MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
STATISTISCHE AFDELING

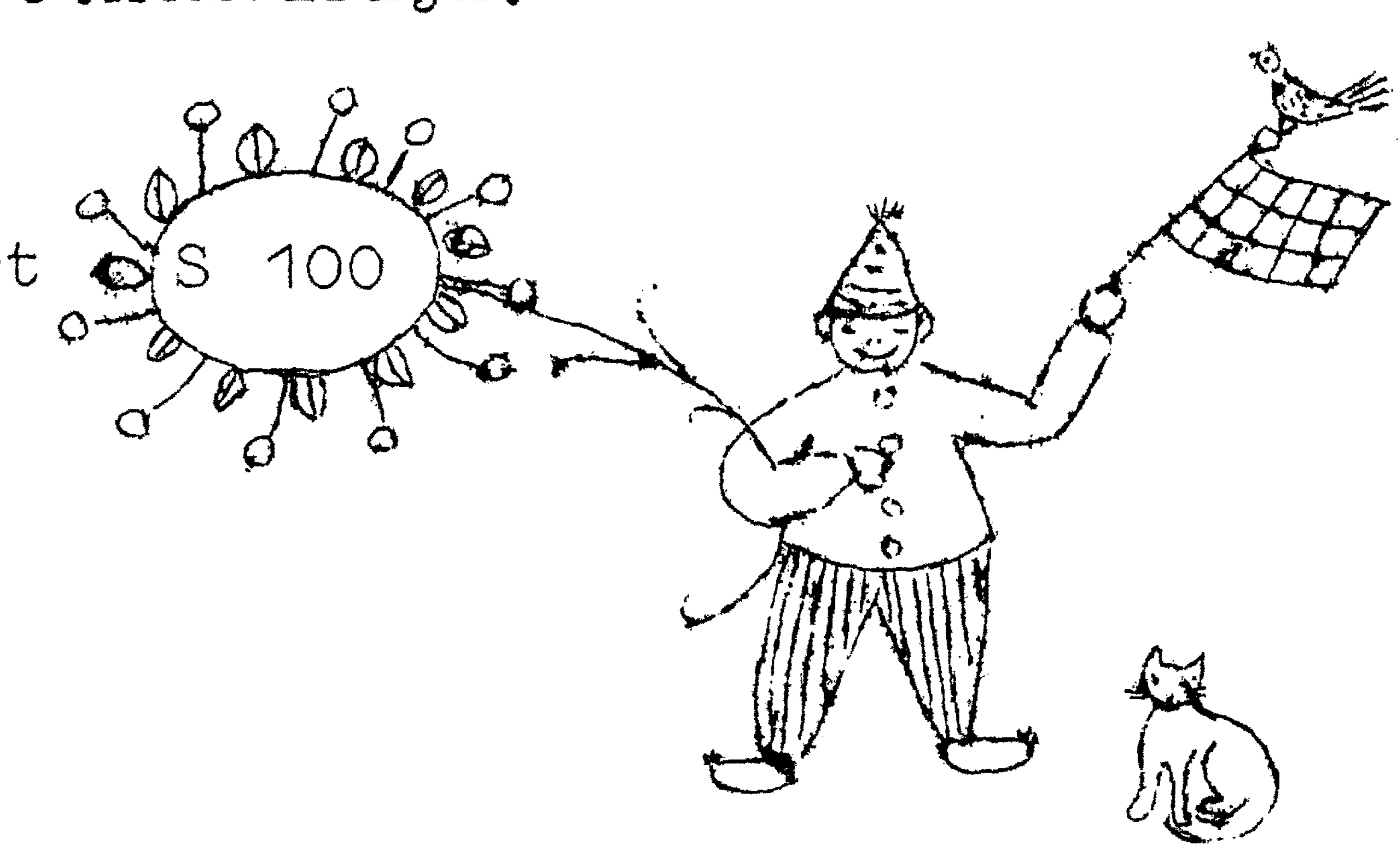
Leiding: Prof. Dr D. van Dantzig
Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

Statistische methoden zonder onderstelling van normaliteit.

door

Prof. Dr J.Hemelrijk.

