

MATHEMATISCH CENTRUM

2e BOERHAAVESTRAAT 49

AMSTERDAM

STATISTISCHE AFDELING

Leiding: Prof. Dr D. van Dantzig

Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

Rapport S 154

Onderzoek naar de gezondheidstoestand van
de bevolking in de Bommeler- en Tielerwaard

door

Hilda A. Kuipers

en

Ann M.J.A. Verbeek

1954

1. Inleiding.

Teneinde een inzicht te krijgen in de gezondheidstoestand van de bevolking der Bommeler- en Tielerwaard, werden ons gegevens verstrekt, voor ieder der jaren 1948 t/m 1952 afzonderlijk, omtrent:

- a. de totale bevolking
- b. het aantal sterfgevallen
- c. het aantal levend geboren kinderen
- d. het aantal doodgeboren kinderen
- e. het aantal overleden zuigelingen (beneden de leeftijd van 1 jaar)
- f. het aantal kleuters (leeftijd 1.- 4 jaar)
- g. het aantal overleden kleuters (leeftijd 1.- 4 jaar)
- h. het aantal aangegeven gevallen van besmettelijke ziekten
- j. het aantal sterfgevallen t.g.v. besmettelijke ziekten
- k. het aantal sterfgevallen t.g.v. tuberculose.

Tevens werden de overeenkomstige gegevens vermeld betreffende de bevolking van de provincie Gelderland en betreffende de bevolking van geheel Nederland.

2. Doel van het onderzoek.

Aan de hand van deze gegevens hebben wij getracht de vraag te beantwoorden of, en in welk opzicht, de gezondheidstoestand van de bevolking in de Bommeler- resp. Tielerwaard verschilt van die in de rest van Gelderland en in de rest van Nederland.

Onder de "rest van Gelderland" wordt hier verstaan de provincie Gelderland met aftrek van het daarmee vergeleken district; evenzo onder de "rest van Nederland" het gehele land behalve het daarmee vergeleken district. Het is duidelijk, dat de hier gekozen interpretatie slechts één van de mogelijkheden is; men zou b.v. voor de "rest van Gelderland" kunnen nemen de provincie Gelderland met aftrek van zowel de Bommeler- als de Tielerwaard en voor de "rest van Nederland" b.v. geheel Nederland behalve de provincie Gelderland. Gezien het vele rekenwerk konden wij slechts één van de mogelijkheden nader uitwerken.

3. Algemene opmerkingen.

a. Om na te gaan of uit het waarnemingsmateriaal tot bepaalde systematische effecten kan worden besloten, maakt men gebruik van wiskundig-statistische methoden, "toetsen" genaamd. Een beschrijving van de algemene gang van zaken bij het toepassen van deze toetsen vindt men in het memorandum S 47 (M 6), hetwelk aan dit verslag is toegevoegd. Voor de onbetrouwbaarheidsdrempel, daar vermeld, werd in ons onderzoek de waarde 0,05 genomen. Daar een star aanhouden van één vaste onbetrouwbaarheidsdrempel bij een onderzoek, waarbij veel toetsingen uitgevoerd worden, een onvolledig inzicht in de resultaten geeft, hebben wij bovendien steeds de (tweezijdige) overschrijdingskansen vermeld. In feite worden de conclusies meer op de waarde van deze overschrijdingskansen gebaseerd dan op het feit alleen of zij groter of kleiner dan 0,05 zijn.

b. De enige toets, die bij dit onderzoek is toegepast, is de methode der dubbele dichotomie; zie memorandum S 47 (M 15). De in dit memorandum beschreven normale benaderingsmethode kan in alle gevallen toegepast worden, ook - zoals bij een nader onderzoek met nauwkeuriger maar bewerkelijker methoden, bleek - in die gevallen, waar kleine aantallen voorkwamen. Bij de methode der dubbele dichotomie kunnen alleen verhoudingscijfers vergeleken worden, b.v. de verhouding tussen het sterftecijfer en de totale bevolking in de Bommelerwaard met die in de rest van Nederland. Bij elk verschijnsel dat onderzocht is, zal vermeld worden op welk totaalcijfer de verhoudingen betrokken zijn.

