sufficiency and statistical tests, *Proc. Royal Society of London*, 160 A (1937), 268-282.
124. BARTLETT, M. S. - Some examples of statistical methods of research in agriculture and applied biology, Supplement to the *J. Royal Stat. Soc.*, 4 (1937), 137-183.
125. BARTLETT, M. S. - The use of transformations, *Biometrics*, 3 (1947), 39-52.
126. BELLINSON, H. R., J. VON NEUMANN, R. H. KENT en B. I. HART - The mean square successive difference, *Annals Math. Stat.*, 12 (1941), 153-162.
127. BENEPE, O. J. - The sensitivity of *t* and *F* to departures from normality, University of Washington, Master's Thesis.
128. BEVERIDGE, W. I. B. - The art of scientific investigation, Heinemann Ltd, London, 1951.
129. BEZEM, Ir J. J. - Toepassing van de transformatie $y = 2bgsin \sqrt{p}$ bij de bepaling van ED_{50} . Verslag v. h. symposium over ED_{50} bepalingen, Utrecht, 1958.
130. BLISS, C. I. - The statistics of bioassay, Academic Press, New York, 1952.
131. BOX, G. E. P. - Non-normality and tests on variances, *Biometrika*, 40 (1953), 318-335.
132. BROSS, I. - Estimates of the LD_{50} : A critique, *Biometrics*, 6 (1950), 413-423.
133. BROWNLEE, K. A. - Industrial experimentation, Chemical Publishing Co., 3rd ed., New York, 1949.
134. BROWNLEE, K. A., J. L. HODGES en M. ROSENBLATT - The up-and-down method with small samples, *J. Am. Stat. Ass.*, 48 (1953), 262-272.
135. BURN, J. H. - Biological standardization, Oxford University Press, London, 1950, 2nd ed.
136. CASIER, H., A. L. DELAUNOIS en C. HEYMANS - Biologische standaardisatie van geneesmiddelen, Standaard-Boekhandel, Antwerpen, 1948.
137. COCHRAN, W. G. - The distribution of the largest of a set of estimated variances as a fraction of their total, *Ann. of Eugenics* 11 (1941), 47-52.

LITERATUUROPGAVE

138. COCHRAN, W. G. - Some consequences when the assumptions for analysis of variance are not satisfied, *Biometrics* 3 (1947), 22.
139. DANTZIG, Prof. Dr D. VAN - De verantwoordelijkheden van de statisticus, *Statistica*, 7 (1953), 199-208.
140. DAVID, F. N. - Tables of the correlation coefficient, Cambridge University Press, Cambridge, 1938.
141. DAVIES, O. L. - Statistical methods in research and production, Oliver & Boyd, Edinburgh, 1949.
142. DAVIES O. L., H. C. HAMAKER, F. J. VAN DUN - Design and analysis of industrial experiments, *Stat. Neerl.* 9 (1955), 189-207.
143. DIXON, W. J. en A. M. MOOD - A method for obtaining and analyzing sensitivity data, *J. Am. Stat. Ass.*, 43 (1948), 109-126.
144. DIXON, W. J. - Analysis of extreme values, *Annals Math. Stat.*, XXI (1950), 488-506.
145. DIXON, W. J. - Ratio's involving extreme values, *Annals Math. Stat.*, XXII (1951), 68-78.
146. DIXON, W. J. - Processing data for outliers, *Biometrics*, 9 (1953), 74-88.
147. DOORNBOS, R. - Significance of the smallest of a set of estimated normal variances, *Statistica Neerl.*, 10 (1956), 117-126.
148. DOORNBOS, R. - Statistische methoden voor het aanwijzen van uitbijters, *Mededelingen Medisch-Biologische Sectie V. v. S.*, 1959, nr. 8.
149. DRION, Dr E. F. - Waarschijnlijkheidsrekening en biologie, Vacantie-cursus 1950: Waarschijnlijkheidsrekening, Mathematisch Centrum, Amsterdam.
150. DUNCAN D. B. - Multiple range and multiple *F* tests, *Biometrics* 11 (1955), 1-42.
151. EMMENS, C. W. - Principles of biological assay, Chapman & Hall, London, 1948.
152. FINNEY, D. J. - Probit analysis, Cambridge University Press, London, 1952, 2nd ed.
153. FINNEY, D. J. - Statistical method in biological assay, Ch. Griffin & Comp., London, 1952.
154. FISHER, R. A. - Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population, *Biometrika* 10 (1915), 507.
155. FISHER, R. A. - On the 'probable error' of a coefficient of correlation deduced from a small sample, *Metron* 1(1921), 1-32.
156. FRASER, D. A. F. - Statistics, An introduction, John Wiley & Sons, New York, 1958.
157. FREEMAN, M. F. & J. W. TUKEY - Transformations related to the angular and the square root, *Ann. Math. Stat.*, 21 (1950), 607-611.
158. GEARY, R. C. & E. S. PEARSON - Tests of normality, *Biometrika Office, University College, London, 1938.*
- 159a. GEARY, R. C. - Moments of the ratio of the mean deviation to the standard deviation for normal samples, *Biometrika*, 28 (1936), 295.
- 159b. GEARY, R. C. - Testing for normality, *Biometrika*, 34 (1947), 209-242.
160. GRUBBS, F. E. - Sample criteria for testing outlying observations, *Annals Math. Stat.*, XXI, (1950), 27-58.
161. HAMAKER, Dr. H. C. - Naar efficiënte experimenten, *Stat. Neerl.* 9 (1955), 7-25.
162. HAMAKER, Dr. H. C. - Proefopzetten in de industrie, *Stat. Neerl.* 9 (1955), 209-232.
163. HAMAKER, Dr. H. C. - Statistiek en experiment, *Stat. Neerl.* 12 (1958), 119-130.

BIJLAGE III

164. HART, B. I. – Significance levels for the ratio of the mean square successive difference to the variance, *Annals Math. Stat.*, 13 (1942), 445–447.
165. HARTLEY, H. O. – The maximum *F*-ratio as a short-cut test for heterogeneity of variance, *Biometrika* 37 (1950), 308–312.
166. HEMELRIJK, Prof. Dr. J. – Het begrip nauwkeurigheid, *Mededelingen Medisch-Biologische Sectie V. v. S.*, nr. 4, 1958.
167. HEMELRIJK, Prof. Dr. J. – Statistische proefopzetten; bewijs en detectie, *Stat. Neerl.* 12 (1958), 111–118.
168. HEY, B. B. – A new method of experimental sampling illustrated on certain non-normal populations, *Biometrika* 30 (1938), 68–80.
169. HOEL, Paul G. – Introduction to mathematical statistics, John Wiley & Sons, New York, 1951.
170. HOTELING, H. – The impact of R. A. FISHER on statistics, *J. Am. Stat. Ass.*, 46 (1951), 35–46.
171. INGLE, D. J. – Principles of research in biology and medicine, Lippincott, Philadelphia, 1958.
172. KÄRBER, G. – Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche, *Arch. f. Exp. Path. u. Pharm.*, 162 (1931), 480–487.
173. KENDALL, M. G. – The advanced theory of statistics, I & II, Griffin & Co., London, 1946.
174. LITCHFIELD, J. T. en F. WILCOXON – A simplified method of evaluating dose-effects experiments, *J. Pharmacol.*, 95 (1949), 99–113.
175. LORD, E. – The use of the range in place of the standard deviation in the *t*-test, *Biometrika* 34 (1947), 41.
176. MADOW, W. G. – Note on tests of departure from normality, *J. Am. Stat. Ass.*, 35 (1940), 515–517.
177. MATHER, K. – R. A. Fisher's 'Statistical methods for research workers', *J. Am. Stat. Ass.*, 46 (1951), 51–54.
178. MAXWELL, A. E. – Experimental design in psychology and the medical sciences, Methuen & Co, London, 1958.
179. MAY, Joyce M. – Extended and corrected tables of the upper percentage points of the 'Studentized' range, *Biometrika*, 39 (1952), 192–193.
180. MERRINGTON, M. & C. M. THOMPSON – Tables for testing the homogeneity of a set of estimated variances, *Biometrika*, 33 (1946), 296–304.
181. NAIR, K. R. – The distribution of the extreme deviate from the sample mean and its Studentized form, *Biometrika* 35 (1948), 118–144.
182. NASS, C. A. G. – Het quantitative verband tussen dosering en letaliteit bij proefdieren, *Statistica*, 1 (1946–47), 257–266.
183. NASS, C. A. G. – The correlation coefficient as a test criterion, *Cursus Gezondheidszorg en Praeventieve Geneeskunde, Syllabus statistiek*, 1955.
184. NEUMANN, J. VON – Distribution of the ratio of the mean square successive difference to the variance, *Annals Math. Stat.*, 12 (1941), 307–395.
185. PEARSON, E. S. – The analysis of variance in cases of non-normal variation, *Biometrika*, 23 (1931), 114–133.
- 185a. PEARSON, E. S. en C. CHANDRA SEKAR – The efficiency of statistical tools and a criterion for the rejection of outlying observations, *Biometrika*, 23 (1936), 308–320.
186. PERRY, W. L. M. – The design of toxicity tests, Report no. 270, Medical Research Council, H. M. Stationery Office, London, 1950.
187. PROOSDIJ-HARTZEMA, Mevr. Dr E. G. – Indirecte bloeddrukmeting bij de rat, *Dissertatie*, Amsterdam, 1953.
188. PROOSDIJ-HARTZEMA, Mevr. Dr E. G. – De graphische bepaling van de ED_{50} volgens LITCHFIELD-WILCOXON, *Verslag van het symposium over ED_{50} bepalingen*, Utrecht, 1958.

LITERATUUROPGAVE

189. QUENOUILLE, M. H. – Associated Measurements, Butterworths Scientific Publications, London, 1952.
190. RÜMKE, Dr Chr. L. – De fouten bij de telling van bloedcellen in een telkamer, N.T.v.G., 98 (1954), 3480–3485.
191. RÜMKE, Dr Chr. L. – De toepassing van aseleceringsprocedures, Meded. Medisch Biologische Sectie V. v. S., 1958, nr. 6.
192. RÜMKE, Dr Chr. L. – Toepassing van statistiek bij experimenten in een farmacologisch laboratorium, Statistica Neerl., 12 (1958), 149–158.
193. RÜMKE, Dr Chr. L. – De fouten in de uitkomsten van de differentiatie van cellen in een uitstrijkpreparaat, N.T.v.G., 102 (1958), 2505–2508.
194. SATTERTHWAITTE, F. E. – An approximate distribution of estimates of variance components, Biometrics Bull., 2 (1946), 110–114.
195. SMITH, H. F. – The problem of comparing the results of experiments with unequal errors, J. Sc. Ind. Research (India), 9 (1936), 211–212.
196. SPEARMAN, C. – The method of 'right and wrong cases' ('constant stimuli') without GAUSS's formulae, Br. J. of Psych., 2 (1908), 227–242.
197. SPITS, J. – Syllabus 'Inleiding variantie-analyse', 1956–57, Psych. Lab. Soc.-Paed. Instituut, Universiteit van Amsterdam.
198. Statistical Research Group, Columbia University – Sequential analysis of statistical data, Applications, rev. ed., Columbia University Press, New York, 1945.
199. STRIK, R. VAN, – Bepaling van de mediaan-effectieve dosis (ED_{50}) en de relatieve effectiviteit met de probit-analyse, Verslag v. h. symposium over ED_{50} bepalingen, Utrecht, 1958.
200. STUDENT (W. S. GOSSET) – The probable error of a mean, Biometrika 6 (1908), 1.
201. STUDENT (W. S. GOSSET) – On the distribution of the means of samples which are not drawn at random, Biometrika, 7 (1909), 210–214.
202. THOMPSON, C. M. – On a criterion for the rejection of observations and the distribution of the ratio of the deviation to the sample standard deviation, Annals Math. Stat., 6 (1935), 214–219.
203. THOMPSON, C. M. – Tables of the percentage points of the χ^2 -distribution, Biometrika, 42 (1941), 188–189.
204. THOMPSON, C. M. – 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 (1947), 115–145.
205. TUKEY, J. W. – Comparing individual means in the analysis of variance, Biometrics, 5 (1949), 99.
206. WELCH, B. L. – The generalization of 'STUDENT's' problem when several different population variances are involved, Biometrika, 34 (1947), 28–35.
207. WEST, Q. M. – The results of applying a simple random sampling process to farm management data, Cornell University, 1951.
208. WILKS, S. S. – Mathematical statistics, Princeton University Press, Princeton, 1943.
209. WILKS, S. S. – Undergraduate statistical education, J. Am. Stat. Ass., 46 (1951), 1–18.
210. YATES, F. – The influence of 'Statistical methods for research workers' on the development of the science of statistics, J. Am. Stat. Ass., 46 (1951), 19–34.
211. YOUNDEN, W. J. – The Fisherian revolution in methods of experimentation, J. A. Stat. Ass., 46 (1951), 47–50.

BIJLAGE III

212. ANSCOMBE, F. J. – The transformation of Poisson, binomial and negative binomial data, *Biometrika*, 35 (1948), 246–254.
213. BARTLETT, M. S. – Square root transformation in analysis of variance, *Suppl. J. Royal. Stat. Soc.*, 3 (1936), 168–178.
214. BARTLETT, M. S., F. J. ANSCOMBE, W. G. COCHRAN, J. BERKSON – Discussion of 'The analysis of variance with various binomial transformations by Sir RONALD FISHER', *Biometrics*, 10 (1954), 140–151. Dit artikel bevat een afzonderlijk commentaar van elk der genoemde personen.
215. BERKSON, J. – Are there two regressions, *J. Am. Stat. Ass.*, 45 (1950), 250.
216. BLISS, C. I. – Confidence limits for measuring the precision of bioassays, *Biometrics*, 12 (1956), 491–526.
217. CLARINGBOLD, P. J., J. D. BIGGERS en C. W. EMMENS – The angular transformation in quantal analysis, *Biometrics*, 9 (1953), 467–484.
218. EISENHART, C. – The interpretation of certain regression methods and their use in biological and industrial research, *Ann. Math. Stat.*, 10 (1939), 162–186.
219. EPSTEIN B. en C. W. CHURCHMAN – On the statistics of sensitivity data, *Ann. Math. Stat.*, 15 (1944), 90–96.
220. FINNEY, D. J. – The choice of a response metameter in bio-assay, *Biometrics*, 5 (1949), 261–272.
221. FISHER, R. A. – The analysis of variance with various binomial transformations, *Biometrics*, 10 (1954), 130–139.
222. GADDUM, J. H. – Lognormal distributions, *Nature*, 156 (1945), 463–466.
223. JERNE, N. K. en E. C. WOOD – The validity and meaning of the results of biological assays, *Biometrics*, 5 (1949), 273–299.
224. JESSOP, W. N. – One line or two, *Applied Stat.*, 1 (1952), 131–137.
225. LINDLEY, D. V. – Regression lines and the linear functional relationship, *J. R. Stat. Soc. Supp.*, 9 (1947), 218.
226. MOSTELLER, F. en R. R. BUSH, *Selected quantitative techniques*, ch. 8, *Handbook of social psychology*, vol. 1, Harvard University Press, 1954.
227. PANNE, C. VAN DE en H. LANGE – Scheefheidstoetsen, *Statistica Neerl.*, 12 (1958), 255–264.
228. PAULSON, E. – An approximate normalization of the analysis of variance distribution, *Ann. Math. Stat.*, 13 (1942), 233–235.
229. WALLIS, W. A. – Statistics of the Kinsey report, *J. Am. Stat. Ass.*, 44 (1949), 463–484.
230. WILSON, E. B. en M. M. HILFERTY – The distribution of chi-square, *Nat. Acad. Sc. Proc.*, 24 (1932), 428–440.
231. KIMBALL, A. W. – Short-cut formulas for the exact partition of χ^2 in contingency tables, *Biometrics*, 10 (1954), 452–458.
232. YATES, F. – The analysis of contingency tables with groupings based on quantitative characters, *Biometrika*, 35 (1948), 176–181.