Rapport S 100



1953.

Statistische methoden zonder onderstelling van normaliteit.

Inleiding. In de vorige hoofdstukken zijn verschillende statistische methoden besproken, die berusten op de onderstelling, dat de waarnemingen uit een normale verdeling afkomstig zijn. Indien aan deze onderstelling niet voldaan is, worden de conclusies minder betrouwbaar dan door de gebruikte onbetrouwbaarheidsdrempel aangegeven wordt. De mate, waarin dit het geval is, is afhankelijk van de mate en van de aard van de afwijking van normaliteit en is vaak niet bekend.

Naast de traditionele, op normaliteit gebaseerde, methoden, beschikt men tegenwoordig over een aantal statistische methoden, waarbij geen onderstellingen over de vorm van de verdeling wordt gemaakt. In dit hoofdstuk worden enkele der eenvoudigste van deze zg. "parametervrije" methoden behandeld.

Voorbeelden.

Niet normale, in het bijzonder scheve, verdelingen treft men op vele gebieden aan, zoals b.v. bij sterkteproeven, bij de verdeling van vezellengten bij sommige textielproducten, de verdeling der inkomens over een bevolking enz. In sommige van deze gevallen is er voldoende bekend over de vorm van de verdeling, om deze door een transformatie tot een normale te maken en daarna de traditionele methoden toe te passen. Vooral als men over veel waarnemingen beschikt, kan men pogingen in die richting doen. Bij zeer onregelmatige verdelingen of als er te weinig waarnemingen beschikbaar zijn om de vorm der verdeling te kunnen onderzoeken, is voorzichtigheid geboden. Dergelijke verdelingen, waarbij men weinig kan zeggen over de vorm, treden vooral op bij medische, biologische en aanverwante onderzoekingen. Wij geven daarom twee aan deze gebieden ontleende voorbeelden.

I. Bij een onderzoek naar de invloed van ergotamine op de zuurgraad van de maaginhoud van patiënten, lijdende aan een maagzweer[1], werd bij 12 patiënten voor en na het toedienen van ergotamine de zuurgraad van de maaginhoud bepaald (de details van het onderzoek, dat veel uitgebreider was, blijven hier buiten beschouwing). Voor ieder der patiënten werd het verschil in zuurgraad na en voor de toediening van het

----

1) Ontleend aan J.F.Visser, Diss., Utrecht 1950, tabel V, groep Ib, 11.45 uur.

ergotamine berekend. Daarbij werden de volgende getallen gevonden (de gebruikte maat doet hier niet ter zake en is daarom niet vermeld).

Tabel I

Verschil in zuurgraad van de maaginhoud na en voor het toedienen van ergotamine bij twaalf patiënten

24; 33; 7; -7; 14; 60; 32; 8; 8; 0; 8; -20.

De bedoeling van het onderzoek was, na te gaan of de zuurgraad door de toediening van het ergotamine wordt beïnvloed.

II. Bij een onderzoek naar het vetgehalte van verschillende groepen duiven[2], uitgedrukt in procenten van het lichaamsgewicht, werden de volgende twee reeksen waarnemingen verkregen.

Tabel II

Vetgehalte bij twee groepen duiven, uitgedrukt in percenten van het lichaamsgewicht

groep a	groep b
1,4	3,0
2,2	3,4
1,7	6,4
1,7	2,9
2,6	2,2
2,5	7,9
7,8	
2,3	

De bedoeling van het onderzoek was in dit geval, na te gaan of er een systematisch verschil tussen de vetgehalten van de twee groepen duiven bestaat.