c. De toetsing met behulp van de methode der dubbele dichotomie werd telkens toegepast op de gegevens van één jaar. De gebruikte demografische cijfers hebben nl. in de opeenvolgende jaren voor een groot deel betrekking op dezelfde personen. Deze cijfers kunnen daarom niet als onderling onafhankelijke waarnemingen worden beschouwd, waardoor het niet mogelijk was de resultaten der opeenvolgende jaren tot één toets te combineren. Tevens moet bij het bestuderen van de resultaten rekening worden gehouden met het feit dat ook de toetsen, toegepast bij verschillende onderzoekingen in hetzelfde jaar, niet geheel onafhankelijk zijn.

d. Een moeilijkheid bij dit onderzoek was, dat een aantal factoren, die de gezondheidstoestand beïnvloeden, onbekend waren, zoals b.v. de leeftijdsverdeling en de inkomstenverdeling van de beschouwde groepen personen. Daar over deze factoren geen gegevens beschikbaar waren, konden zij niet geëlimineerd worden. Wordt er een verschil gevonden, dan kunnen één of meer van deze factoren geheel of gedeeltelijk de oorzaak daarvan zijn. In hoeverre men sommige van deze factoren in de nogal vage term "gezondheidstoestand" kan opnemen, laten wij hier buiten beschouwing. Dat is een kwestie van medische interpretatie, waar de statisticus zich buiten kan houden. Bij de interpretatie van de conclusies dient echter met het bovenstaande wel rekening gehouden te worden.

Bovendien zijn deze conclusies gebaseerd op de waarnemingen, verricht over een betrekkelijk korte periode (5 jaar); de conclusies hebben dus slechts waarde, voor zover deze periode als voldoende representatief beschouwd kan worden.

4. Statistisch onderzoek ¹⁾

Wij hebben in de eerste kolom van de tabel der resultaten (tabel I) aangegeven ~~waarop~~ de schillende verhoudingscijfers betrokken zijn bij de toepassing van de methode der dubbele dichotomie. Waar nodig zullen wij onze keuze in deze paragraaf toelichten; verder bevat deze paragraaf enkele opmerkingen over punten, die bij het statistisch onderzoek tot moeilijkheden aanleiding gaven.

4.1. Bevolkingstoename.

De toename van de bevolking in een bepaalde streek houdt slechts ten dele verband met de daar bestaande gezondheidstoestand, zich afspiegelend in het aantal geboorten en sterfgevallen. Andere factoren (economische enz., die van invloed zijn op de migratie) kunnen hierbij ook een belangrijke rol spelen.

Daar zowel het aantal geboorten (zie 4.2) als het aantal sterfgevallen nog afzonderlijk ter sprake komen, leek ons een vergelijking van de mutaties in de totale bevolking der onderzochte districten met die in de provincie Gelderland resp. in geheel Nederland weinig zinvol. Dit werd dan ook achterwege gelaten.

1) Voor de resultaten van de in deze par. beschreven vergelijkingen zie de paragrafen 5 en 6.

4.2. Geboorten.

Het zou wellicht voor de hand liggen het aantal geboorten te betrekken op het aantal gehuwde vrouwen met inachtneming van de leeftijd. Wij beschikten echter niet over de daarvoor benodigde gegevens en hebben daarom het aantal geboorten betrekken op de totale bevolking.

Bij de beoordeling van de resultaten betreffende het totale aantal geboorten, het aantal levend geboren en het aantal dood geboren, dient men er rekening mee te houden, dat deze drie cijfers zeer nauw samenhangen en dus allerm minst als onafhankelijk beschouwd kunnen worden.

4.3. Zuigelingensterfte.

De zuigelingensterfte levert een vrij zuiver vergelijkingsmateriaal op, omdat hierin de gezondheidstoestand (zuigelingenzorg, hygiënische toestand e.d.) wel rechtstreeks weerspiegeld wordt, al spelen ook hier andere factoren (b.v. van praeinatale aard) een rol. Wij betrokken in dit geval het aantal kinderen, overleden beneden de leeftijd van 1 jaar, op het aantal levend geboren.