BIJLAGE IV

ANTWOORDEN BIJ DE OPGAVEN

De antwoorden bij de opgaven 1.5 t/m 1.11 zijn achterwege gelaten. Zij liggen o.i. voor de hand en een werkelijk volledige bespreking zou te veel plaatsruimte vergen. Om dezelfde reden hebben wij ook de (enkele) antwoorden weggelaten, waarbij een grafiek moest worden getekend.

Hoofdstuk 1

1.2. 2, 4, 8, 9, 10 en 1, 3, 5, 6, 7.

1.3. 11, 13, 15, 16, 19, 20, 22, 26, 27, 30;
 3, 4, 6, 7, 9, 10, 18, 24, 25, 29;
 1, 2, 5, 8, 12, 14, 17, 21, 23, 28.

1.4. Voorschrift *b* is onjuist.

Hoofdstuk 2

(Bij de opgaven 2.1, 2.2 en 2.3 geven wij slechts één der mogelijke indelingen)

2.1.

Klassen van — t/m	<i>f</i>	<i>F</i>
13,9—14,1	1	50
13,6—13,8	4	49
13,3—13,5	5	45
13,0—13,2	11	40
12,7—12,9	12	29
12,4—12,6	10	17
12,1—12,3	4	7
11,8—12,0	2	3
11,5—11,7	1	1
Totaal	50	—

2.2.

Klassen van — t/m	<i>f</i>	<i>f/n</i>
34—36	1	0,0147
31—33	1	0,0147
28—30	1	0,0147
25—27	2	0,0294
22—24	3	0,0441
19—21	2	0,0294
16—18	4	0,0588
13—15	8	0,1176
10—12	9	0,1324
7—9	15	0,2206
4—6	15	0,2206
1—3	7	0,1029
Totaal	68	0,9999

2.3.

		Hersengewicht in gr							<i>f_s</i>	
		1150	1250	1350	1450	1550	1650	1750		1850
Schedellengte in mm	200								1	1
	195					1	1		1	3
	190				1	1	1	1		4
	185			2	1	2	1			6
	180			2	4	4	1	1		12
	175		2	3	7	2	1			15
	170		1	3	3	2	1			10
	165	1	1	2	2					6
	160			1		1				2
	155		1							1
<i>f_s</i>		1	5	13	18	13	6	2	2	60

BIJLAGE IV

- 2.4. t/m 2.7: tekeningen.
- 2.8. a. 1 b. 2 c. 2,6 d. 2,050 e. 1,43.
- 2.9. a. 4 b. 4 c. 4,0 d. 2,200 e. 1,48.
- 2.10. $\mu_a = 15,0$, $\sigma_a = 2,105$.
- 2.11. a. Het gemiddelde en de mediaan worden 10 hoger, de standaardafwijking verandert niet.
- b. Het gemiddelde wordt 2 maal zo groot, dus als de oorspronkelijke waarde \bar{x} was, wordt de nieuwe waarde $2(\bar{x} + 10)$. Hetzelfde geldt voor de mediaan. Als de oorspronkelijke standaardafwijking s is, wordt de nieuwe $2s$.
- c. Het gemiddelde neemt met 1 toe en wordt dus $2\bar{x} + 21$. De mediaan behoudt de onder (b) verkregen waarde. De standaardafwijking kan slechts worden berekend, als men $(x_9 - \bar{x})$ kent.
- 2.12. Mediaan = 387, $\bar{x} = 387,6$, $s = 6,14$.
- 2.13. Mediaan = 9, $\bar{x} = 11,0$.
- 2.14. a. $\bar{x} = 12,84$, $s = 0,5127$; b. $\bar{x} = 12,85$, $s = 0,5096$ (voor de door ons samengestelde frequentieverdeling, zie 2.1).
- 2.15. $\bar{x} = 107,5$, $s = 9,27$.

Hoofdstuk 3

- 3.1. 5 blauw, 1 rood 3.2. a. 1/6 b. 5/18 3.3. 46/135
- 3.4. Enige mogelijkheid: in de zwarte doos 2 zwarte ballen en 1 witte bal, dus in de witte doos 3 zwarte en 3 witte. De gevraagde kans is:

$$C_2^5 \cdot C_1^4 / C_3^9 = 10/21.$$
- 3.5. a. 13/204 b. 13/102 c. 13/102 d. 25/204.
- 3.6. De kans, dat zich onder de eerste 9 lampen 1 defecte lamp bevindt is $C_1^2 \cdot C_8^{10} / C_{10}^{12} = 9/22$. De kans, dat van de resterende 3 lampen het eerste getrokken exemplaar defect is bedraagt 1/3. De gevraagde kans is dus $9/22 \times 1/3 = 3/22$.
- 3.7. a. $\mu_x = 4,2$, $\mu_y = 0,5$, $\sigma_x^2 = 1,56$, $\sigma_y^2 = 0,45$.
- b.

z	2	3	4	5	6	7	8
$P(z=z)$	0,06	0,15	0,25	0,23	0,21	0,08	0,02

 $\mu_z = 4,2 + 0,5 = 4,7$
 $\sigma_z^2 = 1,56 + 0,45 = 2,01$
- c.

w	0	1	2	3	4	5	6
$P(w=w)$	0,01	0,05	0,15	0,23	0,26	0,18	0,12

 $\mu_w = 4,2 - 0,5 = 3,7$
 $\sigma_w^2 = 1,56 + 0,45 = 2,01$
- 3.8.

x	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
216 [P(x=x)]	1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1

 $\mu_x = 3(3,5) = 10,5$, $\sigma_x^2 = 3(17,5/6) = 52,5/6 = 8,75$.
- 3.9. a. 2/7 b. 1/7 c. $1 - 5/42 = 37/42$ d. 9/14 e. 23/42 f. 31/84.
- 3.10. a. 109/350 b. $1 - 109/350 = 241/350$ c. 41/70 d. $1 - 41/70 = 29/70$
e. 9/70 f. 19/35 g. 19/35.
- 3.11. a. 118/375 b. $1 - 118/375 = 257/375$ c. 67/375 d. 6/125 e. 53/75.
- 3.12. a. $\mu_x = 4$, $\mu_y = 4$ b. $\sigma_x^2 = 5,667$, $\sigma_y^2 = 2,833$, $\sigma_z^2 = 2\sigma_x^2$

ANTWOORDEN BIJ DE OPGAVEN

c.

s_x^2	0	0,5	2	4,5	8	12,5	18	24,5
P	6/36	6/36	8/36	4/36	4/36	4/36	2/36	2/36

gemiddelde = 5,667 = σ_x^2

3.13. a.

\bar{x}	0	0,5	1	1,5	2
$P(\bar{x} = \bar{x})$	4/25	8/25	8/25	4/25	1/25

b.

	0	0,5	1	1,5
	1/10	4/10	3/10	2/10

3.14. a.

z	2	1	0
$P(z = z)$	6/55	28/55	21/55

b. $P(w = w) = P(z = z)$.

3.15.

\bar{x}	1	1½	1¾	2	2½	2¾	3	3½	3¾	4	4½	4¾	5	5½	5¾	6	6½	6¾	7	7½	8	
216 · P	1	3	6	10	12	15	16	18	21	19	21	15	15	12	9	10	3	6	3	1		

3.17.

S	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
56 · P	1	1	2	3	4	5	6	6	6	6	5	4	3	2	1	1

3.18. $P_L(3) = 0,20$, $P_R(8) = 0,06$, $P_D(1) = 0,06$, $P_D(2) = 0,16$.

3.19. $P_D(10) = 0,10$, $P_D(8) = 0,70$.

Hoofdstuk 4

4.1. $\mu_x = 0,9$, $\sigma_x^2 = 0,49$.

4.2.

x	0	1	2	3
$P(x=x)$	4/35	18/35	12/35	1/35

$\mu_x = 1^2/7$

$\sigma_x^2 = 2^4/49$

4.3. $\frac{C_5^{11} \cdot C_0^4 + C_4^{11} \cdot C_1^4 + C_3^{11} \cdot C_2^4}{C_5^{15}} = \frac{12}{13}$.

4.4. a. $\frac{3800}{23751}$ b. $\frac{19209}{23751}$ c. $1 - \frac{19209}{23751} = \frac{4542}{23751}$.

4.5. a.

v	0	1	2	3	4	5
525 · P	4	50	180	205	80	6

b.

w	0	1	2	3	4	5
1890 · P	14	235	780	695	160	6

c.

z	0	1	2	3	4	5
429 · P	2	35	140	175	70	7

4.6. a. $P(x=x) = C_x^{50} (1/6)^x (5/6)^{50-x}$, $\mu_x = 8\frac{1}{3}$ en $\sigma_x^2 = 6\frac{17}{18}$.

b. $P(x=x) = \frac{C_x^{13} \cdot C_{8-x}^{39}}{C_5^{52}}$, $\mu_x = 2$ en $\sigma_x^2 = 1^5/17$.

c. $P(x=x) = C_x^3 (0,515)^x (0,485)^{3-x}$, $\mu_x = 1,545$ en $\sigma_x^2 = 0,749325$.

d. Daar deze steekproef wel zonder teruglegging zal worden getrokken, geldt eigenlijk de hypergeometrische verdeling, maar daar de steekproefomvang ongetwijfeld klein is t.o.v. de (niet gegeven) populatie-omvang, kan deze zeer goed worden benaderd door de binomiale verdeling $P(x=x) = C_x^{100} (0,08)^x (0,92)^{100-x}$, $\mu_x = 8$, $\sigma_x^2 = 7,36$.

e. Volgt noch een hypergeometrische, noch een binomiale verdeling.

BIJLAGE IV

f. Kan niet worden gegeven, daar de fractie personen met een bloeddruk hoger dan 150 mm Hg in de populatie onbekend is.

g. $P(x = x) = C_x^{10} (0,005)^x (0,995)^{10-x}$, $\mu_x = 0,05$, $\sigma_x^2 = 0,04975$ (best-passende binomiale verdeling).

4.7. a. $P(x = x) = C_x^{10} (\frac{1}{3})^x (\frac{2}{3})^{10-x}$, daar $n = 10$ en $N = 300$ zal de benadering bevredigend zijn.

b. Geen binomiale benadering toepassen, daar $n = 11$ te groot is t.o.v. $N = 35$.

4.8.	x	0	1	2	3	4	5	6
	$P(x = x)$	1/64	6/64	15/64	20/64	15/64	6/64	1/64

4.9. $1 - \left(\frac{35}{36}\right)^{34} = 0,49128$.

4.10. $\frac{1}{243}$ 4.11. a. $(0,99)^{10}$ b. $10(0,99)^9(0,01)$ c. $(a + b)$ d. $(1 - c)$

e. $(1 - a)$ f. $[a + b + 45(0,99)^9(0,01)^2]$.

4.12. $P(x = 0) = (0,8)^{64}$.

4.13.	x	$P(x=x)$
	0	0,48225
	1	0,38580
	2	0,11574
	3	0,01543
	4	0,00077

4.14.	x	e
	3	40
	2	120
	1	120
	0	40

4.15.	x	e
	0	78
	1	314
	2	420
	3	188

Zonder toetsing is reeds te zeggen, dat de steekproef niet uit een binomiaal verdeelde populatie zal stammen.

4.16. a. 0,41637 b. 0,00001
c. 0,00129 d. 0,05265
e. 0,00209.

4.17. a. 0,00129 b. 0,00129 c. 0,00258 d. 0,00258 e. 0,01182
f. 0,05766 g. 0,11532 h. 0,02070.

4.18. a. $x = 6$ b. $x = 15$ c. $x = 5$ of 15 d. $x = 3$ of 17.

4.19. a. $P(0) \cdot \mu^4/4! = P(0) \cdot \mu^3/3!$, zodat $\mu = 4$, $P(0) = 0,01832$ (tabel 4.6),
 $P(1) = 4(0,01832) = 0,07328$, $P(x \leq 1) = 0,09160$.

b. $\mu = 2$, $P(0) = 0,1353$.

4.20. De grootheid $x = 5$ kruis volgt bij benadering een Poissonverdeling met $\mu = Pn = (1/32) 64 = 2$. Deze verdeling luidt:

x	0	1	2	3	4	5	6	7	8	9	>10	Totaal
P	0,1353	0,2706	0,2706	0,1804	0,0902	0,0361	0,0120	0,0034	0,0009	0,0002	0,0005	1,0002

4.21. $P(0) = 0,6065$ (tabel 4.6), $P(1) = (0,5)(0,6065) = 0,30325$,
 $P(x > 2) = 1 - (0,6065 + 0,30325) = 0,09025$.

4.22. $\mu = 10000(0,00004) = 0,4$, $P(x < 3) = 0,9992$, $P(x > 4) = 0,0008$.

4.23.	x	0	1	2	3	4	5	6	7	8	9	10	11	12	Totaal
	e	12	44	80	98	90	66	40	21	10	4	1	1	0	467

ANTWOORDEN BIJ DE OPGAVEN

4.24. *a* en *b*. $\bar{x}_1 = 0,5$, $\bar{x}_2 = 1,0$, $\bar{x}_3 = 1,5$.

4.25. Zie tabel 9.1 op blz. 174.

Hoofdstuk 5

5.1. *a*. 0,0202 *b*. 0,6700 *c*. 0,6600 *d*. 0,6600 *e*. 0,0027 *f*. 0,4086
g. 0,6154 *h*. 0,1717.

5.2. *a*. 0,5793 *b*. 0,0228 *c*. 0,0456 *d*. 0,0000634 *e*. 0,3108 *f*. 0,0895
g. 0,5000 *h*. 0,3085.

5.3. *a*. 2,054 *b*. 2,878 *c*. -0,583 *d*. -3,09 *e*. -1,960 *f*. 1,645.

5.4. *a*. 0,12 *b*. 0,13 *c*. 0,99 *d*. 0,60 *e*. 0,10 *f*. 0,99 *g*. 0,75 *h*. 0,10.

5.5. *a*. 0,994 *b*. -1,751 *c*. 0,358 *d*. 0,674 of -0,674.

5.6. *a*. 0,9332 *b*. 0,7888 *c*. 0,0228.

5.7. $1 - (0,9772)^3 = 0,06685$.

5.8. *a*. $\mu_x = 100$ of 120 *b*. $\sigma_x = 25$, $P_L(x = 60) = 0,0548$.

5.9. $\sigma_x = 4$, $\mu_x = 74,2$.

5.10.

<i>x</i>	$P(x = x)$ Benaderd Exact	
20	0,0000	0,0000
19	0,0000	0,0000
18	0,0000	0,0000
17	0,0000	0,0000
16	0,0002	0,0003
15	0,0012	0,0013
14	0,0045	0,0049
13	0,0142	0,0146
12	0,0346	0,0355
11	0,0723	0,0710
10	0,1212	0,1171
9	0,1607	0,1597
8	0,1820	0,1797
7	0,1607	0,1659
6	0,1212	0,1244
5	0,0723	0,0746
4	0,0346	0,0350
3	0,0142	0,0124
2	0,0045	0,0031
1	0,0012	0,0005
0	0,0003	0,0000
	0,9999	1,0000

5.11. *a*. 0,00415 *b*. 0,00830

5.12. *a*. 0,0192 *b*. 0,00317
c. 0,7698

5.13.

<i>x</i>	<i>e</i>
51	3
50	4
49	9
48	17
47	26
46	38
45	47
44	49
43	45
42	36
41	25
40	15
39	8
38	3
37	1
36	1
	327

5.14.

<i>x</i>	<i>e</i>
5-8	0
9-12	2
13-16	20
17-20	109
21-24	420
25-28	1124
29-32	1976
33-36	2386
37-40	1963
41-44	1112
45-48	412
49-52	106
53-56	19
57-60	2
61-64	0
	9651

Hoofdstuk 6

6.1. *a*. $Z: x = 16$ t/m 20 , $\alpha_Z = 0,05096$, $\beta_Z = 0,04318$, $1 - \beta_Z = 0,95682$.

b. H_0 wordt verworpen ten gunste van H_1 .

6.2. *a*. $Z: x = 0$ t/m 9 , $\alpha_Z = 0,01716$, $\beta_Z = 0,04798$, $1 - \beta_Z = 0,95202$.

b. H_0 wordt verworpen ten gunste van H_1 .