In beide gevallen wijzen de uitkomsten allerminst op normaliteit. In het bijzonder zijn er in beide gevallen waarnemingen, die zeer sterk van het gemiddelde afwijken. Over de vorm van de verdelingen valt weinig te zeggen. Het betekent dus een aanzienlijke winst, indien deze gegevens geanalyseerd kunnen worden zonder dat daarbij onderstellingen omtrent deze vorm behoeven te worden gemaakt.

----

2) Ontleend aan M.Gruber, Diss., Utrecht 1952, p. 65, groepen IV en V.

De mediaan.

Bij parametervrije methoden beschouwt men vaak de mediaan van een verdeling in plaats van het gemiddelde. De mediaan is die waarde van de grootheid, waar de helft van de op de gehele populatie aangenomen waarden onder en de helft boven ligt. Bij een symmetrische verdeling, zoals b.v. de normale, valt de mediaan dus samen met het gemiddelde. De mediaan van een waarnemingsreeks is de middelste waarneming bij rangschikking naar grootte, of als er een even aantal waarnemingen is, het gemiddelde van de twee "middelste" waarnemingen. De mediaan van een steekproef uit een verdeling is een schatting van de mediaan van de verdeling.

Betrouwbaarheidsinterval voor de mediaan.

Evenals men in het geval van een reeks waarnemingen van een normaal verdeelde grootheid een betrouwbaarheidsinterval voor het gemiddelde kan berekenen, kan men, zonder onderstelling over de vorm van de verdeling, een betrouwbaarheidsinterval voor de mediaan geven.

Daartoe rangschikt men de waarnemingen naar opklimmende grootte. Noemen wij de aldus gerangschikte waarnemingen

$$x_{(1)},\ x_{(2)},\ \ldots,\ x_{(n)},$$

en is α de gekozen onbetrouwbaarheidsdrempel, terwijl a het grootste gehele getal is, waarvoor

$$2^{-n+1} \sum_{i=0}^{a} \binom{n}{i} \leqq \alpha \ ,$$

dan is het interval $[x_{(a+1)}, x_{(n-a)}]$, de grenzen inbegrepen, een tweezijdig betrouwbaarheidsinterval voor de mediaan van de verdeling, met α als onbetrouwbaarheidsdrempel.

Eénzijdige betrouwbaarheidsintervallen voor de mediaan, met onbetrouwbaarheidsdrempel α, verkrijgt men door het grootste gehele getal b op te zoeken, waarvoor

$$2^{-n} \sum_{i=0}^{b} \binom{n}{i} \leqq \alpha \ .$$

De waarneming $x_{(b+1)}$ is dan een ondergrens voor de mediaan, met onbetrouwbaarheidsdrempel α, en $x_{(n-b)}$ is een bovengrens met dezelfde onbetrouwbaarheidsdrempel.
De bij gegeven waarden van n en α behorende waarden van a en b kan men opzoeken in de volgende tabellen.

Tabellen: "Tables of the binomial probability distribution", National Bureau of Standards, Washington 1950 (n=1,...,49); W.J.Dixon and A.M.Mood, "The statistical sign test", Jrn. Am.Stat.Ass. 41 (1946), pp. 557-566 (n=1,...,100); W.J.Dixon and F.J.Massey, "Introduction to statistical analysis", N.Y., Toronto, London, 1951, Table 10 (n=1,...,90). Chr.L.Rümke, "De tekentoets, een methode van statistisch onderzoek", Tdschr voor Medische Analysten 7 (1952), pp. 54-58 (n=1,...,100); A.van Wijngaarden, "Table of the cumulative symmetric binomial distribution", Proc.Kon.Ned.Ak.v.Wet. 53 (1950), pp. 857-868, Indag.Math. 12 (1950), pp. 301-312 (n=1,...,200).