4.4. Besmettelijke ziekten.

Bij het onderzoek van het aantal aangiften van besmettelijke ziekten doet zich een moeilijkheid voor, die het toepassen van de hier gebruikte statistische toetsingsmethode onmogelijk maakt. Deze methode berust nl. op de onderstelling van onafhankelijkheid, d.w.z. op de onderstelling, dat er voor iedere persoon in ieder der twee te vergelijken streken een bepaalde kans op een bepaald verschijnsel (b.v. overlijden in een bepaald jaar) bestaat, maar dat het al of niet optreden daarvan bij de ene persoon het al of niet optreden bij de andere persoon niet beïnvloedt. Dit is bij besmettelijke ziekten echter - juist wegens hun besmettelijkheid - niet het geval en daardoor verliest de gebruikte methode zijn geldigheid. Het aantal gevallen van besmettelijke ziekten is dan ook een onvoldoende gegeven voor het vergelijken van dit aspect van de gezondheidstoestand van twee landstreken. Andere gegevens, zoals het aantal epidemieën en hun uitgebreidheid en verspreiding, kunnen bij een dergelijk onderzoek niet gemist worden. Daar deze gegevens niet aanwezig zijn, kan op dit punt geen analyse uitgevoerd worden.

4.5. Sterfgevallen ten gevolge van besmettelijke ziekten.

Indien men het aantal sterfgevallen ten gevolge van besmettelijke ziekten zou betrekken op de totale bevolking der te vergelijken streken, dan zou zich dezelfde moeilijkheid voordoen als in par. 4.4. voor de aantallen besmettelijke ziekten zelf zijn gesignaleerd. Dit kan dus niet gedaan worden. Wel kan men met iets minder bezwaar het aantal sterfgevallen betrekken op het aantal gevallen van besmettelijke ziekten, hoewel ook hier het eventuele voorkomen van een kwaadaardige epidemie in één van de beide vergeleken streken, waarbij - wegens het kwaadaardige karakter - de sterftekans groter is dan bij andere epidemieën, tot de zelfde moeilijkheden zou kunnen leiden en zodoende tot een schijnresultaat zou kunnen voeren. Daar ons bovendien het aantal gevallen van besmettelijke ziekten niet bekend is, maar alleen het aantal aangegeven gevallen daarvan, moeten wij de toets daarop betrekken. Een verschil in gewoonte, wat het al of niet aangeven van besmettelijke ziekten betreft, dat tussen de te vergelijken streken kan bestaan, kan ons dus bij dit onderzoek parten spelen. Alleen onder veel voorbehoud worden daarom de resultaten van dit deel van het onderzoek in dit rapport opgenomen.

4.6. Sterfgevallen tengevolge van tuberculose.

Bij het onderzoek van par. 4.5. werd de tuberculose buiten beschouwing gelaten. Deze wordt in deze paragraaf apart onderzocht. Hoewel ook de tuberculose een besmettelijke ziekte is, zodat de in par. 4.4. en 4.5. genoemde moeilijkheden zich ook hier voordoen, kan voor deze ziekte toch met meer recht van deze moeilijkheden afgezien worden. Tuberculose is nl. een langzaam werkende ziekte, die minder in korte felle epidemieën optreedt dan andere besmettelijke ziekten. De mate van voorkomen van tuberculose in een bepaalde streek schommelt dan ook veel minder sterk dan die van andere besmettelijke ziekten en kan daarom gezien worden als een facet van de gezondheidstoestand van de streek. Op grond hiervan valt de toepassing van de methode der dubbele dichotomie wel te rechtvaardigen.

Daar het totale aantal gevallen van tuberculose niet gegeven was, kan de vergelijking van het aantal sterfgevallen daar niet op betrokken worden. Wij hebben dit daarom gedaan op de totale bevolking der te vergelijken streken en ook op het totale aantal sterfgevallen. Dit laatste betekent dus,

dat voor de te vergelijken streken onderzocht wordt of tuberculose als doodsoorzaak al of niet even veelvuldig voorkomt. Daarbij moet rekening gehouden worden met de mogelijkheid, dat een grote sterfte ten gevolge van andere besmettelijke ziekten in een bepaald jaar in een bepaalde streek de geldigheid der conclusies kan aantasten.

5. Resultaten.

De resultaten van het in par. 4 beschreven onderzoek zijn in tabel I op de volgende bladzijde samengevat. Deze tabel geeft de waarden van de bij alle vergelijkingen gevonden tweezijdige overschrijdingskansen. Een +, resp. - teken achter een overschrijdingskans betekent, dat de genoemde grootte in het gebied, aangegeven door de eerste letter groter, resp. kleiner is dan in het daarmee vergeleken gebied. Is het vakje hierachter gearceerd, dan is het gevonden verschil duidelijk (overschrijdingskans $k \leq 0,01$); een kruis in dit vakje geeft een minder duidelijke aanwijzing voor een verschil aan ($0,01 \leq k \leq 0,05$); bevat het vakje één streepje, dan is er slechts een zeer geringe aanwijzing voor een verschil in een bepaalde richting ($0,05 \leq k \leq 0,10$); is het bewuste vakje ledig, dan is er geen noemenswaardige aanwijzing voor een systematisch verschil.