BIJLAGE IV

- 6.3. a. $n = 66$, $x_n = 32$ b. $n = 96$, $x_n = 46$ c. $n = 131$, $x_n = 65$.
- 6.4. a. Verwerpen, $0,02 < P_D < 0,05$ i. Verwerpen, $P_R = 0,0367$
 b. Niet verwerpen, $0,02 < P_D < 0,05$ j. Niet verwerpen, $P_D = 0,0734$
 c. Verwerpen, $0,025 < P_L < 0,05$ k. Niet verwerpen, $P_R = 0,0505$
 d. Niet verwerpen, $P_R > 0,995$ l. Niet verwerpen, $P_L = 0,2033$
 e. Verwerpen, $0,01 < P_R < 0,025$ m. Niet verwerpen, $P_D = 0,0512$
 f. Niet verwerpen, $P_D = 0,0818$ n. Verwerpen, $P_R = 0,0256$
 g. Niet verwerpen, $P_D = 0,0142$ o. Niet verwerpen, $P_D = 0,0178$
 h. Niet verwerpen, $P_R = 0,9906$
 (De kansen bij f t/m o gelden uiteraard bij benadering)
- 6.5. $T_0 = -2,83$, $P_L = 0,00233$, H_0 wordt verworpen t.g.v. H_1 .
- 6.6. $T_0 = 2,91$, $P_R = 0,00181$, H_0 wordt verworpen t.g.v. H_1 .
- 6.7. De vergelijkingen van de lijnen zijn:
 a. $L_0: a_n = 0,7327n - 1,340$, $L_1: b_n = 0,7327n + 1,340$.
 b. $L_0: a_n = 0,7327n - 2,091$, $L_1: b_n = 0,7327n + 2,091$.
 c. $L_0: a_n = 0,6609n - 2,124$, $L_1: b_n = 0,6609n + 2,124$.
 d. $L_0: a_n = 0,6609n - 3,314$, $L_1: b_n = 0,6609n + 3,314$.
- 6.8. a.

P	GSG
0	2
0,5	6
0,73	10
0,9	8
1	6

 b.

P	GSG
0	3
0,5	9
0,73	23
0,9	13
1	8

 c.

P	GSG
0	4
0,5	12
0,66	21
0,8	14
1	7

 d.

P	GSG
0	6
0,5	21
0,66	50
0,8	24
1	10
- 6.9. a. $T = 2,13$, $P_R = 0,0166$, H_0 wordt verworpen.
 b. Vergelijkingen: $a_n = 0,186n - 1,67$, $b_n = 0,186n + 2,14$, H_0 wordt bij de 10-de waarneming verworpen.

Hoofdstuk 7

- 7.1. a. Door de steekproefomvang te vergroten. b. De betrouwbaarheidscoëfficiënt wordt kleiner.
- 7.2. a. $0,35 < P < 0,48$ b. $0,3605 < P < 0,4819$ c. $0,3588 < P < 0,4812$.
- 7.3. a. 246 b. 369 c. 384 d. 381 e. 369.
- 7.4. a. 0,3074 b. 0,2933.
- 7.5. a. $0,352 < P < 0,448$ b. $0,568 < P < 0,632$ c. $0,780 < P < 0,820$
 $0,337 < P < 0,463$ $0,558 < P < 0,642$ $0,774 < P < 0,826$
 d. $0,184 < P < 0,216$ e. $0,435 < P < 0,565$ f. $0,461 < P < 0,539$
 $0,179 < P < 0,221$ $0,414 < P < 0,586$ $0,448 < P < 0,552$
 g. $0,880 < P < 0,920$ h. $0,890 < P < 0,910$
 $0,874 < P < 0,926$ $0,887 < P < 0,913$
- 7.6. a. 246 b. 984 c. 6144 d. 24576.
- 7.7. Bij de volgende opmerkingen gebruiken wij steeds fracties i.p.v. percentages. De fractie manifest linkshandigen (ML) in de steekproef van $n_1 = 50$ personen van de auteur noemen wij $p_1 = 0,06$ (zijn populatiefractie is P_1). De fractie ML in de steekproef van HUTTNER (van de onbekende omvang n_2) noemen wij $p_2 = 0$ (en zijn populatiefractie P_2).

ANTWOORDEN BIJ DE OPGAVEN

3e alinea. De auteur gebruikt voor zijn betrouwbaarheidsinterval de formules (7.3) en (7.4) met $T = 3$, zodat hij met de betrouwbaarheidscoëfficiënt 0,9973 werkt. Men kan deze formules echter slechts voor het berekenen van een benaderend interval gebruiken, als n groot is en p niet te dicht bij 0 of 1 ligt. Aan deze voorwaarden is niet voldaan, zodat hij het *onjuiste* interval $-0,041 < P_1 < 0,161$ verkrijgt (wat niet alleen de waarde $P = 0$, maar zelfs negatieve waarden bevat!). Maakt men gebruik van de tabellen van MAINLAND (45), dan vindt men als 95% betrouwbaarheidsinterval $0,013 < P_1 < 0,166$ en als 99% interval $0,007 < P_1 < 0,203$. Deze niet-symmetrische intervallen bevatten de waarde nul niet — vanzelfsprekend, want de auteur kan op grond van zijn steekproef met zekerheid beweren, dat zijn populatie manifest links-handigen telt.

4e alinea. HUTTER bedoelt uiteraard dat het totaal ontbreken van ML in zijn *steekproef* een zeldzaam voorkomen in zijn *populatie* niet uitsluit. Hoé zeldzaam hangt van zijn steekproefomvang (n_2) af. Dit blijkt, als men voor verschillende waarden van n de 95% betrouwbaarheidsintervallen van P_2 beschouwt, die gelden als $p_2 = 0$; deze zijn eveneens ontleend aan MAINLAND:

$n = 20, 0 \leq P < 0,167; n = 50, 0 \leq P < 0,071; n = 100, 0 \leq P < 0,036;$
 $n = 500, 0 \leq P < 0,0074; n = 1000, 0 \leq P < 0,0037.$

2e alinea. Als men onderstelt, dat de steekproeven van de auteur en van HUTTER aselekt getrokken zijn, dan kan de hypothese $H_0: P_1 = P_2$ met de toets van FISHER (zie 9.5) worden getoetst. Men moet dan echter wel degelijk weten, hoe groot de steekproef van HUTTER is. Is deze steekproef betrekkelijk klein, dan zal het verschil tussen de bevindingen van de auteur en HUTTER niet significant zijn. Bestaat deze steekproef echter b.v. uit 300 personen, dan is het verschil reeds significant bij een 5% drempel.

De toets voor het vergelijken van twee percentages (fracties), die de auteur geeft, is bij een juist gebruik equivalent aan de χ^2 -benadering van de toets van FISHER. Men moet dan nemen [zie formule (24.11) op blz. 624]:

$$T = \frac{p_1 - p_2 \pm 2/n}{\sqrt{\frac{p_0 q_0}{n_1} + \frac{p_0 q_0}{n_2}}}, \text{ waarin } p_0 = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}, q_0 = 1 - p_0$$

en waarin $\pm 2/n$ een continuïteitscorrectie is. De toets kan echter alleen worden gebruikt, als de steekproeven niet te klein zijn en als p_1 en p_2 niet te dicht bij 0 of 1 liggen (dit komt overeen met de voorwaarde bij de χ^2 -benadering van de toets van FISHER, dat de verwachte frequenties niet te laag mogen zijn).

Hoofdstuk 9

- 9.2. a. $0,005 < P < 0,01$ b. $0,025 < P < 0,05$ c. $0,10 < P < 0,20$
 d. $0,50 < P < 0,70$ e. $0,95 < P < 0,99$ f. $0,20 < P < 0,30$
 g. $0,98 < P < 0,99$ h. $0,01 < P < 0,02$ i. $0,01 < P < 0,05$
 j. $P < 0,001.$
- 9.3. a. $P_R = 0,0078$ b. $P_R = 0,3131.$
- 9.5. $X_0^2(3) = 7,96, 0,025 < P < 0,05, \text{ verwerpen.}$

BIJLAGE IV

- 9.6. a. $H_0: P(x=1) = P(x=2) = \dots = P(x=6) = 1/6$, H_1 : de dobbelsteen is onzuiver, $X_0^2(5) = 11,48$, $0,025 < P < 0,05$, H_0 wordt verworpen. Het lijkt, alsof er een preferentie is voor de uitkomsten 4, 5 en 6.
- b. $H_0: P(x < 3) = P(x > 4)$, $H_1: P(x < 3) < P(x > 4)$, dus links éézijdige toetsing. $X_0^2(1) = 10,83$, $P \cong 0,0005$, H_0 wordt verworpen.
- 9.7. a.

1.	$X_0^2(3) = 33,79$ sign.
2.	93,66 sign.
3.	16,30 sign.
4.	12,71 sign.
5.	3,39 n.sign.
6.	2,13 n.sign.
7.	5,24 n.sign.
8.	1,91 n.sign.
9.	8,91 sign.
10.	1,51 n.sign.
11.	1,19 n.sign.
12.	1,95 n.sign.
13.	0,61 n.sign.
14.	3,54 n.sign.
15.	11,60 sign.

b. 1.	$X_0^2(1) = 0,20$ n.sign.
2.	0,20 id.
3.	0,42 id.
4.	0,12 id.
5.	0,56 id.
6.	0,12 id.
7.	0,12 id.
8.	0,12 id.
9.	0,30 id.
10.	0,12 id.
11.	0,12 id.
12.	0,20 id.
13.	0,12 id.
14.	0,12 id.
15.	0,42 id.

Zeer in 't kort valt nu over de uitkomsten het volgende op te merken:

1. Bij vergelijking van de volledige uitkomsten (in vier categorieën) blijken 6 van de 15 onderzoekers uitkomsten te produceren, die significant van het genetisch model afwijken.

2. Beschouwt men alleen de indeling in 'starchy' en 'sweet', dan zij echter de uitkomsten per onderzoeker zonder enige uitzondering ver van significant. Bij 14 van de 15 onderzoekers is echter de 'starchy'-telling hoger (dus de 'sweet'-telling lager) dan de verwachting. Volgens de tekentoets is deze uitkomst (14 plus, 1 min) significant bij een 1% drempel, zodat tot een weliswaar klein, maar systematisch verschil met de verwachting (ten gunste van 'starchy') kan worden geconcludeerd.

3. Bij de als 'starchy' geklassificeerde korrels treft men significant meer gele korrels aan, dan er verwacht worden (tekentoets, yellow starchy 14 maal hoger dan de verwachting). Bij de als 'sweet' geklassificeerde korrels is dit niet het geval (yellow sweet is 7 maal hoger, 8 maal lager dan de verwachting). Wel valt hier op, dat onderzoeker 2 blijkbaar geen juist onderscheid tussen gele en witte korrels heeft weten te maken. T.a.v. de kleurverdeling binnen de 'starchy' soort zijn er twee mogelijkheden; óf er komen meer gele korrels voor, dan met het genetisch model verenigbaar is, óf het kleuronscheid is hier moeilijker (dan bij de 'sweet' soort) en de onderzoekers maken een systematische fout ten gunste van de kleur geel.

9.8. Volgens de gegeven 'wet' vindt men de volgende verwachtingen:

R	0	0,04	0,10	0,20	0,50	1,00
$\log e$	2,4700	2,3908	2,2720	2,0740	1,4800	0,4900
e	295,1	245,9	187,1	118,6	30,2	3,1

ANTWOORDEN BIJ DE OPGAVEN

De overeenstemming tussen f en e blijkt zo fraai te zijn, dat toetsing wel achterwege kan blijven. Wil men echter toetsen, dan dienen f en e in frequenties per regenhoeveelheid te worden omgezet:

Regenval in inches	0	0,01-0,04	0,05-0,10	0,11-0,20	0,21-0,50	0,51-0,99	> 1
f	69	50	59	68	89	27	3
e'	69,9	49,2	58,8	68,5	88,4	27,1	3,1

9.9. Hij zal soort B kopen (tekentoets, $0,025 < P_L < 0,05$).

9.10. Zie 9.7. 9.11. a. 62 en 88 b. 109 en 141 c. 227 en 273.

9.12. $Q_0 = \frac{4 [5 \times 242 - (32)^2]}{5 \times 32 - 64} = 7,75 \cong \chi^2(4)$, P ongeveer gelijk aan 0,10.

De getoetste hypothese wordt niet verworpen.

9.13. De kansverdeling van de grootheid g luidt:

a	0	1	2	3	4	5	6	7
$11440 \cdot P$	1	63	756	2940	4410	2646	588	36

$P_D(g=2) = \frac{1 + 63 + 756 + 588 + 36}{11440} = 0,1262$, H_0 wordt niet verworpen.

9.14. De kansverdeling van g is:

a	0	1	2	3	4	5	6	7	8
$P(g=a)$	0,00393	0,05030	0,20538	0,35208	0,27507	0,09780	0,01467	0,00076	0,00001

$P_D(g=1) = 0,00393 + 0,05030 + 0,01467 + 0,00076 + 0,00001 = 0,06967$, H_0 wordt niet verworpen.

9.15. $X_0^2(1) = 1,40$, $0,2 < P < 0,3$. Het verschil is dus niet significant.

9.16. De hypothese, dat de percentages tumoren in de proef- en controle-groep toevallig verschillen, dient voor de groepen A en B afzonderlijk te worden getoetst. Uitkomsten:

Wijfjes-dieren: $X^2(1) = 7,13$, $0,005 < P_D < 0,01$, significant.

Mannetjes-dieren: $X_0^2(1) = 3,53$, $0,05 < P_D < 0,10$, (net) niet significant.

Het verschil tussen deze uitkomsten is vnl. een gevolg van de kleinere omvang van de steekproeven bij de mannetjes. De uitkomsten van beide toetsingen kunnen volgens de in 12.3 beschreven methode worden gecombineerd.

9.17. a. Hoe zijn beide groepen samengesteld? Alleen indien de toewijzing van een behandeling aselekt geschiedde, kan men bij een significant verschil in mortaliteit tot een verschillend effect van de behandelingen besluiten.

b. $X_0^2(1) = 1,78$, $0,1 < P_D < 0,2$, niet significant.

BIJLAGE IV

9.18. De tabel is uit de volgende 2×2 tabellen samengesteld (+ = overleden, — = overleefd, C = cirrose, E = extra-hepatisch block):

		<i>Porta-cava shunt</i>			<i>Spleno-renale shunt</i>		
		C	E		C	E	
Blakemore	+	13	0	$\alpha = 12,2$	11	1	$\alpha = 8,0$
	—	52	4		43	26	
		65	4	69	54	27	81
Linton	+	3	0	$\alpha = 2,2$	1	6	$\alpha = 2,9$
	—	8	4		18	21	
		11	4	15	19	27	46
Child	+	4	—		0	0	
	—	27	—		27	1	
		31	—	31	1	3	4

De uitkomsten van Child blijken niets te zeggen, daar bij de PCS geen gegevens betreffende extra-hepatisch block beschikbaar zijn en bij de SRS slechts zeer kleine aantallen zonder sterfte gegeven zijn. Beschouwen wij de gegevens van Blakemore en Linton, dan blijkt het volgende. Bij de PCS ligt bij beide onderzoekers de sterfte bij cirrose iets hoger dan de verwachting, maar de verschillen zijn zo miniem, dat zij niets zeggen en dat toetsing wel achterwege kan blijven. Bij de SRS ziet men bij Blakemore een sterfte bij de cirrose-groep, die iets hoger is dan de verwachting (niet significant, $X_0^2(1) = 2,75$), maar bij Linton is zelfs precies het tegendeel het geval (hoewel ook hier van significantie geen sprake is). De conclusie, dat onafhankelijk van de toegepaste methode (d.i. dus: bij elke methode afzonderlijk) de sterfte bij levercirrose aanzienlijk (!) hoger is, blijkt dus volledig ongewettigd te zijn.