Indien n, het aantal waarnemingen, te groot is, om van deze tabellen gebruik te kunnen maken, of indien n niet zeer klein is en de tabellen niet beschikbaar zijn, kan men ook een benaderingsmethode gebruiken, waarvoor een tabel van de normale verdeling met gemiddelde 0 en spreiding 1 nodig is.

Zij ξ_α die waarde van een normaal verdeelde grootheid met gemiddelde 0 en spreiding 1, waarbij een éénzijdige overschrijdingskans α behoort. Voor α=0,10, 0,05, 0,025 resp. 0,01 is ξ_α=1,28, 1,64, 1,96 resp. 2,32. Zij evenzo $\xi_{\frac{1}{2}\alpha}$ die waarde, waarvoor de rechteroverschrijdingskans gelijk aan $\frac{1}{2}\alpha$ is. Het getal a wordt dan gevonden als het grootste gehele getal, dat $\leq \frac{1}{2}(n-1- \xi_{\frac{1}{2}\alpha}\sqrt{n})$ is, terwijl b het grootste gehele getal is, dat $\leq \frac{1}{2}(n-1- \xi_\alpha \sqrt{n})$ is. De methode blijft verder ongewijzigd.

Voorbeeld I. Indien bij voorbeeld I de ergotamine geen invloed op de zuurgraad uitoefent, is het optreden van een positief verschil even waarschijnlijk als van een negatief verschil, zodat de mediaan van de verdeling der verschillen in dat geval gelijk aan 0 is. Wij komen dus over de eventuele werking van het ergotamine iets te weten, indien wij een betrouwbaarheidsinterval voor de mediaan bepalen.

Nemen wij α=0,05,dan vinden wij uit de eerstgenoemde der bovenstaande tabellen, a=2. Een tweezijdig betrouwbaarheidsinterval voor de mediaan wordt dus gevormd door de op twee na kleinste en de op twee na grootste waarneming, zodat wij dus (zie tabel I) als uitkomst verkrijgen, dat de mediaan, behoudens een onbetrouwbaarheid 0,05, tussen de waarden 0 en 32 ligt, deze grenzen inbegrepen.

Passen wij de benaderingsmethode toe, eveneens voor een tweezijdig betrouwbaarheidsinterval, dan vinden wij voor $\frac{1}{2}(n-1- \xi_{\frac{1}{2}\alpha}\sqrt{n})$, daar $\xi_{0,025}$=1,96 en n=12 is, de waarde 2,11, zodat wij voor a de waarde 2 verkrijgen, die gelijk is aan de exact berekende.

De berekening van éénzijdige betrouwbaarheidsintervallen verloopt analoog.

De tekentoets.

Bij het eerste voorbeeld gaat het er eigenlijk niet zozeer om, een betrouwbaarheidsinterval voor de mediaan der verschillen te bepalen, als wel, om na te gaan, of de kans op een positief verschil gelijk is aan die op een negatief verschil, onder verwaarlozing van de kans op een verschil, dat gelijk aan 0 is. Immers, indien de ergotamine niet werkt, is niet alleen de mediaan van de verdeling van de verschillen gelijk aan 0, maar de kans op een positief verschil is gelijk aan die op een negatief en als de ergotamine wel invloed op de zuurgraad uitoefent, zijn deze kansen ongelijk. De verschillen, die gelijk aan 0 zijn kunnen daarom buiten beschouwing gelaten worden, indien het er uitsluitend om gaat, te toetsen, of de ergotamine invloed heeft of niet.

Deze toetsing kan nu als volgt uitgevoerd worden.

Wij geven het aantal waarnemingen, met weglating van de waarnemingen, die gelijk aan nul zijn, aan met n en het aantal positieve waarnemingen met x.
De exacte tweezijdige overschrijdingskans, behorende bij n en x, is dan, als $x > \frac{1}{2}n$ is, gelijk aan

$$2^{-n+1} \sum_{i=x}^{n} \binom{n}{i}$$

en, als $x < \frac{1}{2}n$ is

$$2^{-n+1} \sum_{i=0}^{x} \binom{n}{i},$$

terwijl voor $x=\frac{1}{2}n$ deze overschrijdingskans gelijk aan 1 is.

