

Statistische Afdeling
van het
Mathematisch Centrum,
Amsterdam

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

Rapport S 32.

Door Mejr. C. van Zede: en

Dr J. Hemelrijk Jr.

Vergelijking van het histaminegehalte van het
bloed van normale mannen en vrouwen lijdende
aan schizofrenie of andere psychosen.

Heenk 1252.

1. Inleiding.

Bij een aantal normalen, schizofrenen en lijders aan andere psychosen werd het histaminegehalte van het bloed enige malen bepaald

1. met tussenpozen van 24 uur ("dagwaarnemingen"),
2. met tussenpozen van 1 minuut ("minuutwaarnemingen"),
3. vrijwel zonder tussenpozen ("secondewaarnemingen").

De laatste waarnemingen werden alleen bij schizofrenen verricht. ¹⁾

De vraag is nu:

1. of er een verschil is tussen de histaminegehalten van het bloed van normalen, schizofrenen en lijders aan andere psychosen.
2. of het histaminegehalte van het bloed bij schizofrenen en lijders aan andere psychosen sterker wisselt dan bij normalen.

Bij de schizofrenen is een persoon (no.36) met een zeer hoog histaminegehalte, ook in vergelijking met de andere schizofrenen. Wij hebben deze patient als een uitzonderlijk geval beschouwd en hem derhalve niet in de statistische analyse betrokken. In de bij dit rapport gevoegde graphiek wordt deze patient vertegenwoordigd door het geheel rechts gelegen kringetje.

Alle proefpersonen waren mannen.

2. Methode van onderzoek en resultaten:

2.1. Onderzoek naar verschil in histaminegehalte:

Voor ieder der personen en voor ieder der bovengenoemde waarnemingsreeksen apart is berekend het gemiddelde $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ van de waarnemingen op de verschillende tijdstippen.

- 1) Deze gegevens zijn gepubliceerd in: J.A.H. Gooszen, Onderzoekingen over de histaminestofwisseling bij schizofrenen, Diss., Utrecht 1951, hoofdstuk VII.

De histaminegehalten van het bloed van de verschillende groepen personen zijn nu vergeleken door deze gemiddelden te vergelijken met behulp van de toets van Wilcoxon (zie bijlage S 47 (M 6) en (M 7)). In tabel I vinden we de resultaten voor de dagwaarnemingen; tabel II geeft de resultaten voor de minuutwaarnemingen:

Tabel I:²⁾

Vergelijking van het histaminegehalte van het bloed van normalen, schizofrenen en lijders aan andere psychosen met de toets van Wilcoxon:

(dagwaarnemingen):

	Overschrijdings- kans
normalen- andere psychosen	0,003-
normalen- schizofrenen	0,004-
andere psychosen schizofrenen	0,45-

Tabel II:³⁾

Vergelijking van het histaminegehalte van het bloed van normalen, schizofrenen en lijders aan andere psychosen met de toets van Wilcoxon:

(minuutwaarnemingen):

	Overschrijdings- kans
normalen- andere psychosen	0,04+
normalen- schizofrenen	0,32-
andere psychosen- schizofrenen	0,005-

2) Een + teken bij een overschrijdingskans betekent dat het histaminegehalte bij de eerstgenoemde groep personen hoger ligt dan bij de tweede groep: een - teken het omgekeerde.

3) Zie voetnoot 1.

Met de toets van Wilcoxon zijn ook nog de dag-,minuut- en secondenwaarnemingen voor de verschillende groepen personen vergeleken. Bij de schizofrenen is vergelijking van minuut- en secondenwaarnemingen met dagwaarnemingen niet goed mogelijk, daar enige personen zowel bij de minuut- (seconde-) als bij de dagwaarnemingen voorkomen. Laten wij deze personen buiten beschouwing, dan wordt het aantal waarnemingen te klein.

Tabel III geeft de vergelijking van dag- en minuutwaarnemingen voor normalen en lijders aan andere psychosen. In tabel IV vinden we de vergelijking van minuut- en seconde-waarnemingen voor schizofrenen:

Tabel III:⁴⁾

Vergelijking van dag- en minuutwaarnemingen bij normalen en lijders aan andere psychosen met de toets van Wilcoxon:

	Overschrijdingskans
normalen	0,09-
andere psychosen	0,003+

Tabel IV:⁵⁾

Vergelijking van minuut- en secondenwaarnemingen met de toets van Wilcoxon:

	Overschrijdingskans
schizofrenen	0,15+

De interpretatie van deze uitkomsten wordt in § 3 gegeven.

4) Een + teken bij een overschrijdingskans betekent dat het histaminegehalte bij de eerstgenoemde waarnemingen hoger ligt dan bij de tweede; een - teken het omgekeerde.

5) Zie voetnoot 4.

2.2. Onderzoek naar een sterkere wisseling van het histaminegehalte bij schizofrenen en lijders aan andere psychosen dan bij normalen:

Voor ieder der personen is de spreiding

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

berekend van de waarnemingen op de verschillende tijdstippen.

In een grafiek (zie figuur 1) zijn deze spreidingen uitgezet tegen de gemiddelden $\bar{x}$, voor de dag-, minuut- en seconde-waarnemingen afzonderlijk.