Opmerking: Bij de PCS blijken de uitkomsten van de drie onderzoekers bij cirrose niet significant te verschillen:

<i>f</i>	<i>B</i>	<i>L</i>	<i>C</i>	<i>e</i>	<i>B</i>	<i>L</i>	<i>C</i>	<i>e</i>
+	13	3	4	20	12,1	2,1	5,8	20,0
—	52	8	27	87	52,9	8,9	25,2	87,0
				65	11	31	107	$X_0^2(2) = 1,25$ $0,5 < P < 0,7$
				65,0	11,0	31,0	107,0	

Vergelijkt men nu de totaalcijfers voor levercirrose bij de PCS met die voor extra-hepatisch block, dan vindt men evenmin een significante uitkomst:

		PCS (totaal)		
		C	E	
+	20	0	20	$\alpha = 18,6$
—	87	8	95	
		107	8	115

9.19. $X_0^2(2) = 9,97$, $0,005 < P_D < 0,01$, verwerpen.

ANTWOORDEN BIJ DE OPGAVEN

- 9.20. Past men de χ^2 -toets toe, dan dient men tot combinatie van klassen over te gaan, teneinde te lage verwachtingen te voorkomen. Deze combinatie is echter arbitrair, zodat verschillende uitkomsten kunnen ontstaan, wat een belangrijk nadeel van deze toetsingsmethode is. Combineert men bv. 5, 6, 7 en 8 en hoger, dan is de uitkomst $X_0^2(5) = 5,96$, $0,3 < P < 0,5$. Men kan deze moeilijkheid vermijden door op de gegevens de toets van WILCOXON toe te passen. Men vindt dan (normale benadering met correctie voor gelijken): $T = 1,66$, $P_D = 0,0970$.
- 9.21. $X_0^2(4) = 4,04$, $0,3 < P < 0,5$, niet verwerpen.
- 9.22. $X_0^2(9) = 68,0$, $P \ll 0,001$, verwerpen. Bestudering van de verschillen tussen de waargenomen en verwachte frequenties en van hun bijdragen tot X^2 laat zien, dat deze zeer significante uitkomst vnl. wordt veroorzaakt door het te veel voorkomen van de meningococce vorm bij kinderen beneden 1 jaar (te weinig bij kinderen van 1 en 2 jaar) en het te weinig voorkomen van de tuberculeuze vorm bij deze kinderen (te veel bij de 1 en 2 jarigen).
- 9.23. $X_0^2(1) = 9,70$, $0,005 < P_D < 0,001$, verwerpen.
- 9.24. $X_0^2(4) = 97,25$, $P \ll 0,001$, verwerpen.

Hoofdstuk 10

- 10.1.¹ a. $X_0^2(2) = 3,47$, $0,1 < P < 0,2$, niet significant.
 b. $X_0^2(2) = 152,55$, $P \ll 0,001$, zeer significant.
 c. $X_0^2(8) = 8,93$, $0,3 < P < 0,5$, niet significant.
 d. $X_0^2(8) = 979,18$, $P \ll 0,001$, zeer significant.
 Deze waarden zijn berekend uit de bij de betrokken opgaven in deze antwoorden opgenomen verwachte verdelingen.

10.2. H_0 wordt alleen voor geval (a) verworpen.

10.3. a.

F_i/n	0	0,1	0,1	0,6	1,0	$D_0 = \max. = 0,6 - 0,1 = 0,5$ $D_{0,05}(10) = 0,410$, verwerpen.
E_i/n	0,2	0,4	0,6	0,8	1,0	

b. Men kan met dit voorbeeld even goed 'bewijzen', dat de χ^2 -toets superieur is. Combineert men nl. 1 + 2 + 3 en 4 + 5, dan vindt men $X_0^2(1) = 10,42$ (met continuïteitscorrectie: 8,44), zodat men zelfs tot verwerpen kan besluiten met een onbetrouwbaarheid van 0,005 (men dient echter de uitkomst van de χ^2 -toets wegens de lage verwachtingen met enige reserve te hanteren). Uit dit voorbeeld blijkt, dat de χ^2 -toets in dit soort gevallen een uitkomst oplevert, die in sterke mate afhankelijk is van de wijze van groeperen.

- 10.4. $V_0 = 100$, $0,02 < P_D < 0,05$, verwerpen.
- 10.5. $V_0 = 54$, $n = 11$, $0,01 < P_D < 0,02$, verwerpen.
- 10.6. a. $V_0 = 71$, $n = 13$, $0,01 < P_D < 0,02$, het verschil is dus significant.
 b. De waarnemingen zijn verricht in de volgorde: 1E, 1C, 2E, 2C, enz. Deze reeks blijkt zeer regelmatig op te klimmen (deze trend is zeer significant: past men de toets van SPEARMAN op de 30 opeenvolgende

¹ a en b betreffen de opgaven 4.14, resp. 4.15 en niet, zoals op blz. 217 abusievelijk is aangegeven, 4.13 en 4.14.

BIJLAGE IV

waarnemingen toe, dan vindt men $R_0 = 28$, $r_r = 0,998$, $T = 5,35$, $P \ll 0,001$. GREENBERG geeft hiervoor de volgende verklaring: 'The larvae were relatively large in size and clung to the various surfaces of the needle and syringe causing incomplete evacuation. The syringe was then reinserted into the stock solution containing the infecting larvae, and after stirring by flushing the syringe several times, 0,05 cc was again drawn off. Due to the incomplete evacuation, the solution containing the larvae was gradually being made more virulent in terms of the relative number of larvae. . . . A flushing of the syringe in clear saline solution before reinserting it into the stock solution revealed that sufficient larvae remained in the syringe to raise each subsequent dosage by two to five larvae. Thus, each of the thirty inocula, on the average, exceeded the previous inoculum by this amount.'

- c. Het spuitje na elk gebruik grondig reinigen. De oneven-even volgorde vervangen door een aselechte volgorde, dus aseleteren. Zie verder het —ook overigens interessante — artikel van GREENBERG.
- 10.7. Natrium: $V_0 = 74$, $n = 15$, $0,02 < P < 0,05$, de hypothese dat de toename van het natriumgehalte toevallig is, wordt verworpen.
Kalium: $V_0 = 63$, $n = 14$, $P_D \cong 0,05$, verwerpen.
Opmerking: Men zal echter niet zonder meer tot een effect van de behandeling met Diamox mogen besluiten, daar geen controles aanwezig zijn. Men weet immers niet, welk effect zou zijn opgetreden als geen behandeling was toegepast.
- 10.8. $V_0 = 48$, $n = 12$, $0,05 < P_D < 0,10$, niet verworpen. Merk op, dat deze niet significante uitkomst veroorzaakt wordt door één negatieve 'uitbijter' ($N - V = -39$).
- 10.9. De waarnemingen zijn zeer grof afgerond (af rondingsinterval 5 mm kwik). In werkelijkheid zullen de twee waarden van 70 mm bij proefpersoon 4 dus wel niet precies gelijk zijn. Onderstelt men, dat de eerste waarneming lager is dan de tweede, dan vindt men $K_0 = 72$ met $P = 0,0016$. In het omgekeerde geval geldt, dat $K_0 = 74$ met $P = 0,00094$. De hypothese, dat de waargenomen verschillen toevallig zijn kan dus in ieder geval worden verworpen. Werkt men met gemiddelde rangnummers, dan vindt men $K_0 = 72,5$.
- 10.10. H_0 : de waargenomen verschillen tussen de dagen zijn toevallig, $K_0 = 6,50$, $P \gg 0,135$, niet verworpen.
- 10.11. a. $K_0 = 429,5$, $X_0^2(5) = 24,54$, $P < 0,001$, verwerpen.
b. $K_0 = 86,5$, $X_0^2(4) = 5,77$, met correctie $X_0^2 = 5,82$, $0,2 < P < 0,3$, niet verworpen.
- 10.12. $K_0 = 866$, $X_0^2(7) = 28,87$, $P < 0,001$, verwerpen; $C_0 = 0,825$.
- 10.13. $S_0 = 58,5$, $T_0 = 1,93$, $P_D = 0,054$, niet verworpen.
- 10.14. Toets toegepast op de gewichtstoename: $S_0 = 25$, $W_0 = 64$, $0,02 < P_D < 0,05$, groep B vertoont een systematisch grotere toename in gewicht dan groep A.
- 10.15. $S_0 = 403$, $T_0 = 1,51$, $P_D = 0,1310$, niet verworpen.
- 10.16. $S_0 = 92$, $T_0 = -3,00$, $P_D = 0,0027$, verwerpen. Het gehalte bij de lijders is systematisch hoger.

ANTWOORDEN BIJ DE OPGAVEN

- 10.17. De theoretische vitale capaciteit van de groepen A en B verschilt niet significant: $S_0 = 250$, $T_0 = -0,51$, $P_D = 0,61$. In dit opzicht kunnen de groepen dus als gelijkwaardig worden beschouwd. Vergelijkt men echter de verschillen tussen de actuele en de theoretische v.c. (II — I) van beide groepen, dan vindt men : $S_0 = 172$, $T_0 = 3,45$, $0,000466 < P_D < 0,000674$. De verschillen zijn dus bij groep A systematisch groter dan bij groep B .
- 10.18. $\Sigma S_j^2/n_j = 14117,75$, $X_0^2(2) = 16,19$, $P < 0,001$, verwerpen.
- 10.19. $\Sigma S_j^2/n_j = 13348,54$, $X_0^2(2) = 9,26$, $0,005 < P < 0,01$, verwerpen.
- 10.20. De twee gelijken (8,8) zouden bij nauwkeuriger meting waarschijnlijk verschillend zijn. Onderstelt men, dat de waarde bij 'eik' het laagst is, dan zijn de rangsommen resp. 25, 37 en 58, zodat $X_0^2 = 5,58$, met (volgens tabel M) een overschrijdingskans, die even groter is dan 0,051. Onderstelt men het tegendeel, dan vindt men de sommen 26, 36 en 58 en $X_0^2 = 5,37$, met een iets grotere overschrijdingskans. Werkt men met de gemiddelde rangnummers (8,5), dan vindt men de tussenliggende waarde $X_0^2 = 5,465$. Hoewel er bij de 5% drempel dus juist geen significantie wordt bereikt, leveren de uitkomsten een voldoende indicatie voor een nader onderzoek.
- 10.21. $R_0 = 78$, $n = 12$, zodat in ieder geval geldt dat $0,01 < P_D < 0,02$.
- 10.22. $R_0 = 366,5$, $n = 22$, $P_D \ll 0,01$ (tabel N). Met de normale benadering vindt men: $T_0 = 3,63$, $P_D \cong 0,00025$. H_0 wordt dus verworpen. De χ^2 -toets blijkt dus inderdaad veel minder scherp te zijn.
- 10.23. $R_0 = 1655$, $n = 24$, $P_L > 0,05$, niet verwerpen.
- 10.24. $S_0 = 34$, $n = 12$, $0,02 < P_D < 0,05$, verwerpen.
- 10.25. $S_0 = 144$, $n = 22$, $P_D \ll 0,01$, verwerpen.

Hoofdstuk 11

- 11.1. Serietoets, $u_0 = 8$, $m_1 = m_2 = 25$, $n = 50$, $P_D \ll 0,01$. Met de normale benadering vindt men $T_0 = -5,00$. De waarnemingsreeks vertoont een systematisch dalend verloop.
- 11.2. Serietoets, $u_0 = 45$, $m_1 = 35$, $m_2 = 33$, $n = 68$, $\mu_u = 34,97 = 35$, $\sigma_u = 4,09$, $T_0 = 2,32$, $P_D = 0,0204$. H_0 verwerpen, er zijn te veel series (deze zijn met opzet aangebracht).
- 11.3. Rangcorrelatietoets van SPEARMAN. Experimental: $R_0 = 2,5$, $n = 15$, $P_L \ll 0,01$; Control: $R_0 = 6$, $n = 15$, $P_L \ll 0,01$. In beide gevallen bestaat er een significante positieve correlatie, dus een systematisch stijgend verloop. Zie ook het antwoord bij 10.6.
- 11.4. Serietoets, $u_0 = 8$, $m_1 = m_2 = 10$, $n = 20$, $P_L > 0,05$, H_0 niet verwerpen. Past men echter de toets van SPEARMAN toe, dan is de uitkomst: $R_0 = 2227$, $P_R < 0,005$, dus negatieve correlatie en een significant dalend verloop.
- 11.5. a. $u_0 = 5$, $m_1 = m_2 = 8$, $n = 16$, $0,05 < P_D < 0,10$, niet verwerpen.
b. $W_0 = 109$ (feitelijk: 108 óf 110), $0,02 < P_D < 0,05$, verwerpen.
- 11.6. $u_0 = 7$, $m_1 = m_2 = \frac{1}{2}n = 13$, $P_L < 0,005$, hypothese, dat het verloop aselekt is verwerpen t.g.v. cyclische trend.

BIJLAGE IV

Hoofdstuk 12

12.1. TERPSTRA:
$$T = \frac{W}{\sigma_W} = \frac{n_1(n+1) - 2S_1}{n_1 n_2 \sqrt{\frac{1}{3} \left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_1 n_2} \right)}} = \frac{n_1(n+1) - 2S_1}{\sqrt{n_1 n_2 (n+1)/3}}$$

WILCOXON:
$$T = \frac{W - \mu_W}{\sigma_W} = \frac{2n_1 n_2 + n_1(n_1 + 1) - 2S_1 - n_1 n_2}{\sqrt{n_1 n_2 (n+1)/3}} = \frac{n_1(n+1) - 2S_1}{\sqrt{n_1 n_2 (n+1)/3}}$$

Laatstgenoemde formule kan dus ook —in plaats van (10.12)— worden gebruikt bij de normale benadering van de toets van WILCOXON.

12.2. H_0 : geen verloop, H_1 : stijgend verloop; $S_{1,2} = 70,5$, $S_{1,3} = 56$, $S_{1,4} = 55$, $S_{2,3} = 69$, $S_{2,4} = 55$, $S_{3,4} = 67$, $\Sigma S_{i,j} = 372,5$, $W_0 = 5,15$, $\sigma_W = 0,8287$, $T_0 = 6,215$, $P_R < 10^{-8}$, H_0 verwerpen.

12.3. H_0 : geen verloop, H_1 : dalend of stijgend verloop. $S_{1,2} = 48,5$, $S_{1,3} = 50$, $S_{1,4} = 39$, $S_{2,3} = 76,5$, $S_{2,4} = 56$, $S_{3,4} = 193,5$, $W_0 = 2,5388$, $\sigma_W = 0,9210$, $T_0 = 2,76$, $P_D = 0,00578$, H_0 verwerpen t.g.v. een stijgend verloop.

12.4. $S_j - \bar{S}$: -13 -13 -3 -3 11 11 10, $K_0 = 698$, $X_0^2(6) = 24,93$, $P < 0,001$, $S_j(n+1) - \bar{S}$: -16 -15 -4 -3 12 13 13, $K_0 = 988$, $X_0^2(7) = 30,245$. Daar de tweede waarde van χ^2 groter is dan de eerste, wordt geconcludeerd dat een stijgend verloop optreedt.

12.5. H_0 : Dezelfde sterfte in kliniek en stad, H_1 : Verschillende sterfte in kliniek en stad. Het is duidelijk, dat men bij de verschillende gewichtsverdelingen in de kliniek en in de stad (met verschillende sterfte per gewichtsklasse) *niet* de totaalcijfers mag vergelijken. Deze zijn immers wegens het relatief grote aantal 'zware' kinderen (met geringe sterfte) en het relatief kleine aantal 'lichte' kinderen (met grote sterfte) gewogen in het voordeel van de kliniek. Men zal dus met de volgende viervelden-tabellen per gewichtsklasse moeten werken:

	500-1000 g		1001-1500 g		1501-2000 g		2001-2500 g	
	K	S	K	S	K	S	K	S
Overleden	2	14	15	13	8	11	6	7
Niet overl.	1	1	16	9	56	60	146	83
	3	15	31	22	64	71	152	90
	16	15	28	25	116	135	229	242

Op deze vier 2x2 tabellen kan de in 12.3.2 beschreven toets worden toegepast (de alternatieve hypothese luidt, dat de verschillen tussen de sterftetekansen van kliniek en stad overwegend hetzelfde teken bezitten, zodat tweezijdig wordt getoetst). Het feit, dat in de eerste tabel te lage verwachtingen voorkomen is hierbij een bezwaar, dat moeilijk kan worden ondervangen. Men vindt dan: $M_0 = -0,4419$, $\sigma_M^2 = 0,066091$, $\sigma_M = 0,2571$, $T_0 = -1,72$, $P_D = 0,0854$. De nulhypothese kan dus bij een 5% drempel niet worden verworpen. Voor een juiste interpretatie van deze gegevens zou men echter over meer gegevens betreffende de in de kliniek en in de stad geboren kinderen moeten beschikken (selectie?!).