Is deze overschrijdingskans kleiner dan de van tevoren gekozen onbetrouwbaarheidsdrempel, dan wordt de getoetste hypothese verworpen.

De linker-éénzijdige overschrijdingskans is steeds

$$2^{-n} \sum_{i=0}^{x} \binom{n}{i}$$

en de rechter-éénzijdige

$$2^{-n} \sum_{i=x}^{n} \binom{n}{i}.$$

De overschrijdingskansen kunnen ook bij goede benadering berekend worden met behulp van een tabel van de normale verdeling. Daartoe berekent men in het tweezijdige geval de grootheid

$$y=\frac{|x-\frac{1}{2}n|-\frac{1}{2}}{\frac{1}{2}\sqrt{n}},$$

in het linkereénzijdige

$$y_l=\frac{x-\frac{1}{2}n+\frac{1}{2}}{\frac{1}{2}\sqrt{n}}$$

en in het rechtereénzijdige

$$y_r=\frac{x-\frac{1}{2}n-\frac{1}{2}}{\frac{1}{2}\sqrt{n}}$$

en zoekt in een tabel van de normale verdeling (met gemiddelde 0 en spreiding 1) de bij y behorende tweezijdige resp. de bij y_l behorende linker- of bij y_r behorende rechtereénzijdige overschrijdingskans op.

In voorbeeld I is nu n=11 en x=9. Daar bij het betrokken onderzoek zowel verlaging als verhoging van de zuurgraad door de ergotamine mogelijk werd geacht, wordt de toets tweezijdig toegepast. De tweezijdige overschrijdingskans is

$$2^{-10}\left\{\binom{11}{9}+\binom{11}{10}+\binom{11}{11}\right\},$$

hetgeen bij berekening, of bij opzoeken in één der genoemde tabellen, gelijk blijkt te zijn aan 0,065. De benadering met behulp van de normale verdeling geeft

$$y=\frac{|9-5\frac{1}{2}|-\frac{1}{2}}{\frac{1}{2}\sqrt{11}}=1,81.$$

De bijbehorende overschrijdingskans is 0,07. Voor grotere waarden van n wordt de benadering snel beter.

Gebruiken wij $\alpha=0,05$, dan kunnen wij dus de getoetste hypothese, die inhoudt, dat de ergotamine geen invloed op de zuurgraad heeft, niet verwerpen.

Opmerking. De tekentoets kan algemeen gebruikt worden, om de hypothese te toetsen, dat een waarschijnlijkheid gelijk aan $\frac{1}{2}$ is. Geven wij de uitkomst, waar het om gaat, aan met A, dan kan men de hypothese, dat de kans op A gelijk is aan $\frac{1}{2}$, toetsen door bij een reeks van n onafhankelijke experimenten het aantal malen, dat er A uit komt te tellen. Dit aantal is dan de grootheid x in de bovenstaande formules.

De toets van Wilcoxon.

Bij het tweede voorbeeld (zie tabel II) gaat het erom de hypothese H_0 te toetsen, inhoudende dat de beide waarnemingsreeksen uit eenzelfde verdeling afkomstig zijn. De traditionele toets voor dit geval is de toets van Student voor het verschil van de gemiddelden van twee normaal verdeelde grootheden met gelijke spreidingen. De onderstelling van normaliteit is niet nodig voor toepassing van de toets van Wilcoxon.

De toetsingsgrootheid van deze toets wordt als volgt berekend: Voor ieder der waarnemingen van groep a telt men, hoeveel waarnemingen van groep b kleiner zijn, terwijl bij gelijkheid $\frac{1}{2}$ geteld wordt. Deze aantallen worden opgeteld en de som wordt U genoemd.