6. Samenvatting der conclusies.

Bij de beschouwing van tabel I kunnen wij de volgende conclusies trekken:

Voor de Bommelerwaard:

1. Het aantal sterfgevallen per hoofd van de bevolking was in 1950 zeer duidelijk hoger dan in de rest van Gelderland en in de rest van Nederland. In 1948 lag het ook hoger dan in de rest van Nederland, terwijl er een geringe aanwijzing in dezelfde richting bestond ten aanzien van de rest van Gelderland. De resultaten van de overige jaren vertoonden geen afwijkingen, die voldoende duidelijk zijn om ze als systematisch te kenmerken.
2. De zuigelingensterfte betrokken op het aantal levend geboren was in 1950 en 1951 duidelijk hoger dan in de rest van Gelderland resp. Nederland. In de andere jaren vertoonde deze sterfte geen afwijkingen van betekenis.

Tabel I

Overzicht van de tweezijdige overschrijdingskansen

Jaar		1948	1949	1950	1951	1952
Sterfte (op totale bevolking)	BG ²⁾	0,09+	0,19 +	<10 ⁻⁶ +	0,17-	0,50+
	BN	0,03+	0,11+	<10 ⁻⁶ +	0,53-	0,19+
	TG	0,01+	<10 ⁻⁴ +	0,17+	0,11+	0,01+
	TN	0,001+	<10 ⁻⁴ +	0,13+	0,001+	5x10 ⁻⁴ +
Geboorten (op totale bevolking)	BG	0,44+	0,84 -	0,56+	0,79+	0,94-
	BN	0,11+	0,45 +	0,11+	0,12+	0,14+
	TG	0,33-	0,03 -	0,12-	0,03-	0,01-
	TN	0,77+	0,47 -	0,98+	0,81-	0,77-
Levend geborenen (op totale bevolking)	BG	0,38+	0,67 -	0,58+	0,86+	0,99+
	BN	0,11+	0,68 +	0,13+	0,16+	0,13+
	TG	0,28-	0,03 -	0,14-	0,03-	0,01-
	TN	0,95+	0,37 -	0,99-	0,76-	0,83-
Dood ge- borenen (op aantal geboorten)	BG	0,48+	0,13 +	0,91+	0,59+	0,52-
	BN	0,84-	0,02 +	0,58+	0,44+	0,65-
	TG	0,44+	0,74 +	0,72-	0,83+	0,42-
	TN	0,10+	0,16 +	0,79+	0,66+	0,60-
Zuigelin- gensterfte (op aantal levend ge- borenen)	BG	0,76+	0,49 -	0,003+	0,03+	0,36-
	BN	0,54+	0,77 -	2x10 ⁻⁴ +	0,005+	0,59-
	TG	0,23-	0,99 +	0,90-	0,11-	0,28+
	TN	0,46-	0,52 +	0,45+	0,35-	0,09+
Kleuter- sterfte (op aantal kleuters)	BG	0,99-	0,09 +	0,38+	0,78+	0,38-
	BN	0,70-	0,10 +	0,33+	0,70+	0,37-
	TG	0,91+	0,45 +	0,96-	0,87-	0,40+
	TN	0,65-	0,48 +	0,96-	0,98-	0,45+
Sterfte aan besmet- telijke ziekten (op aantal aangiften hiervan)	BG	0,49+	0,10 +	0,37-	0,18-	0,23+
	BN	0,20+	0,02 +	0,43-	0,30-	0,27+
	TG	0,04+	0,90 -	0,26-	0,95+	0,32+
	TN	0,001+	0,51 +	0,40-	0,45+	0,37+
Sterfte aan tuber- culose (op totale sterfte)	BG	0,16+	0,17 -	0,06-	0,46-	0,03+
	BN	0,15+	0,28 -	0,14-	0,61-	0,02+
	TG	0,06-	0,15 -	0,97-	0,98-	0,14-
	TN	0,08-	0,35 -	0,56+	0,67+	0,22-
Sterfte aan tuber- culose (op totale be- volking)	BG	0,07+	0,24 -	0,43-	0,36-	0,02+
	BN	0,05+	0,40 -	0,70-	0,56-	0,008+
	TG	0,16-	0,43 -	0,85+	0,82+	0,23-
	TN	0,22-	0,81 -	0,35+	0,39+	0,39-

1) Leeftijd 1 - 4 jaar.