ANTWOORDEN BIJ DE OPGAVEN

Hoofdstuk 13

13.2. $T_0 = \frac{1782/25 - 70}{2,5/\sqrt{25}} = 2,56, P_R \cong 0,005.$

13.3. $T_0 = \frac{12,8 - 12,0}{2/\sqrt{25}} = 2, P_R = 0,0228, 2,28\%.$

13.4. a. $\chi^2_0(19) = 19 \times 13/10 = 24,7, 0,2 > P_R > 0,1.$ Met de *W-H*-transformatie vindt men (ongeveer) $T = 0,90, P_R = 0,1841.$

b. $\chi^2_0(19) = 11,4, 0,05 < P_L < 0,10.$ Met de *W-H*-transformatie: $T = -1,40, P_L = 0,0808.$

c. Hier kan men het best direct de transformatie toepassen. Bij $s^2 = 8$ behoort dan $T = -0,61$ en bij $s^2 = 12$ behoort $T = 0,63$, zodat de gevraagde intervalkans ongeveer 0,4648 bedraagt.

13.5. a. $T = 2,8, P_R = 0,00256$ b. $T = -4,2, P_L = 0,0000133$

c. $T_1 = -1,4, T_2 = +1,4, P(T_1 < \underline{T} < T_2) = 0,8384.$

13.6. a. 0,99 b. 0,25 c. 0,05 d. 0,05 e. 0,95 f. $0,005 < P_R < 0,01$
g. $0,10 < P_D < 0,20$ h. 0,95 i. $0,80 < P_D < 0,90$ j. $0,90 < P_D < 0,95.$

13.7. a. -4,303 of 4,303 b. -2,571 of 2,571 c. -2,201 of 2,201
d. -2,069 of 2,069 e. -1,98 of 1,98 f. -1,96 of 1,96.

13.8. $F_{0,95} = 5,80, F_{0,99} = 14,0, F_{0,05} = 0,348, F_{0,01} = 0,226, \text{modus} = 0,6,$
 $\mu_F = 2.$

13.9. a. 0,223 en 6,68 b. 0,150 en 4,48.

13.10.	$F_{0,01}$	$F_{0,05}$	$F_{0,95}$	$F_{0,99}$
a.	0,0361	0,112	4,35	8,45
b.	0,118	0,230	8,89	27,7
c.	0,143	0,264	3,79	6,99
d.	0,158	0,285	2,71	4,14
e.	0,172	0,303	2,17	2,95
f.	0,339	0,461	3,30	5,82
g.	0,472	0,588	1,84	2,40
h.	0,452	0,578	1,52	1,79
i.	0,680	0,758	1,39	1,60

13.11. $T = 2, P_R = 0,0228.$ 13.12. $n = 20.$ 13.13. $s^2 = 7,663953, s = 2,58.$

Hoofdstuk 14

14.1.	χ^2_0	ν	P_D	95% betr. interval
a.	4	2	0,20 — 0,40	$5,42 < \sigma^2 < 790,51$
b.	10	5	0,10 — 0,20	$7,79 < \sigma^2 < 120,34$
c.	16	8	0,05 — 0,10	$9,12 < \sigma^2 < 73,39$
d.	30	15	0,02 — 0,04	$10,91 < \sigma^2 < 47,91$
e.	48	24	0,002 — 0,01	$12,19 < \sigma^2 < 38,71$
f.	70	35	0,00043	$13,17 < \sigma^2 < 33,96$
g.	4,5	9	0,20 — 0,40	$2,37 < \sigma^2 < 16,67$
h.	7,2	9	0,60 — 1,00	$3,78 < \sigma^2 < 26,67$
i.	11,7	9	0,40 — 0,60	$6,15 < \sigma^2 < 43,33$
j.	14,4	9	0,20 — 0,40	$7,57 < \sigma^2 < 53,33$
k.	19,8	9	0,02 — 0,04	$10,41 < \sigma^2 < 73,33$
j.	23,4	9	0,01 — 0,02	$12,30 < \sigma^2 < 86,67$

BIJLAGE IV

- 14.2. a. $6,56 < \sigma^2 < 62,13$ b. $4,37 < \sigma^2 < 19,16$ c. $10,47 < \sigma^2 < 23,29$
d. $7,41 < \sigma^2 < 12,98$ e. $8,13 < \sigma^2 < 12,06$ f. $8,80 < \sigma^2 < 11,28$
- 14.3. a. $F_0(9, 11) = 1,96$, $F_{0,975} = 3,59$, niet significant
b. $F_0(20, 12) = 2,73$, $F_{0,975} = 3,07$, niet significant
c. $F_0(12, 20) = 2,73$, $F_{0,975} = 2,68$, significant
d. $F_0(60, 120) = 1,90$, $F_{0,975} = 1,53$, significant.
- 14.4. $s_1^2 = 0,05785$, $s_2^2 = 0,05563$, $F_0(9, 9) = 1,04$, niet significant.
- 14.5. Toets van BARTLETT:
a. $B_0 = 7,67$, $k - 1 = 3$, $0,05 < P < 0,10$
b. $B_0 = 5,37$, $k - 1 = 5$, $0,30 < P < 0,50$
c. $B_0 = 9,34$, $k - 1 = 6$, $0,10 < P < 0,20$.
- Toets van HARTLEY:
a. Geen uitspraak mogelijk, b. $F_0(\max) = 2,615$, H_0 niet verwerpen,
c. $F_0(\max) = 3,17$, vermoedelijk H_0 niet verwerpen.
- 14.6. $k_0 = 0,8648$, $T_0 = 4,59$, $P_D \ll 0,001$.

Hoofdstuk 15

- 15.1. $T_0 = -2$, $P_D = 0,0456$, verwerpen.
- 15.2. $t_0 = -3,273$, $0,002 < P_D < 0,01$, verwerpen.
- 15.3. a. 44, b. 35, c. 25, d. 87, e. 64, f. 41, g. 153, h. 90, i. 62.
- 15.4. a. 5, b. 1, c. 1 en 2.
- 15.5. $t_0(7) = 3,5$, $P_D \cong 0,01$.
- 15.6. a. Rangtekentoets, $V_0 = 87$, $n = 14$, $P_D < 0,01$, b. $t(15) = 3,464$,
 $0,002 < P_D < 0,01$.
- 15.7. $t_0(19) = -2,269$, $0,02 < P_D < 0,05$, significant.
- 15.8. a. $14,0 < \mu < 16,0$ e. $32,8 < \mu < 47,2$ i. $38,0 < \mu < 42,0$
b. $14,6 < \mu < 21,4$ f. $37,3 < \mu < 42,7$ j. $37,8 < \mu < 42,2$
c. $14,6 < \mu < 21,4$ g. $38,4 < \mu < 41,6$ k. $36,75 < \mu < 43,25$
d. $31,8 < \mu < 48,2$ h. $38,5 < \mu < 41,5$ l. $35,7 < \mu < 44,3$.
- 15.9. a. $65,8 < \mu < 68,2$ en $65,3 < \mu < 68,7$
b. $18,2 < \mu < 23,8$ en $17,2 < \mu < 24,8$
- 15.10. $n = 107$. Men onderstelt, dat de glucosegehalten in de betreffende populatie ongeveer normaal verdeeld zijn.
- 15.11. a. Exacte kans $P_D = 2 \cdot 2^{-16} = 2^{-15} = 0,00003052$, b. Als (a),
c. $t_0(15) = 17,356$ (!), overschrijdingskans astronomisch klein.
d. De verschillen stammen uit een (ongeveer) normaal verdeelde populatie. Onderstelling o.i. gerechtvaardigd. e. $7,3 < \mu < 10,2$.
- 15.12. $1,4 < \mu_v < 4,1$. 15.13. P_D iets groter dan 0,05.
- 15.14. a. Methode I: $s^2 = 0,001490$, $s = 0,039$, methode II: $s^2 = 0,0001359$,
 $s = 0,012$. b. $F_0(10, 10) = 11,04$, $P_D < 0,01$.
- 15.15. a. 2400 — 4000, b. 4000 — 6000, c. 6000 — 8400, d. 8400 — 11200
- 15.16. a. 0,21—2,9 pct, b. 0,48—1,8 pct, c. 21—30 pct, d. 10—21 pct.

ANTWOORDEN BIJ DE OPGAVEN

Hoofdstuk 16

- 16.1. a. $T_0 = -2,79$, $P_D = 0,00528$, H_0 verwerpen,
 b. $-1,48 > \mu_1 - \mu_2 > -8,52$.
- 16.2. a. $t_0(40) = 1,258$, $0,2 < P_D < 0,5$, H_0 wordt niet verwerpen,
 b. $-2,18 < \mu_1 - \mu_2 < 9,38$.
- 16.3. a. $t_0(9) = 1,880$, $t_{0,95} = 1,833$, H_0 verwerpen, b. $t_0 = 1,820$, $f = 7,1$,
 H_0 niet verwerpen.
- 16.4. $F_0(405, 389) = 1,88$. De populatievarianties dienen ongelijk te worden
 ondersteld. Met (16.10) vindt men : $T_0 = 4,82$, zodat H_0 wordt verwerpen.
- 16.5. a. $0,0125 < \mu_1 - \mu_2 < 0,0675$, b. $0,000 < \mu_1 - \mu_2 < 0,080$.
- 16.6. a. $n_1 = 520$, $n_2 = 693$, b. $n_1 = n_2 = 73$.
- 16.7. a. $n_1 = 129$, $n_2 = 257$, b. $n_1 = n_2 = 123$.
- 16.8. a. $T_0 = 3,37$, $P_D < 0,001$, b. $t(16) = 5,235$, $P_D \ll 0,001$,
 c. $t_0(12) = 5,235$, $P_D \ll 0,001$, d. $t_0(16) = 5,285$, $P_D \ll 0,001$,
 e. $t_0(16) = 5,110$, $P_D \ll 0,001$, f. $t'_0 = 0,399$, $P_D < 0,01$,
 g. $t'_0 = 0,383$, $P_D < 0,01$.
- 16.9. a. $t_0(14) = 3,34$, $0,002 < P_D < 0,01$, $0,016 < \mu_1 - \mu_2 < 0,269$,
 b. $t'_0 = 0,264$, $0,01 < P_D < 0,05$. De onderstelling, dat de populatie-
 varianties gelijk zijn lijkt gerechtvaardigd, daar $F_0(7, 7) = 1,305$.
- 16.10. $t_0 = 2,28$, $f = 29,6$, $t_{0,975}(29) = 2,045$, verwerpen.
- 16.11. $0,53 < \mu_A - \mu_B < 4,19$.

Hoofdstuk 17

- 17.1. a. Tussen 3798,82 9 } $F_0(9, 40) = 4,36$, $F_{0,95} = 2,12$,
 Binnen 3876,80 40 } H_0 verwerpen.
 b. De gemiddelden worden gesplitst in de volgende 'groepen':

Populatie	5	2	9	8	10	4	7	1	3	6
\bar{x}	10,4	21,8	23,0	23,4	24,6	25,2	25,4	25,8	41,2	41,8

- 17.2. Het is zonder meer duidelijk, dat de varianties niet significant ver-
 schillen; $F_0(5, 29) = 1,10$, $F_{0,95} = 2,52$, niet verwerpen.
- 17.3. a. Toets van HARTLEY, $F_{\max} = 17/12 = 1,42$, niet verwerpen,
 b. $w_0 = 12,64$, verwerpen.
- 17.4. Toets van HARTLEY op varianties: $F_{\max} = 4,77$, insignificant.
 Tussen 31,9693 2 } $F_0(2, 12) = 3,97$, $F_{0,95} = 3,89$, verwerpen.
 Binnen 48,2680 12 }
 De methode van TUKEY laat zien, dat de significante uitkomst voor-
 namelijk een gevolg is van het hogere gemiddelde bij de populier.
- 17.5. Toets van HARTLEY op varianties: $F_{\max} = 1,58$; niet significant.
 Tussen 73,0167 2 } $F_0(2, 117) = 7,90$, $F_{0,95} \cong 4,80$, verwerpen.
 Binnen 540,7750 117 }
 Het gemiddelde van type 9D bijt uit.

BIJLAGE IV

17.6. Kolommen 70 2 }
 Rijen 150 4 } Kolommen: $F_0(2, 8) = 15,56, P_R < 0,005,$
 Residueel 18 8 } Rijen: $F_0(4, 8) = 16,67, P_R < 0,005.$

17.7. Toets van HARTLEY op de kolomvarianties: $F_{\max} = 2,47, F_{0,95} = 5,67,$
 op rijvarianties: $F_{\max} = 15,79, F_{0,95} = 48,0.$

Uren 2721,34 3 $F_0(3, 30) = 1,72, F_{0,95} = 2,92,$ niet verwer-
 Personen 47698,18 10 pen; daar de geconstateerde gemiddelde ver-
 Residueel 15816,91 30 schillen dus nog zeer goed op toeval kunnen
 berusten, heeft het geen zin hiervan hoogste
 en laagste waarden op te geven.

17.8. Toets van HARTLEY op de varianties: $F_{\max} = 4,59, F_{0,95} = 7,11.$
 Tussen 227,8 4 } $F_0(4, 45) = 5,67, F_{0,995} = 4,32,$ verwerpen, dus:
 Binnen 452,2 45 } *aselecteren!*

17.9. a. De varianties van de $kn = 12$ steekproefjes van $m = 5$ bepalingen
 verschillen niet significant: $F_{\max} = 5,39, F_{0,95} = 51,4.$

	S	V.g.
Gehalten	257,3258	3
Analysten	1,4653	2
$G \times A$	0,4467	6
Residueel	2,1720	48
Totaal	261,4098	59

Interactie:
 $F_0(6, 48) = 1,65, F_{0,95} \cong 2,30.$
 Analysten:
 $F_0(2, 48) = 16,19, F_{0,995} \cong 5,93.$
 De verschillen tussen de anal-
 ysten zijn dus sterk significant.
 Uit de eerste stap van de methode van TUKEY blijkt, dat het gemiddelde van B te laag

is; A en C verschillen niet significant.

b. $\sqrt{s^2_{\text{res}}} = \sqrt{2,1720/48} = 0,213.$

Hoofdstuk 18

18.1. b. 23,25, 23,95, c. 7, d. 23,25, 23,90, 7,1.

18.2. a. $\hat{y} = 0,77 + 1,12x,$ b. $\hat{y} = 67,0 - 1,36x.$

18.3. $\sigma_{xy} = 136, \rho = \frac{1}{2} \sqrt{2} = 0,7071.$ 18.4. $B_{yx} = 0, \rho = 0,$ neen.

18.5. a. $b_{xy} \cdot b_{yx} > 1,$ b. $r_0 = 1543/\sqrt{1560 \times 1136} = 1,159,$ c. Negatieve waarden van de b's zijn onverenigbaar met een positieve waarde van r,
 d. $B_{yx} = 0,5 \times 10/4 = 1,25.$

18.6. $\rho = -0,25.$

Hoofdstuk 19

19.1. b. $\hat{y} = 0,718 + 0,248x,$ c. $s_{y,x} = 0,086, 0,234 < B_{yx} < 0,262,$ mits het linear-normale model bij benadering geldt.

d.