In het voorbeeld is de eerste waarneming van groep I kleiner dan alle waarnemingen van groep II; deze geeft dus een bijdrage 0 tot U. De tweede waarneming (2,2) is gelijk aan één waarneming van groep II en kleiner dan de overige; dit geeft dus een bijdrage $\frac{1}{2}$. De derde en vierde geven weer een bijdrage 0, de vijfde en zesde geven ieder een bijdrage 1, de zevende 5 en de achtste 1. Wij vinden dus $U=8\frac{1}{2}$.

De waarschijnlijkheidsverdeling van U, onder de getoetste hypothese H_0, is bekend. Voor waarnemingsreeksen met minder dan 11 waarnemingen ieder zijn tabellen van deze verdeling beschikbaar.

Tabellen: H.B.Mann and D.R.Whitney, "On a test of whether one of two random variables is stochastically larger than the other", Ann.Math.Stat. 18 (1947), pp. 50-60; H.R.van der Vaart, "Gebruiksaanwijzing voor de toets van Wilcoxon", Rapport S 32 (M 4) van het Mathematisch Centrum, Amsterdam 1950.

Is één der beide aantallen waarnemingen, die wij met m en n aangeven, groter dan 10, dan kan men van een benaderingsmethode gebruik maken. Men vindt dan een benadering van de tweezijdige overschrijdingskans, door de grootheid

$$\frac{|U-\frac{1}{2}mn|-\frac{1}{2}}{\sqrt{\frac{1}{12}\,mn(m+n+1)}}$$

te berekenen en de bij deze waarde behorende tweezijdige overschrijdingskans op te zoeken in een tabel van de normale verdeling (met gemiddelde 0 van spreiding 1).

De linker- resp. rechteréénzijdige overschrijdingskans van U vindt men als de linker- resp. rechteréénzijdige overschrijdingskans van de grootheden

$$\frac{U-\frac{1}{2}mn+\frac{1}{2}}{\sqrt{\frac{1}{12}\,mn(m+n+1)}} \quad \text{resp.} \quad \frac{U-\frac{1}{2}mn-\frac{1}{2}}{\sqrt{\frac{1}{12}\,mn(m+n+1)}}\;,$$

evenzo opgezocht in een tabel van de normale verdeling.

De grootheden $\frac{1}{2}mn$ en $\sqrt{\frac{1}{12}\,mn(m+n+1)}$ zijn voor m en $n \leqq 100$ getabelleerd in "Auxiliary Table for Wilcoxon's two sample test," Report R 132/S 86 of the Mathematical Centre, Amsterdam. 1952.

In voorbeeld II is m=8 en n=6, terwijl $U=8\frac{1}{2}$ is. Voor U=9 [1]) vinden wij in de bovengenoemde tabel van de verdeling van U als tweezijdige overschrijdingskans de waarde 0,059 en voor U=8 de waarde 0,042. Deze uitkomst is, als wij voor α de waarde 0,05 nemen, juist op het randje.

Passen wij de benadering toe, dan vinden wij, bij tweezijdige toetsing,

$$\frac{|U-\frac{1}{2}mn|-\frac{1}{2}}{\sqrt{\frac{1}{12}\,mn(m+n+1)}} = \frac{|8\frac{1}{2}-24|-\frac{1}{2}}{\sqrt{\frac{1}{12}\,8.6.15}} = 1{,}93.$$

De bijbehorende tweezijdige overschrijdingskans, opgezocht in een tabel van de normale verdeling, is gelijk aan 0,054, hetgeen goed overeenstemt met de exacte uitkomsten. Er is wel een aanwijzing, dat de getoetste hypothese niet juist is, maar deze aanwijzing is niet zeer overtuigend.

---

1) De tabellen voor de toets van Wilcoxon houden geen rekening met het optreden van gelijke waarnemingen. Daarom komen alleen gehele waarden van U in deze tabellen voor.