2) De volgende afkortingen worden gebruikt:

BG = vergelijking van de Bommelerwaard met de rest van Geld.

BN = " " " " " " " " Nederl.

TG = " " " Tielerwaard " " " " Geld.

TN = " " " " " " " " Nederl.

3. De sterfte aan tuberculose betrokken op de totale bevolking was in 1952 hoger dan in de rest van Gelderland resp. Nederland; een aanwijzing in dezelfde richting vonden wij voor 1948. In de tussen liggende jaren was de TBC-sterfte lager, zonder dat hiérbij een systematisch verschil aangetoond kan worden. Als deze sterfte betrokken wordt op de totale sterfte, is het effect van 1952 minder duidelijk, terwijl dan van een systematisch verschil in 1948 niets meer blijkt.

4. Bij de andere punten van onderzoek werden slechts geïsoleerde resultaten gevonden; het jaar 1949 levert een aanwijzing, dat het aantal doodgeborenen, betrokken op het totale aantal geboorten, groter was dan in de rest van Nederland, terwijl hetzelfde het geval was voor de sterfte aan besmettelijke ziekten. Overigens werden geen afwijkingen van betekenis gevonden, hetgeen deze aanwijzingen aanzienlijk verzwakt.

Over het algemeen kan men wel zeggen, dat alle resultaten, die enigszins duidelijk zijn, wijzen op een slechtere gezondheidstoestand in de Bommelerwaard dan elders in Gelderland of Nederland. In ieder der jaren zijn op één of meer punten aanwijzingen (waaronder sterke) voor afwijkingen in de ongunstige richting, terwijl er slechts één zeer zwakke aanwijzing in de tegenovergestelde richting is. (nl. voor de sterfte aan tuberculose in 1950).

Voor de Tielerwaard.

1. Het aantal sterfgevallen per hoofd van de bevolking was over alle jaren van onderzoek duidelijk hoger dan in de rest van Nederland; ~~het was in de jaren 1948, 1949 en 1952 eveneens duidelijk hoger dan in de rest van Gelderland~~ en in de andere jaren lagen de resultaten in dezelfde richting.

2. Het aantal geboorten betrokken op de totale bevolking was in 1949, 1951 en 1952 lager dan elders in Gelderland. Hetzelfde geldt voor het aantal levend geboren. In de andere jaren lagen de resultaten in dezelfde richting. Vergelijken wij deze geboortecijfers met die van de rest van Nederland, dan zijn er echter geen afwijkingen van betekenis te constateren.

3. De sterfte aan besmettelijke ziekten betrokken op het aantal aangegeven gevallen, was in 1948 hoger dan elders in Gelderland resp. Nederland. De andere jaren vertonen geen afwijkingen van betekenis.

4. Op geen enkel ander punt werden afwijkingen gevonden van voldoende betekenis.

Hoewel in de Tielerwaard de sterfte over de gehele linie ongunstig is, zijn er hier aanzienlijk minder **aanwijzingen** voor een slechte gezondheidstoestand op andere punten. Dit maant tot voorzichtigheid bij het trekken van een conclusie omtrent de gezondheidstoestand, daar b.v. ook een verschil in leeftijdsverdeling ~~een~~ verschil in sterfte en geboorte tot gevolg kan hebben.

Algemene gang van zaken bij het toetsen van een ¹⁾
hypothese.

De toetsing van een hypothese H_0 berust steeds op een aantal waarnemingen $x_1, x_2, \dots, x_n$ van één of meer stochastische grootheden²⁾, of op enige groepen van waarnemingen (bv. twee steekproeven).

Bij een toets behoort een toetsingsgrootheid u (soms meer dan één), die een functie is van bovengenoemde stochastische grootheden en die, voor de waargenomen waarden $x_1, x_2, \dots, x_n$ een waarde aanneemt, die berekend kan worden (bv.: het gemiddelde der waarnemingen, of de spreiding, of het verschil van de gemiddelden van twee waarnemingen).

De toetsingsgrootheid wordt steeds zo gekozen, dat men, op grond van de onderstelling, dat H_0 juist is, de waarschijnlijkheidsverdeling van deze grootheid kan berekenen.