	S	V.g.
Regressie	10,0663	1
Gemiddelden	0,0278	7
Tussen	10,0941	8
Binnen	0,1500	17
Totaal	10,2441	25

Toets op rechtlijnigheid:
 $F_0(7, 17) = 0,45, F_0$ is kleiner dan 1 en de aanpassing is dus zeer goed.
 Opmerking: de hypothese $B=0$ kan op grond van het bij (c) berekende interval bij een 5% drempel worden verworpen.

ANTWOORDEN BIJ DE OPGAVEN

Deze hypothese kan ook worden getoetst door te berekenen

$$F_0(1, 24) = \frac{10,0663/1}{(0,0278 + 0,1500)/24} = 1358,8 (!)$$

- 19.2. b. $\hat{y} = 15,00 + 7,15x$, c. $s_{y,x} = 11,62$, $t_0 = 10,43$, d. Toets van HARTLEY, $F_{\max} = 4,26$, $F_{0,95} = 6,31$, niet verwerpen.

e.

	S	V.g.
Regressie	14688,7826	1
Gemiddelden	443,8174	2
Tussen	15132,6000	3
Binnen	4684,8000	36
Totaal	19817,4000	39

Toets op rechtlijnigheid:
 $F_0(2, 36) = 1,71$, $F_{0,95} = 3,27$.

Opmerking: Ondanks de niet significante uitkomst van de toetsen op de varianties en de rechtlijnigheid krijgt men bij het bestuderen van het spreidingsdiagram (a) de indruk, dat het lineair-normale model bij deze gegevens niet bevredigend is. De varianties blijken groter te worden naarmate de gemiddelde letale dosis toeneemt en het lijkt erop, dat de regressie feitelijk kromlijinig is. Men bereikt veel betere resultaten met de getransformeerde gegevens $x = \log$ (injectiesnelheid) en $y = \sqrt{d} + \sqrt{d+1}$ ($d =$ letale dosis $- 50$). De systematische stijging van de varianties is dan opgeheven, F_{\max} bedraagt slechts 1,29 en de toets op de rechtlijnigheid wijst op een bijzondere fraaie aanpassing (van de regressielijn $\hat{y} = 8,42 + 9,126x$), daar $F_0(2, 36) = 0,032$.

- 19.3. a. De regressie lijkt kromlijinig, de spreiding neemt af naarmate de sterkte van de oplossing afneemt.
 b. Op het oog lijken deze bezwaren opgeheven.
 c. Toets van HARTLEY, $F_{\max} = 4,69$, $F_{0,95} = 7,80$.
 d. $\hat{y} = 4,927 - 3,089x$. e. $F_0(4, 54) = 1,23$, $F_{0,95} = 2,55$.

Hoofdstuk 20

20.1.	r	formule	$n = 5$	$n = 26$	$n = 82$	$n = 401$
0,06		a	0,104	0,294	0,538	1,201
		b	0,120	0,300	0,540	1,200
0,20		a	0,354	1,000	1,826	4,077
		b	0,400	1,000	1,800	4,000
0,60		a	1,299	3,674	6,708	14,981
		b	1,200	3,000	5,400	12,000

De overeenstemming is beter, naarmate r_0 dichter bij 0 ligt en n groter is.

- 20.2. 5%: a. niet, b. wel, c. wel, d. niet, e. wel, f. wel.
 1%: a. niet, b. niet, c. wel, d. niet, e. niet, f. wel.
- 20.3. a. $r_0 = 0,821$, verwerpen, b. $0,611 < \rho < 0,923$.

BIJLAGE IV

- 20.4. $a. r_0 = 0,740$, $b. T = 1,34$, niet significant, $c. \hat{y} = 1,69 + 0,22x$,
 $\hat{x} = 136,07 + 2,49y$.
- 20.5. $r_0 = 0,855$, 95%: $0,730 < \rho < 0,925$, 99%: $0,674 < \rho < 0,939$.
- 20.6. $a. r_0 = -0,928$, $b. 0,838 < \rho < 0,969$ (indien populatieverdeling twee-
dimensionaal normaal wordt ondersteld), $c. \hat{W} = 121,71 - 0,498 H$, of
als $x = \log$ hemoglobine en $y = \log$ hartgewicht: $\hat{y} = 0,715 - 0,498x$.
- 20.7. $c. r_{xx} = 0,9849$, $d. \hat{x} = 0,376 + 0,163x$, $\hat{x} = -1,978 + 5,962x$,
 $s^2_{x..} = 0,017415$, $s^2_{x..} = 0,637763$, $s_{x..} = 0,132$, $s_{x..} = 0,799$.
- 20.8. $a. \bar{x} = 24,17$, $\bar{y} = 69$, $s_x = 11,767$, $s_y = 7,649$, $b. r_0 = 0,333$,
 $-0,054 < \rho < 0,746$, $c. \hat{y} = 63,76 + 0,217x$, $s_{y..} = 7,35$.
- 20.9. $T_0 = (1,045371 - 0,792814)/\sqrt{1/200 + 1/150} = 2,34$, verwerpen.
- 20.10. $X_0^2(4) = 1,846$, $0,7 < P_R < 0,8$. 20.11. $X_0^2(6) = 22,556$, $P_R < 0,001$.

BIJLAGE II (deel I)

Paragraaf 1

1. $a. \sum_{i=8}^{12} x_i$ $b. \sum_{i=2}^5 x_i$ $c. a \sum_{i=3}^5 x_i y_i$ $d. \sum_{i=1}^4 x_i - 4a$.
2. $a. 59$ $b. 7$ $c. 59/15$ $d. 15$ $e. 21$ $f. 3,6$ $g. 3481$ $h. 271$ $i. 124$.

Paragraaf 2

- (blz. 320) 3. $a. (n+1)n$ $b. 7$ $c. 5$ $d. 7$ $e. 14$ $f. 6!$
- (blz. 322) 2. $a. 560 + 350 + 56 = 966$ $b. 1287$ $3. 1526$ $4. 3^{15} = 14.348.907$.

Paragraaf 3

1. $a. 2^3 5^6$ $b. 2^{10} 3^{12}$ $c. 2^4 3^5 5^3 7^2$ $d. 2^{15} 3^{12}$.
2. $a. \frac{a^5 d^3}{b^7 c^3}$ $b. \frac{pt^5}{q^4 r^3 s}$ $c. \frac{c^5}{a^6 b d^7}$ $d. \frac{w^3 x}{y^2 z^{11}}$.
3. $a. \sqrt[4]{x^8}$ $b. \sqrt{c+d}$ $c. \sqrt[4]{y^8}$ $d. \sqrt[4]{b^4 + c}$ $e. \sqrt{3^2}$.
4. $a. 0,025298$ $b. 63,2456$ $c. 533,73$ $d. 2^3 3^2 5^{-2}$ $e. 16,39$
 $f. 0,045891$ $g. 37,4032$ $h. 0,26857$.

Paragraaf 4

2. $a. 1,09235$ $b. 1,60531$ $c. 0,06144 - 3$ $d. 0,37181 - 1$
 $e. 4,30103$ $f. 3,73038$ $g. 0,20412 - 3$ $h. 0,13259$.
3. $a. 15,3$ $b. 4222$ $c. 0,6628$ $d. 0,0005594$ $e. 1,387$ $f. 999,9$
 $g. 0,0553$ $h. 1,01$.
4. $a. 0,2713$ $b. 2,021$ $c. 131,5$ $d. 204,3$ $e. 44,79$.

Paragraaf 5

1. $a. 77,77$ $b. 22,2$ $c. 17,1$ $d. 7,924$ $e. 75,7$ $f. 65,13$ $g. 0,37$.
2. $a. 625$ $b. 911,0$ $c. 12,86$ $d. 2,924$ $e. 92,9$.

Paragraaf 6

3. $a. y = 3x$ $b. y = -1\frac{1}{2}x$ $c. y = 2x + 7$ $d. y = -2\frac{7}{8} + \frac{1}{8}x$ $e. y = -7\frac{7}{8} - \frac{3}{8}x$.

REGISTER

- A*, constante in regressievergelijking
a, schatting van *A*
 α (alpha), kans op fout eerste soort, onbetrouwbaarheid
AV, ware activiteitsverhouding
av, schatting van *AV*
 Aanpassing, χ^2 -toets voor, 215-216
 nadelen van, 553
 vergelijking met *K-S* toets, 217
 vrijheidsgraden, 215
 Abscis, 329
 Activiteitsverhouding, 589-590
 bij directe ijking, 591-592
 bij indirecte ijking, 593-604
 Additiviteit, 441, 450
 Afknotting, bij sequente toetsing, 151
 Afronden, 36, 326
 Afrondingsinterval, 3, 4, 33
 Aftrekken van constante, 62-64
 Algebra, begrippen uit, 318-330
 Alternatief (zie dichotomie)
 Alternatieve hypothese, 121, 185
 één- of tweezijdig, 135
 Arbitraire klassewaarden, 618
 Aselect, 5, 6, 19
 Aselecte
 getallen, 9, 10, 315
 permutaties, 562, 649
 splitsing, 11-12
 steekproeven, 6, 9-11, 18
 volgorde, 171, 262, 379
 Aselecteren, 11-12, 14
 Aselectiviteit, 14, 261
 Asymmetrie (zie scheefheid)
 Attribuut, 2, 3, 43

B, populatieregressiecoëfficiënt
b, schatting van *B*
 β (beta), kans op fout tweede soort
 BARTLETT, toets op gelijkheid van varianties, 373-375, 442, 450, 491
 Beeldgrafieken, 48
 Behandeling, 19, 23
 BELLINSON, toets tegen verloop, 378

 Benaderende getallen, 325
 Benadering, van verdelingen
 binomiale door normale, 109
 binomiale door POISSON, 90
 hypergeometr. door binom., 88
 POISSON door normale, 405
 X^2 door χ^2 , 175-176
 Beslissingsschema, 123-124, 134
 Best-passende
 binomiale verdeling, 87
 χ^2 -verdeling, 176, 177, 180
 normale verdeling, 113
 POISSON-verdeling, 93, 405
 Betrouwbaarheid, van een waarnemingsmethode, 401
 Betrouwbaarheidscoëfficiënt, 156, 165, 391
 Betrouwbaarheidsgordel, 159
 Betrouwbaarheidsgrenzen, 155-159, 165-166; van
 activiteitsverhouding, 592, 599, 603
 correlatiecoëfficiënt, 525-526
 *ED*₅₀, 608, 610, 613, 615
 fractie, 156-164, 405-407
 gemiddelde, 390-395, 437
 gemiddelde verschil, 398
 metingen, 399-403
 percentage (zie fractie)
 spreiding, 367-371
 standaarddeviatie, 367-371
 tellingen, 405-407
 variantie, 367-371
 verschil van twee gemiddelden, 420-422
 Betrouwbaarheidsgrenzen bij regressie, van
 B, 498-502
 A, 501
 μ_y , 500
 $\mu_{y,j}$, 500
 Betrouwbaarheidsinterval (zie betrouwbaarheidsgrenzen)
 Betrouwbare cijfers, 326
 Bewijsmiddel, statistiek als, 15, 21, 562

REGISTER

- Bgsin transformatie (zie transformaties)
 Bimodaal (zie tweetoppig)
 Binomiaalcoëfficiënt, 82-83
 Binomiale
 toetsen, 120, 142, 560
 schattingen, 156
 Binomiale verdelingen, 82-88
 berekening, 86-87
 best-passende, 87
 formule, gemiddelde, variantie, 82
 normale benadering, 109
 Poisson-benadering, 90
 Biologische standaardisatie (ijking),
 588-617
 Biseriaal correlatiecoëfficiënt, 540
 Blanco
 groep, 19
 waarnemingen, 20, 601
 Breedte
 klasse-, 3, 35, 36, 43, 44
 van schattingsinterval, 156, 163,
 165, 395, 422
 BROWNLEE c.s., sequente ijkingsme-
 thode, 616-617
 Bruikbaarheid (van schatter), 155
 Buigpunt, 103
- c*, (constante) klassebreedte
 CAMP en MEIDELL, ongelijkheid van,
 119
 Cartogram, 48
 Categorisch systeem, 2-4, 32, 67,
 170-171 (zie ook dichotomie)
 toetsen betr., 180-181, 190-213,
 618-629
 Causaliteit, 532
 Cellen, telling van, 403-407
 Centrale
 maten = gemiddelden
 tendentie, 34, 115
 χ^2 (chi-kwadraat)-toetsen
 aanpassing, 215, 217, 220, 553
 categorische systeem, 181-186
 dubbele dichotomie, 213
 stochastische onafhankelijkheid,
 211-213
 2×2 tabel, 197-200, 541, 624
 $2 \times r$ tabel, 204-207
 $k \times 2$ tabel, 208
 $k \times r$ tabel, 209-210, 614-629
 variantie, 365-366
 χ^2 -verdelingen, 176-179, 343-344
 best-passende, 176, 180
 fractielen, 178
- gemiddelde en modus, 180
 relatie met *F*-, *t*- en *T*-verdelingen,
 352
 transformatie tot *T*, 179, 344
 Cirkeldiagram, 48
 Cluster, 586
 COCHRAN,
 toets betreffende *k* dichotome popu-
 laties, 190-192
 toets op *k* varianties, 377
 Codelijst, 41
 Coëfficiënt, zie
 betrouwbaarheids-
 binomiaal-
 kurtosis-
 overeenstemmings-
 phi-
 rangcorrelatie-
 regressie-
 scheefheids-
 variatie-
 Combinaties, 75, 321
 Combinatie van categorieën, 184, 207
 212, 216
 Componenten, bij variantie-analyse,
 441-446, 450
 Conclusie(s), 7, 16-17, 22, 23
 Conditie, experimentele, 582-583
 Consequentie, van fouten bij toetsing,
 124
 Continu (zie grootheid)
 Continue kansverdelingen, 103
 Continuïteitscorrectie, principe, 111
 χ^2 -toets op 2×2 tabel, 199-200
 andere toetsen, 225, 227, 241, 242,
 251, 257, 269, 272, 624
 Contrôlegroep, 19
 Correctie
 (zie continuïteits-)
 (zie gelijken)
 Correlatie
 causaliteit, 532
 heterogeniteit bij, 534
 negatieve, 250, 483-484
 positieve, 250, 483-484
 problemen, 479-482
 pseudo-, 533
 rang-, 249-260
 rekening, 514-545
 tabel, 38, 529
 tussen gemiddelde en variantie, 346,
 361
 uitvallers bij, 535
 volkomen, 484