Vervolgens kiest men een verzameling Z van mogelijke uitkomsten van u , en wel op zodanige wijze, dat de kans, dat u een in Z gelegen waarde aanneemt, onder de hypothese H_0 , gelijk is aan een gegeven getal α , zodat Z dus van α afhankelijk is. Z heet de kritieke zône van de toets, α de onbetrouwbaarheidsdrempel (Engels: level of significance). Voor α neemt men veelal de waarde 0,05 of 0,01.

Men verworpt nu H_0 op grond van de waarnemingen $x_1, x_2, \dots, x_n$, indien de bij deze waarnemingen behorende waarde van u in Z ligt. Dit wordt vaak uitgedrukt door te zeggen, dat het resultaat van het experiment "significant" is. De waarde van α moet dan echter worden vermeld. De kans, dat dit zal gebeuren, is, indien H_0 juist is, gelijk aan α . Derhalve is α de kans op ten onrechte verwerping van de juiste hypothese, ook de kans op een fout van de eerste soort genoemd. Indien men deze methode toepast, met $\alpha = 0,05$ resp. 0,01, zal men in gemiddeld ongeveer één op 20 resp. op 100 van de gevallen, waarin de hypothese die men toetst juist is, deze toch verwerpen.

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

2) Een stochastische grootheid is een grootheid, die een waarschijnlijkheidsverdeling bezit, of, anders gezegd, een grootheid, die voor de elementen van een collectie(universum, populatie) gedefinieerd is en daarop allerlei waarden aanneemt. Stochastische grootheden worden aangegeven door onderstreepte letters.

3) Soms kan men slechts bereiken, dat deze kans $\leq \alpha$ is.

De toetsingstheorie biedt in het algemeen geen mogelijkheid om tot aanvaarding van een hypothese te komen. Indien een bepaalde hypothese H_0 niet verworpen kan worden, is dit gewoonlijk met een hele verzameling van hypothesen tegelijk het geval. Niet-verwerpen staat dus niet gelijk met aanvaarden.

Wel zal men vaak in de loop van een statistische analyse bepaalde onderstellingen, die plausibel schijnen en voor de verdere analyse van nut zijn, toetsen, alvorens ze bij de verdere bewerking van het materiaal te gebruiken. Worden zij dan op grond van de toets niet verworpen, dan houdt dit in zo verre een rechtvaardiging van die onderstellingen in, dat een grote afwijking door de toets veelal wel zou zijn ontdekt. Indien men dan verder de onderstellingen gebruikt, verwaarloost men eventueel aanwezige afwijkingen van onbekende grootte, die echter niet zo groot zijn, dat zij door de toets zijn ontdekt.

Vele toetsen gelden zelf alleen onder bepaalde onderstellingen omtrent de waarschijnlijkheidsverdelingen der stochastische grootheden, waarvan waarnemingen zijn verricht. Deze nevenvoorwaarden dienen steeds uitdrukkelijk te worden vermeld en, zo mogelijk, zelf te worden getoetst.

In plaats van de onbetrouwbaarheidsdrempel α wordt vaak bij de uitslag van een toetsing de overschrijdingskans k opgegeven; dit is de kleinste waarde van α , waarbij in het betrokken geval, nog tot verworping van H_0 , zou zijn overgegaan; anders gezegd: de kleinste α , waarvoor de gevonden waarde der toetsingsgrootte nog juist in de (bij α behorende) kritieke zone Z ligt. Wordt dus de waarde k opgegeven en werkt men met onbetrouwbaarheidsdrempel α , dan wordt verworpen, indien $k \leq \alpha$ is.

Voor het onderscheid tussen één- en tweezijdige toetsing en de keuze tussen deze twee mogelijkheden vergelijk men bv. de tweede hieronder gegeven literatuurplaats. Wij moeten hier volstaan met de opmerking, dat éénzijdige toetsing veelal eerder tot verworping van H_0 leidt, maar dat deze slechts onder bijzondere omstandigheden kan worden toegepast.

Litteratuur:

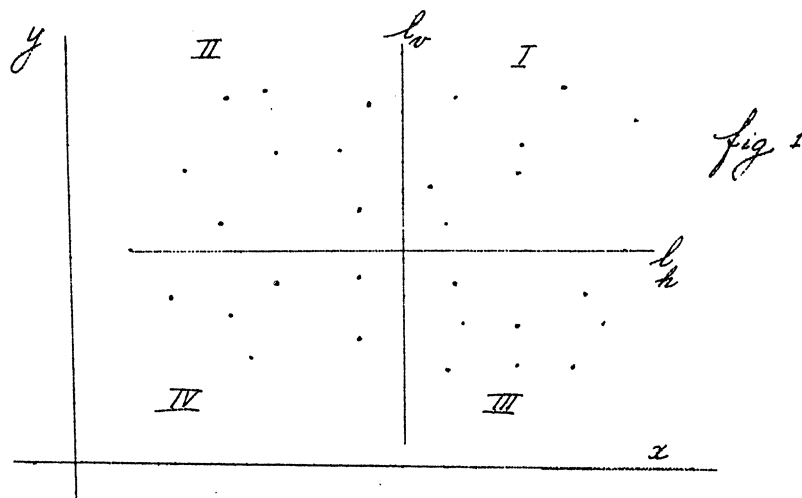
J.Neyman, First course in probability and statistics, New York, 1950, Chapter 5.

J.Hemelrijk en H.R. van der Vaart, Het gebruik van één- en tweezijdige overschrijdingskansen voor het toetsen van hypothesen, Statistica 4 (1950) p.54-66.

S 47 (M 15).

Toets voor onafhankelijkheid van 2 grootheden
met behulp van de methode der
Dubbele Dichotomie ¹⁾

Deze methode dient o.a. om de onafhankelijkheid te toetsen van twee continue grootheden. $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$ zijn waarnemingsparen van de stochastische grootheden x en y . Deze getallenparen worden als punten in een vlak getekend.



We verdelen de puntenwolk door een verticale en een horizontale rechte (l_v resp. l_h) zo, dat links en rechts van l_v , en ook boven en onder l_h evenveel punten liggen. Het vlak is nu verdeeld in vier gebieden. Is er geen afhankelijkheid dan verwachten we dat in alle vier gebieden ongeveer evenveel punten zullen liggen. Een exacte afleiding, (geschikt voor kleine aantallen waarnemingen) verkrijgen we als volgt:

We beschouwen de punten in ons vlak als waarnemingen van elementen die ieder tegelijk twee kenmerken bezitten; A of $\bar{A}$ (rechts resp. links van verticale streep) en B of $\bar{B}$ (boven resp. onder horizontale streep). Er ontstaan dus vier groepen elementen met de kenmerken

AB	$A\bar{B}$	$\bar{A}B$	$\bar{A}\bar{B}$
vak I	vak III	vak II	vak IV

We tekenen dit als volgt

in een 2 x 2 tabel: (Zie blz 2)

¹⁾ Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

^{''}) A betekent: "non-A"

	A	$\bar{A}$	
B	a	$n-a$	n
$\bar{B}$	b	$m-b$	m
	n	s	N

Het totale aantal elementen is N . Het aantal elementen dat kenmerk B bezit, is n , het aantal dat kenmerk $\bar{B}$ bezit, is m . Hetzelfde geldt voor de kenmerken A en $\bar{A}$ met de aantallen n resp. s . In het bovenstaande voorbeeld werd de puntenwolk zo verdeeld, dat $n = m = n = s = \frac{1}{2} N$

Dit principe wordt ^{verp.} toegepast, maar is niet noodzakelijk. De toets kan ook worden gebruikt met andere waarden van m, n, n, s mits deze steeds op grond van de proefopstelling, niet naar aanleiding van de uitkomst van het experiment worden gekozen.

Onafhankelijkheid van de kenmerken wil zeggen, dat de kans dat een element het kenmerk B bezit even groot is, indien het tevens het kenmerk A bezit, als het tevens het kenmerk $\bar{A}$ bezit. Wij toetsen nu deze hypothese van onafhankelijkheid.

Daar de rand-totalen gegeven zijn, is er nog één vrijheidsgraad, d.w.z. indien één van de waarden in één der vier binnenvakjes bepaald is, volgen daaruit de andere waarden. Wij beschouwen het getal a in het linker-boven vakje. Dit kan verschillende waarden aannemen en is dus een stochastische grootheid a , waarvan de waarschijnlijkheidsverdeling, onder aanname van de te toetsen hypothese, een z.g. hypergeometrische verdeling is, die gegeven wordt door

$$P[a = a] = \frac{\binom{n}{a} \binom{m}{n-a}}{\binom{m+n}{n}} = \frac{n! m! n! s!}{a! b! (n-a)! (m-b)! N!}$$

Het gemiddelde van deze verdeling is $E a = \frac{mn}{N}$ en het spreidings kwadraat is $\sigma_a^2 = \frac{n m n s}{N^2 (N-1)}$

Als voorbeeld nemen wij $N=20$, $n=s=m=n=10$

De waarschijnlijkheidsverdeling van a is nu in fig. 2 en tabel I weergegeven.

getallen v.b.