BIJLAGE V

- Correlatiecoëfficiënt, 483-485, 521-543
 als correlatiemaat, 532
 als toetsingsgrootheid, 521-528, 531, 539-542
 berekening, 483; uit correlatietabel, 528-531
 betrouwbaarheidsgrenzen, 526
 biserial, 540
 formules, 483, 530, 532, 543
 gebruik van, 531-536
 kansverdeling, 521-523
 rangcorrelatie, 253, 258
 standaardfout, 522-523
 toetsing hypothesen, 521-527
 z -transformatie van, 524-528
 Covariantie, 482
 -analyse, 521, 591
 Cumulatieve frequentie (zie somfrequentie)
 Cyclisch verloop, 262, 267, 271, 274-276, 378
 Cijfers, aantal betrouwbare, 326
- Dagritme, 404, 566
 Deciel, 58
 Deelverzameling, 1, 488, 503
 Delen, door constante, 62-64
 Detectie, 16, 23, 27, 276, 563
 Diagram (zie grafiek)
 Dichotomie, 2, 79-86, 120, 154-155, schatting, toetsing, betr. P (zie fractie), 183-184
 toetsen, 190-192, 193-201, 208, 213, 282-283, 286-290, 624
 Differentiatie van cellen, 406
 Directe standaardisatie (ijking), 589-592
 Discontinu (discreet), 3
 Discrepancie, tussen waargenomen en verwachte frequenties, 175-176, 187
 DIXON, uitbijtertoets, 409
 DIXON en MOOD, sequente ijkingsmethode, 613-616
 Doeltreffendheid, 155
 DOORNBOS, uitbijtertoets, 378
 Dosis, 588-589
 individueel-effectieve, 589, 593
 mediaan-effectieve, ED_{50} , 604
 schalen, 590
 Dosis-werkingskromme, 510
 Drempelwaarde (zie onbetrouwbaarheidsdrempel)
 Driehoek van PASCAL, 83
- Dubbel blinde proef, 19
 Dubbele dichotomie, 213
 Dubbel logaritmisch papier, 50
 Dummy, 20
 Duplo, 401-403
- e , grondtal natuurlijke logaritmen, 90, 325
 e , verwachte frequentie
 η (eta), zie kolom-component
 ED_{50} (zie mediaan-effectieve dosis)
 Eéndimensionale verdelingen, 464
 Eéntoppigheid, 115
 Eénzijdige
 overschrijdingskans, 77, 139
 toetsing, voorwaarden voor, 135
 Element, 1
 Ellipsen, bij tweedimensionaal normale verdelingen, 515-517
 Enquête, 21
 Ethiek, medische, bij experiment, 20
 Exactheid door toetsing, 7, 153
 Experiment, 18-19, 566-583
 Exponent, 422
 Exponentiële groeicurve, 491
 Extrapolatie
 bij regressieproblemen, 493
 in de tijd, 563
 Evenwijdigheid van regressielijnen, toets op, 595-597
- F , somfrequentie
 F (zie F -verdelingen)
 F_{\max} (zie toets van HARTLEY)
 f , frequentie
 f , aantal vrijheidsgraden bij toets van WELCH
 Factor (bij experiment), 567
 Factoriële proefschema's, 567-578
 Faculteit, 320
 F -toets van FISHER
 twee varianties, 371-372
 bij regressie-analyse, 506
 bij variantie-analyse, 427, 435, 436, 439, 445, 456-458, 570, 575, 579, 581
 F -verdelingen, 349-351
 definitie en toelichting, 349
 fractielen, 349-351
 gemiddelde en modus, 349
 relatie met t , χ^2 en T , 352-353
 transformatie tot T , 561

REGISTER

FISHER

- 'Design of experiments', 566
- F*-toets (zie aldaar)
- normaliteitstoetsen, 553-557
- toets op 2×2 tabel, 193-196
- transformatie van χ^2 tot *T*, 179
- z*-transformatie, 524-528
- Fluctuatie, 2, 37, 565
- Fout
 - van de eerste soort (α), van de tweede soort (β), 125, 132, 137, 153, 387
 - systematische, 400, 546
 - waarnemings-, 399-408
- Fractie
 - betrouwbaarheidsgrenzen, 156-164, 405-407
 - toetsing, 133-152, 560
- Fractielen, 58, 106-107, 119, 178, 348, 349-351, 522
- Frequentie
 - absolute, relatieve, som-, 3
 - marginale, rand-, 197, 199
 - verwachte, 87, 173, 175, 176, 216
- Frequentiedichtheid, 44
- Frequentiepolygoon, 43, 45, 46
- Frequentiequotiënt, 3, 66
- Frequentieverdeling, 3, 4
 - ééntoppige, 115
 - gemiddelde van, 55
 - grafieken van, 43-46
 - karakterisering van, 114-119
 - mediaan van, 54
 - modus van, 53
 - opstellen van, 32-38
 - pseudo-normale, 114, 117-118
 - relatieve, 3, 46, 114, 116
 - scheve, 85, 115, 119
 - spreiding van, 60
 - tweedimensionale, 464-466, 482-485
 - waargenomen, 5
- FRIEDMAN, toets op *k* verwante steekproeven, 230-235
- Functie, 329
- Functionele relatie, 489, 543-544

- g_1 , scheefheidscoëfficiënt, 554
- g_2 , kurtosiscoëfficiënt, 555
- GAUSS-LAPLACE (zie normale verdeling)
- GEARY, toets tegen kurtosis, 558
- Gelaagde, gelede steekproef (zie laagsgewijs)

- Gelijken, correctie voor rangtekentoets, 226-228
- serietoets, 270, 272
- toets van FRIEDMAN, 233-234
- toets van KENDALL, 257-258
- toets van KRUSKAL en WALLIS, 348
- toets van SPEARMAN, 253
- toets van WILCOXON, 241-242
- Gemiddelde (rekenkundig)
 - berekening, 54-56; via hulpgrootheid, 63-64
 - betrouwbaarheidsgrenzen, 390-395, 398, 437
 - bij karakterisering, 117
 - gewogen, 57
 - kansverdeling, 340-342, 347-348
 - standaardfout, 341
 - toetsen inzake, 382-389, 396-398
 - van kansverdeling, 68, 71
 - van normale verdeling, 102
 - van (relatieve) frequentieverdeling, 55
- Gemiddelde
 - afwijking, 58
 - rangnummers, 621 (zie ook bij gelijken)
 - steekproefgrootte, 149-150
- Gemiddelden, 53-57
 - twee steekproeven, verschil van schatting, 420-422
 - toetsing, 411-420
 - k* steekproeven (zie variantie-analyse)
- Gepaarde waarnemingen, 170, 188, 221, 396, 449
- Getallen, aselecte, 9, 10, 315
- Gewogen gemiddelde, 57
- Grafieken, grafische voorstellingen, 43-53
 - beeldgrafiek, 48, 49
 - cartogram, 48
 - cirkeldiagram, 48, 50
 - frequentiepolygoon, 45-46
 - histogram (= kolommendiagram)
 - kolommendiagram, 43
 - sequente toetsing, 148
 - somfrequenties, 45
 - spreidingsdiagram, 38, 47-48
 - staafdiagram, 43-44, 47
- Grafiekenpapier, 50
- Grieks-Latijns vierkant, 583
- Groeikromme, logaritmische, 479
- Groepen, steekproeven bestaande uit, 586-587

BIJLAGE V

- Groondtal, 322
 Groothed
 continu, discreet, 3, 325
 stochastische, 67
 GSG-kromme, 150-151
- H*, hypothese (H_0 , H_1)
 Halve kwartielsafstand (zie kwartie-
 leninterval)
- HARTLEY, toets op homogeniteit van
 twee varianties, 376-377, 429, 442,
 450, 491, 594
- Histogram, 43
- Hoofdeffect, 572
- Homogeniteit van varianties, 376-377,
 419
- Hoektoets, 172
- H*-toets (zie KRUSKAL en WALLIS)
- Hulpgrootheden
 bij berekening van gemiddelde en
 spreiding, 63-64
 bij correlatierekening, 528-531
 bij regressie-analyse, 508-511
 keuze van nulpunt, 63
- Hypergeometrische verdelingen, 79-
 81, 95
 benadering van, 88
 berekening van, 80-81
- Hypothese, 121-132
 enkelvoudig, samengesteld, 121, 152
- i*, breedte van betrouwbaarheidsinter-
 val
- Indirecte standaardisatie (ijking), 593,
 604
- Individueel-effectieve dosis, 589, 593
- Informatie, verlies aan,
 bij kwalitatieve toetsen op waarden,
 192, 201
 bij combinatie van categorieën, 184,
 553
- Integreren, 104
- Interactie, 441, 450, 454, 456, 459,
 568, 571, 584, 632
 -component, 450-451, 457
 -kwadratensom, 453, 455, 574, 575,
 577, 578
 -variantie, 454, 579
- Interval, 3, 117
 -breedte (zie schattings-)
 -grenzen (zie klassegrenzen)
 -kans, 104-105
 -schattingen, 155-156
- Inventaris, 12, 13
- Iteratie(s), 266, 605
- $k \times r$ tabel, 209-212, 621-629
- $k \times 2$ tabel, 208, 282
- Kans (definitie), 66-67
 -regels, 69
 -rekening, 66-78
- Kansverdelingen, discrete, 67
 gemiddelde, 68
 herleiden van, 71
 standaardiseren, 68, 98
 variantie, 67
- Kansverdelingen, continue, 103-104
- Karakterisering, 114-118
- KENDALL, rangcorrelatietoets, 254-
 260, 264
- Kenmerk, 2
- Kettingen, 266
- Klasse, 3
 -breedte, 3, 35, 36, 43, 44
 -grenzen, 3, 35, 111
 -midden, -waarde, 3, 34-36
 modale, 53, 114
 open, 37
- Klassieke
 methoden, 8, 338-339, 546-552
 proefopzetten, 567
- KLERK-GROBBEN en PRINS, 200-201
- KOLMOGOROV-SMIRNOV, toets, 217-
 220, 553
- Kolom-component, 442, 450
- Kolomendiagram, 43
- Kritieke
 waarde, 123
 zone, 123
- Kromlijngige regressie, 469, 473, 477-
 479, 511
- KRUSKAL en WALLIS, toets op k ase-
 lecte steekproeven, 246-249, 427,
 623
- K-S* toets (zie KOLMOGOROV-SMIRNOV)
- Kwadrant, 328
- Kwadraten
 methode der kleinste, 469, 473-479
 -som, 426, 428, 442-444, 448, 453,
 455, 572, 573-575, 579, 632
- Kwartiel, 58, 558
- Kwartieleninterval, 58
- Kurtosis, 114, 553, 555, 558
 coëfficiënt, 555
 toetsen, 555, 558

REGISTER

- Laagsgewijs aselechte steekproeven, 585
 Lambda (λ), zie interactiecomponent
 Latijns vierkant, 579-582
 LD_{50} (zie mediaan-letale dosis)
 Lengte, van voortschrijdend gemiddelde, 611
 Leptokurtosis, 114, 554
 LINDLEY, methode van, 543 (bij functionele relaties)
 Lineariteitstoets, 503-507
 Lineair-normale regressie, 488-511 (zie ook regressie-analyse)
 LITCHFIELD EN WILCOXON, grafische ijkingsmethode, 605
 Logarithmen, 323-325
 Logarithmische
 groeikromme, 479
 transformatie (zie transformaties)
 Lognormale verdelingen, 546, 559, 590
 LORD, toets op 2 aselechte steekproeven, 416
 Loterijmethode, 9
 Lotingstabellen, 11, 562, 649
- $M = \log AV$
 $m = \log ED_{50}$
 μ (mu), populatiegemiddelde
 Marginale frequenties (zie randfrequenties)
 Mantisse, 324
 Mark-sensing, 42
 Mediaan, 54, 170, 172, 188, 266
 Mediaan-effectieve dosis, 604
 Mediaantoeets
 k steekproeven, 208-209
 2 steekproeven, 201-202
 Meervoudige regressie-analyse, 512
 Meten, 18, 22, 325
 Metingen, nauwkeurigheid van, 399-403
 Midden, klasse- (zie klassewaarde)
 Millimeterpapier, 50
 Modale klasse, 53, 114
 Model, 216, 339
 bij correlatie, 479, 514-516, 521
 bij regressie, 467, 488-489, 490-491, 495
 bij variantie-analyse, 425, 440-441, 450, 456, 457, 570, 571, 579
 normaal, 339, 546, 547
 Modus, 53
 Monster (zie steekproef)
 Monte-Carlo methode, 76, 357
- Mood, toets op *k* verwante steekproeven, 192
 MOIVRE, DE (zie normale verdeling)
 Multinomiale verdelingen, 95-96
- N*, populatie-omvang
N (μ, σ)-verdeling (zie normale verdeling)
n, steekproefomvang
 ν (nu), aantal vrijheidsgraden
 Naaldsorteren, 40
 Natuurlijke logarithmen, 325
 Nauwkeurigheid van
 tellingen, 403-407
 waarnemingen, 399-403
 Negatieve correlatie, 250, 483
 Niet-normale verdelingen, 114-119, 169, 339, 341, 546, 550
 Niveau (bij experiment), 567
 Normaal waarschijnlijkheidspapier, 113, 553
 Normale benadering, van
 binomiale verdelingen, 98-100, 109-112
 χ^2 -verdelingen, 179, 344
 F-verdelingen, 561
 hypergeometrische verdelingen, 197
 Poisson-verdelingen, 405
 STUDENT (*t*)-verdelingen, 347, 561
 verdeling van de spreiding, 346
 verdeling van de correlatiecoëfficiënt, 521
 verdelingen, andere, 225, 241, 251, 257, 269-272, 273, 281, 283, 288
 Normale verdelingen, 98-113
 als limiet van binomiale, 98-100
 best-passende, 109, 113
 buigpunt, 103
 fractielen, 106-107
 transformatie tot *T*, 102
 tweedimensionaal, 514-515
 relatie tot *F*, *t* en χ^2 , 352-353
 vorm, 102
 Normaliteitstoetsen, 550, 553-558
 χ^2 -toets voor aanpassing, 553
 van FISHER, 553-556
 van GEARY (kurtosis), 558
 van KOLMOGOROV-SMIRNOV, 553
 van PEARSON, 553
 scheefheidstoets, 558
 Normaliteitsonderstelling, 169, 339, 546-552
 Nulhypothese, 16, 121

BIJLAGE V

- OC-curve (zie werkingskromme)
 Omvang, 1, 6, 138, 153, 385
 Onafhankelijkheid, (zie stochastische)
 Onbesliste gevallen, 189, 192
 Onbetrouwbaarheid, 7 (zie ook onbetrouwbaarheidsdrempel)
 Onbetrouwbaarheidsdrempel, 133, 563-564
 Onderscheidingsvermogen, 125, 135-137, 174, 385, 413
 Onderzoek, 15-27
 Onderzoeking, 17, 18-19
 Ongelijkheid
 van CAMP en MEIDELL, 119
 van TCHEBYCHEFF, 119
 van varianties, 419
 Onzuiverheid, 13
 Oog, op het oog trekken van regressielijnen, 468-469
 Oorsprong, 328
 Operationele definitie, 1
 Open klasse, 37
 Optelbaarheid
 van χ^2 -waarden, 289-290
 van componenten bij variantie-analyse, 441, 450
 Orde (van interacties), 572-573
 Ordinaat, 329
 Orthogonale polynomen, 479
 Oscillatie, 262, 378
 toetsen inzake, 265, 271, 378
 Overeenstemmingscoëfficiënt, 234
 Overschrijdingskans, 77, 104-105, 139, 563
- P*, kans, populatiefractie
p, steekproef fractie
 Parabool, 478
 Parallel line assay, 593, 594-600
 Parameter, 73, 156, 165
 Parameter vrij (zie verdelingsvrij)
 PASCAL, driehoek van, 83
 PAULSON, transformatie van *F* en *t*, 561
 Percentage (zie fractie)
 Percentiel, 58
 Permutaties, 320, 231, 250, 256, 266
 aselecte, 562, 649
 Phi-coëfficiënt, 541
 Placebo, 19, 20
 -effect, 20
 -groep, 19, 20
 Planning, 15, 16, 17, 22, 562 e.v.
- Platheid (platykurtosis), 114, 554
 Plus-mintoets, tegen verloop, 272-273
 POISSON-verdelingen, 90-94
 als benadering van binomiale, 90
 als exacte kansverdeling, 93
 berekening, 92
 best-passende, 93, 405
 normale benadering, 405
 gemiddelde = variantie, 91-93
 Polygoon (zie frequentie-)
 Polynomen (zie orthogonale)
 Ponsconcept (-document), 21, 42
 Ponskaartensysteem, 21, 41
 Populatie, 1
 -verdeling, 4
 Positieve correlatie, 250, 483
 Potency ratio (zie activiteitsverhouding)
 Probit, 605
 -analyse, 605
 -transformatie, 605
 Productregel, 69
 Proefdieren
 aseleteren van, 14
 gevoeligheid, homogeniteit, 600
 Proefgroep, 19
 Proefneming (zie experiment)
 Proefopzet, 15-24, 132, 144, 170-171, 385-389, 395, 412, 422, 466, 479, 562-587, 589, 593, 604
 Proefschema's, statistische, 22, 567-585
 factoriële, 567-576
 fractioneel factoriële, 567
 Grieks-Latijnse vierkanten, 583
 Latijnse vierkanten, 579-582
 symmetrische $2k$ -punts, 594-600
 symmetrische ($2k + 1$)-punts, 600-604
 voordelen van, 568
 Pseudo-correlatie, 533
 Pseudo-normale verdelingen, 114, 117, 118, 339, 342, 344, 357, 400, 491, 546, 549, 550, 551, 552
 Puntendiagram, 37
 Puntenwolk, 518, 519, 534, 535, 536
 Puntchatting, 155
- Q*-toets (van COCHRAN), 190-192
 Quasi-normaal (zie pseudo-normaal)
- R*, spreidingsbreedte
r, steekproefcorrelatiecoëfficiënt
ρ (rho), populatiecorrelatiecoëfficiënt

REGISTER

- Randfrequenties, 197, 199, 205, 207, 209
- Randponskaart, 40
- Rangnummertoeetsen, 168, 549
 FRIEDMAN (methode der m rangschikkingen), 230-235
 KRUSKAL en WALLIS (H -toets) 246-249
 rang-tekenoets (symmetrietoets van WILCOXON), 221-231
 twee-steekproevetoets van WILCOXON, 237-243, 269, 540, 547, 620
- Rangcorrelatiecoëfficiënt, van
 KENDALL, 258
 SPEARMAN, 253, 539
- Rangcorrelatietoets, van
 KENDALL, 254-258, 264
 SPEARMAN, 249-254, 264, 539
 keuze, 258-260
- Rangschikken, 234
- Rangschikkingen (zie permutaties)
 methode der m (zie FRIEDMAN-toets)
- Reactie, 589, 604
- Rechtlijnigheid, 468, 473, 484-485
 bij lineair-normale model, 488-489
 bij tweedimensionaal normale model, 514-517
 toets betreffende, 503-507
- Reciproke, 199, 322
 transformatie (zie transformaties)
- Recursieformule, voor
 binomiale verdeling, 86
 Poisson-verdeling, 92
- Reduceren, 6, 68
- Reductie, 32, 37, 68
- Registratie, 21, 23, 25
- Regressie, 466-579
 -coëfficiënt, 475-479, 483, 498, 502, 508-509, 530, 593, 597, 601-602
 -lijn, 468-472, 473-479, 483-485, 488-491, 504, 514-516, 543
 -problemen, 466
 -vergelijking, 468, 473-479, 483, 488-490, 514, 517, 519, 530, 543
- Regressie-analyse, lineair-normaal
 berekening via hulpgrootheden, 508-510
 betrouwbaarheidsgrenzen en toetsing:
 van A , 501
 van B , 498, 502
 van $\mu_{y,j}$, 499-500
 van μ_y , 500
- van $\sigma^2_{y,x}$, 497
 grensgevallen, 489
 meervoudige, 511
 model, 488-489
 toets op evenwijdigheid, 595-597
 toets op rechtlijnigheid, 503-507
- Regressie-analyse
 bij parallel line assay, 595-597
 bij probitmethode, 605
 bij slope ratio assay, 600-601
 via bgsin transformatie, 606-607
- Rekenkundig gemiddelde (zie gemiddelde)
- Replicatie, 450, 567, 571
- Representatief, 13
- Reproduceerbaarheid, 401
- Responsie, 467, 567, 588, 593
- Residuele (rest-) variantie, 443, 448, 581, 583, 585
- Restrictie, 180
- Runtoets (zie serietoets)
- Rij-component, 442, 450
- S , kwadratensom, of som van rangnummers
- s , steekproefspreiding (-standaarddeviatie)
- σ , populatiespreiding
- Schatten, 7, 154-156, 367-371, 390-394
- Schatter, 154-155
- Schatting, 154-155
 aannemelijkste, 155
 bruikbaarheid van, 155
 doeltreffendheid van, 155
 interval-, 155-156
 punt-, 155
 zuiverheid van, 154
- Schattingsinterval, 156 (zie verder betrouwbaarheidsgrenzen)
- Schattingstheorie (zie schatten)
- Scheefheid, 85, 100-101, 115, 119, 553-554, 558
- Scheefheids
 -coëfficiënt, 554, 558
 -toets van FISHER, 554-555
 -toets, via kwartielen, 558
- Scherpte (zie onderscheidingsvermogen)
- Scheurlijn, 43
- Schijnbehandeling, -middel, 19-20
- Score, 567, 618
- Seizoens
 invloed, 20, 23

BIJLAGE V

verloop, 566
 Selectie, 23-26, 185
 Semi-logarithmisch papier, 50
 Sequentie ijkingsmethoden, van
 BROWNLEE, c.s., 616-617
 DIXON en MOOD, 613-616
 Sequentie toetsing, 144-152
 afknotting, 151
 grafische methode, 148
 GSG-kromme, 149-150
 tabellarische methode, 147
 tweezijdige toetsing, 152
 Serietoets, op
 twee aselechte steekproeven, 269-270
 verschillen, 271-272
 waarnemingen, 265-269
 Significant(ie), 134, 153, 564
 Simultane verdeling, 71, 346, 361,
 483-484
 Slope ratio assay, 593, 600-604
 Somfrequentie, 3, 4
 -polygoon, 45
 -verdeling, 4, 37, 46, 176-177
 Som
 -functie, 3
 -regel, 69
 -teken, 318, 426, 454
 Sorteert
 -apparaat, 41
 -kaartjes, 39
 -methoden, 39-42
 SPEARMAN, rangcorrelatietoets, 249-
 253, 264, 269, 539
 SPEARMAN en KÄRBER, ijkingsmetho-
 de, 607-611
 Spitsheid, 114, 342, 553
 Splitsing in groepen, van gemiddelden,
 431-435
 Spreiding (zie variabiliteit)
 Spreiding (standaarddeviatie)
 berekening, 54-62, via hulpgroot-
 heid, 63-64
 betrouwbaarheids grenzen, 367-371
 bij karakterisering, 117
 kansverdeling, 345-346
 schatting uit 2 (h) steekproeven,
 355-356
 standaardfout, 346
 toetsen betreffende, 365-367
 van een kansverdeling, 68, 71
 van een frequentieverdeling, 60-61
 van een normale verdeling, 102
 Spreidingen, toetsen inzake (zie va-
 rianties)

Spreidings
 -breedte, 35, 57, 416, 431
 -diagram, 38, 47-48
 -maten, 32, 57-62
 Staafdiagram, 43-44, 47
 Standaardfout, van een
 correlatiecoëfficiënt, 522-523
 gemiddelde, 341
 spreiding (standaarddeviatie), 346
 verschil tussen gemiddelden, 421,
 452
 Standaardiseren, 68, 98, 100, 102
 Standaardisatie, biologische, 588-617
 directe, 589
 indirecte, 593, 604
 microbiologische, 593
 Standaardnormale verdeling, 102, 105-
 107
 Standaardpreparaat, 588-590
 Statistica, 6
 Statistiek
 beschrijvende, 6-8
 bevolkings-, 8
 mathematische, 6-8
 Statistische
 analyse, 7, 21-27
 grootheid (zie steekproeffunctie)
 Steekproef (zie ook steekproeven)
 -experimenten, 151, 342, 357, 549
 -functie, 73-74, 121, 358
 -regel, 6
 -ruimte, 74, 121-122
 -trekken, 5, 9-11, 17
 -technieken, 9-15, 585-587
 -verdeling, 74, 341
 Steekproefomvang, 6, 22, 153, 548-
 550, bij:
 binomiale schatting, 163
 binomiale toetsing, 130-132
 χ^2 -toets, 183
 correlatiecoëfficiënt, 523-524
 schatting, 156
 sequente toetsing, 149-150
 schatting van gemiddelde, 395
 schatting van verschil tussen twee
 gemiddelden, 422
 toetsing, 127, 132-133
 toetsing betr. gemiddelde, 385-390
 toetsing betr. twee gemiddelden,
 412
 Steekproeven
 aselechte, 6, 9-11, 19
 doelbewuste, 13
 gelaagde = laagsgewijze, 585

REGISTER

- lukrake, 14
 met teruglegging, 5, 9, 78, 82
 representatieve, 13
 sequente, 144, 613
 soorten, 170-171
 systematische, 12
 zonder teruglegging, 5, 10, 79, 88
 Stereogram, 464-465
 Stimulus, 588
 Stochastische
 grootheid, 67
 onafhankelijkheid, 69, 71, 171, 211-
 213, 249, 254, 260, 346, 466,
 479, 484, 485, 531
 Stragglers, 432
 STUDENT, toetsen van (*t*-toetsen)
 één gemiddelde, 383-389
 gemiddeld verschil van gepaarde
 waarnemingen, 396-398, 449, 542
 twee gemiddelden, 411-420, 439-
 440, 539, 620
 STUDENT-verdelingen (*t*-verdelingen)
 definitie en toelichting, 347
 fractielen, 348
 gemiddelde en variantie, 347
 relatie met *F*, *T* en χ^2 , 352-353
 transformatie tot *T*, 561
 Sygmoïde, 605
 Symmetrietoets van WILCOXON (zie
 rangnummertoeetsen)
 Symmetrische
 2*k*-punts proef, 594-600
 (2*k* + 1)-punts proef, 600-604
 Systematische
 fout, 400
 steekproef, 12-13

T-toetsen
 één gemiddelde, 382
 twee gemiddelden, 411-413
T-verdeling (zie standaardnormale ver-
 deling)
t-toetsen (zie STUDENT-toetsen)
t-verdelingen (zie STUDENT)
 TCHEBYCHEFF, ongelijkheid van, 118
 Tegenhypothese (zie alternatieve hy-
 pothese)
 Tekens, 319
 Tekentoets, 188-190
 Telkamer, 93, 94, 404-406
 Tellingen, 403-407
 Tendentie, centrale, 34
 TERPSTRA, toets tegen verloop voor *k*
 aselechte steekproeven, 278-283

 Testpreparaat, 589-590
 Therapeutische experimenten, 18-20
 THOMPSON, ijkingsmethode van, 611-
 613
 Toets, 121
 Toetsen
 en schatten, 165-166, 371, 548
 klassieke, 338-545, 568-585, 618-
 624, 629-633
 klassieke of verdelingsvrije, 546-552
 normaliteits, 553-558
 verdelingsvrije, 133-152, 168, 173-
 290, 624-629
 Toetsing, 121, 144, 152-153, 563-564
 één- of tweezijdige, 135
 Toetsings
 grootheid, 121
 theorie, 7, 120-132
 Trappenlijnen, 157-158
 Transformaties
 bgsin, 559, 606
 χ^2 tot *T*, 179, 344, 561
F tot *T*, 561
 logaritmische, 117, 510, 546, 559,
 590-591, 594, 604
 probit-, 605
 reciproke, 559
t tot *T*, 561
 vierkantswortel, 117, 419, 559
z, van correlatiecoëfficiënt, 524-528
 Trekking
 met of zonder teruglegging, 5 (zie
 ook steekproeven)
 Trend (zie verloop)
 TUKBY, methode van (zie splitsing van
 gemiddelden)
 Turftabel, 33
 2 × 2 tabel, 193-202, 213, 286, 624
 2 × *r* tabel, 204-207, 618-621
 Tweedeling (zie dichotomie)
 Tweedimensionaal-normale correlatie,
 514-536
 correlatiecoëfficiënt (zie aldaar)
 grafieken, 515-517
 model, 514-515
 puntenwolk, 518-519
 Tweedimensionale (simultane) verdel-
 ingen, 464-465, 482-485
 Tweetoppigheid, 116, 552
 Tweezijdige
 overschrijdingskans, 77, 139, 141,
 196
 toetsing, 135-136
 Tijdreeks, 261, 533

BIJLAGE V

- Uitbijter, 37, 408, 449, 550, 552, 558, 585
 -toetsen, 409-410
 Uitkomsttypen, 97, 174
 Uitspraak (zie conclusie)
 Unimodaal (zie ééntoppig)
- v.g.*, vrijheidsgraden
 Variabele (grootheid), 2, 3
 Variabiliteit, 2
 Variantie
 berekening, 54-62, via hulpgroot-
 heid, 63-64
 betrouwbaarheidsgrenzen, 367-371
 kansverdeling, 343-345
 residuele (rest-), 443, 448, 581, 583,
 585
 schatting, 355-356
 toetsen, 365-367
 van een kansverdeling, 68, 71
 vrijheidsgraden, 344
 Variantie-analyse
 algebraïsche relaties van kwadraten-
 sommen, 426, 442-444, 453, 503-
 505
 bij factoriële schema's, 568-585
 bij toets op rechtlijnigheid, 503-507
 bij standaardisatie, 595
 enkelvoudige, 425-438
 interactie (zie aldaar)
 kritieke zone, 427
 splitsing van gemiddelden in groe-
 pen, 431-435
 tweevoudige, 440-459
 vermenging, 450
 via spreidingsbreedte, 431
 Varianties, toetsen betreffende
 één variantie, 365-367
 twee varianties, *F*-toets, 371-373
 h varianties, 373-378
 Variatie-coëfficiënt, 559
 Veranderlijke (zie variable)
 Verdeling (zie frequentie-, kans-, simul-
 tane, tweedimensionale, binomiale,
 χ^2 -, *F*-, van gemiddelde, hyper-
 geometrische, multinomiale, norma-
 le, Poisson-, van spreiding, stan-
 daardnormale, STUDENT-, van va-
 riantie)
 Verdelfingsfout, 404, 406
 Verdelfingsvrije methoden, 7, 168-172,
 546-552
 Verdunningsfactor, 404
- Vergelijking (zie ook regressie-)
 van rechte, 329-330
 van de graad *h*, 477
 Verloop, in één waarnemingsreeks
 cyclisch, dalend, oscillatie (zaag-
 tandpatroon), sprongsgewijs, stij-
 gend, 262-276, 378-380
 Verloop, in
 h aselechte steekproeven, 278-283,
 466 e.v.
 h 2 × 2 tabellen, 286-290
 h verwante steekproeven, 284-286
 Vermenging, 450
 Verwachte frequenties, 87, 173, 175-
 176, 197, 216
 lage, 181, 207, 212, 226, 553
 Verwachting (mathematische), 68
 Verwante
 steekproeven, waarnemingen, 170-
 171
 Verwerpen, 127
 Verzameling, 1
 Vierkant (zie Latijns, Grieks-Latijns)
 Vierkantswortel transformatie (zie
 transformatie)
 Vierveldentabel (zie 2 × 2 tabel)
 Vlakheid, 114, 342
 Vooronderzoek, 18
 Voortschrijdend gemiddelde, 611
 Vraagstelling, 16
 Vragen (bij enquête), 21
 Vrijheidsgraden
 begrip, 180, 344
 bij χ^2 -verdeling, 178-179
 bij *F*-verdeling, 349-351
 bij STUDENT(*t*)-verdeling, 347-348
- Waarde, 2, 33
 absolute, 58
 klasse, 3, 34-36
 kritieke, 123
 Waarnemen, 3, 20, 325-326
 Waarnemingen, 3, 6, 21
 fout bij, 400, 408
 nauwkeurigheid, 399-407
 onafhankelijke, 6
 ontbrekende, 25-27
 verwante, 170-171
 Waarnemingsparen, 171, 464
 Waarnemingsuitkomsten (zie waar-
 nemingen)
 Waarschijnlijkheid (zie kans)
 Waarschijnlijkheidspapier (zie nor-
 maal)

REGISTER

- | | |
|--|--|
| <p>WELCH, toets voor twee gemiddelden, 417</p> <p>Werkingskromme, 136-139, 153, 385-386</p> <p>WILCOXON, symmetrietoets van (zie rangnummertoetsen) twee-steekproeven toets, 237-243, 269, 540, 547, 620</p> <p>WILSON-HILFERTY, transformatie van χ^2 tot T, 344-345</p> <p>Worteltrekken, 322</p> <p>Wijzer, 324</p> <p>X^2, maat voor discrepantie tussen waargenomen en verwachte frequenties, 175-176</p> | <p>χ^2, zie chi-kwadraat</p> <p>ξ (xi), zie rij-component</p> <p>YATES, toets betr. $k \times r$ tabel, 618-624</p> <p>IJking, 400
biologische (zie standaardisatie)</p> <p>Z, kritieke zone,
z, (zie transformaties)</p> <p>Zaagtandpatroon, 262</p> <p>Zone, kritieke, 122-132</p> <p>Zuiver, 67, 68</p> <p>Zuiverheid, van
schatting, 154-155
waarnemingsmethode, 400</p> |
|--|--|