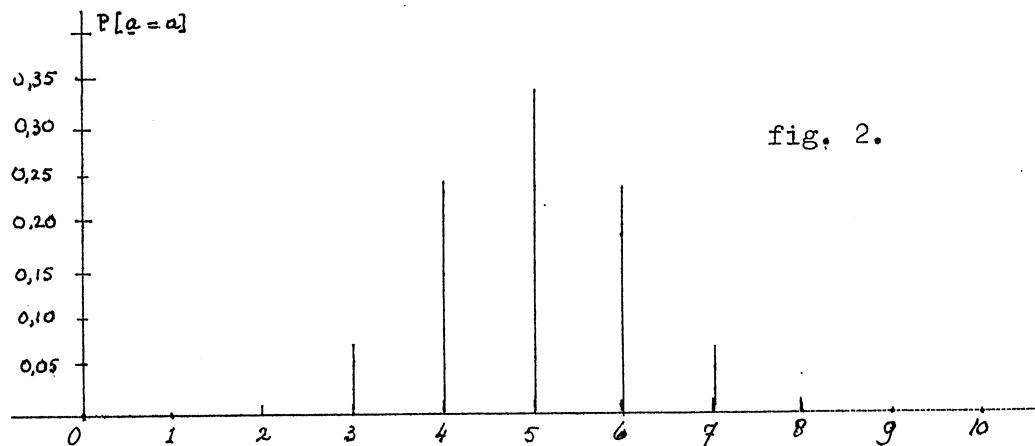


fig. 2.

mogelijkheden:

0 10	1 9	2 8	3 7	4 6	5 5	6 4	7 3	8 2	9 1	10 0
10 0	9 1	8 2	7 3	6 4	5 5	4 6	3 7	2 8	1 9	0 10

Tabel I

a	$P[\underline{a} = a]$
0	0,00001
1	0,00054
2	0,01096
3	0,07794
4	0,24356
5	0,34372
6	0,24356
7	0,07794
8	0,01096
9	0,00054
10	0,00001

Stel nu dat de volgende waarden gevonden zijn:

	A	$\bar{A}$	
B	2	8	10
$\bar{B}$	8	2	10
	10	10	20

We berekenen

$$P[\underline{a} = 2] = \frac{(10!)^4}{20!(8!)^2(2!)^2} = 0,01096$$

en zoeken verder alle waarden van a bijeen, waarvoor $P[\underline{a} = a] \leq P[\underline{a} = 2]$ is

Vervolgens tellen we de daarbij behorende waarschijnlijkheden

$P[\underline{a} = a]$ op. (In ons voorbeeld dus voor $\underline{a} = 0, 1, 2, 8, 9, 10$).

Deze som is per definitie de overschrijdingskans, behorende bij het gevonden resultaat, (in ons voorbeeld 0,02302) ¹⁾

¹⁾ De bij deze definitie van de overschrijdingskans behorende kritieke zône bestaat uit alle waarden voor a met overschrijdingskans $\leq \alpha$, in ons voorbeeld 0,1,2,8,9 en 10.

Is deze α (de onbetrouwbaarheidsdrempel), dan wordt de hypotheese van onafhankelijkheid verworpen (In ons voorbeeld treedt dus, als men $\alpha = 0,05$ neemt, verworping op).

Voor grote aantallen waarnemingen maken we gebruik van het feit dat $\frac{a - E_a}{\sigma_a}$ bij benadering normaal verdeeld is, met gemiddelde 0 en spreiding 1.

Continuïteitscorrectie

Deze behoeft alleen bij kleine aantallen toegepast te worden en bestaat daarin, dat men alle getallen $a, b, n-a, n-b$ met $\frac{1}{2}$ vermindert of vermindert, zodanig dat de randtotalen dezelfde blijven $|a - \bar{a}|$ kleiner wordt.

Literatuur:

- M.G.Kendall, The advanced theory of Statistics, Vol.I,
(London 1947), p.303;
E.S.Pearson, The choice of statistical tests illustrated on
the interpretation of data classed in a 2 x 2
table, Biometrika 34 (1949) p.139-167.