3. Interpretatie der resultaten en conclusies.

3.1. Niveauverschillen tussen de verschillende groepen proefpersonen.

De resultaten, vermeld in tabel I en II, stemmen slecht overeen. Opvallend is vooral, dat volgens tabel I de normalen een lager histamineniveau hebben dan de lijders aan andere psychosen, terwijl dit volgens tabel II juist andersom is. Voor dit verschijnsel bezitten wij geen verklaring. Uit tabel III blijkt bovendien, dat de lijders aan andere psychosen bij de dagwaarnemingen een hoger histamineniveau bezaten dan bij de minuutwaarnemingen, terwijl bij de normalen en schizofrenen dergelijke verschillen niet konden worden aangetoond (tabel III en IV; vgl. ook de grafiek).

De kleine overschrijdingskans (0,003; zie tabel III), die optreedt bij de vergelijking van de gemiddelde gehalten van de twee groepen lijders aan andere psychosen, wijst erop dat er systematische invloeden zijn geweest, die aan de aandacht zijn ontsnapt. Ten gevolge hiervan zullen verdere conclusies, uit dit onderzoek getrokken, slechts met voorbehoud kunnen worden geformuleerd.

De enige conclusie, die verder bij het onderzoek der niveauverschillen getrokken kan worden, is dat de normalen een lager histamineniveau bezitten dan de schizofrenen. Volgens tabel III is ook de overschrijdingskans bij vergelijking van de dag- en minuutwaarnemingen bij normalen nogal klein. Dit kan echter een gevolg zijn van het feit, dat de dagwaarnemingen zijn verricht bij het personeel van een psychiatrische inrichting, terwijl de minuutwaarnemingen van dienstplichtigen afkomstig zijn. Hier kunnen dus b.v.

leeftijds- of milieu-invloeden aanwezig zijn geweest. De conclusie, dat normalen doorgaans een lager histaminegehalte hebben dan schizofrenen is overigens ook reeds bij een andere reeks waarnemingen gevonden (zie hoofdstuk VI van het in voetnoot 1 vermelde proefschrift, waarin speciaal deze vraag aan de hand van andere waarnemingsmateriaal is onderzocht. Bij dit waarnemingsmateriaal werd, in tegenstelling tot het hier geanalyseerde, bij ieder proefpersoon slechts één bepaling verricht; het is daarom niet opnieuw in het huidige onderzoek betrokken).

3.2. Snelle schommelingen van het histaminegehalte.

Bij beschouwing van de grafiek krijgt men de indruk, dat bij de dagwaarnemingen de spreiding een weinig stijgt met het gemiddelde. Een dergelijke stijging zou betekenen, dat bij personen met een hoog gemiddeld gehalte de histaminevoorraad in het bloed hevigere wisselingen ondergaat dan bij personen met een laag gemiddelde. Deze eventuele afhankelijkheid tussen gemiddelde en spreiding komt echter in de waarnemingsresultaten slechts zo zwak tot uiting, dat het niet mogelijk geacht moet worden met het nu beschikbare waarnemingsmateriaal tot een conclusie te komen. Een nader onderzoek van dit verschijnsel zou, indien het hiervoor van voldoende belang wordt geacht, aan te bevelen zijn.

Een dergelijke afhankelijkheid tussen gemiddelde en spreiding is bij de minuut- en seconden-waarden niet te bespeuren. De spreidingen liggen hier bovendien in het algemeen lager dan die van de dagwaarnemingen. Dit heeft tot gevolg, dat er geen enkele reden aanwezig is, om te onderstellen, dat er ook bij de minuut- en seconde -waarnemingen variaties in de histaminegehalten van een proefpersoon zouden zijn opgetreden. De spreidingen kunnen, voor zover wij kunnen nagaan, zeer goed het gevolg van de meetnauwkeurigheden zijn geweest. Er is ook geen aanwijzing voor een verschil in deze spreidingen bij de drie verschillende groepen van proefpersonen. Daar verder de meetnauwkeurigheid bij de hier gebruikte waarnemingsresultaten geenszins als constant kan worden beschouwd (vgl. rapport S 75 van het Mathematisch Centrum), kunnen de klassieke analysemethoden

zoals de variantieanalyse niet worden toegepast.

3.3. Conclusies.

1. In overeenstemming met vroegere onderzoeken vertoont het waarnemingsmateriaal een vrij duidelijke aanwijzing, dat normalen een lager histamineniveau bezitten dan schizofrenen.

2. De conclusies omtrent het histamineniveau van lijders aan andere psychosen zijn bij dag- en minuutwaarnemingen met elkaar in strijd. De dagwaarnemingen wijzen op een hoger niveau dan bij normalen en de minuutwaarnemingen op een lager. Een verklaring hiervan hebben wij uit de beschikbare gegevens niet kunnen afleiden.

3. Er is in de minuut- en secondewaarnemingen geen aanwijzing gevonden voor snelle wisselingen van het histaminegehalte. Dit kan echter het gevolg zijn van de moeilijkheden, die de veranderlijke meetnauwkeurigheid bij de statistische analyse veroorzaakt, zodat hieruit niet geconcludeerd kan worden, dat dergelijke schommelingen in werkelijkheid niet toch zouden kunnen optreden. Bij de dagwaarnemingen bevatten de gegevens wel een suggestie voor een toeneming van de schommelingen met het gemiddelde gehalte, die echter veel te zwak is, om op grond van dit weinig uitgebreide materiaal bewijskracht te bezitten.

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 grootheden ²⁾, 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 zone van de toets, α de onbetrouwbaarheidsdrempel (Engels: level of significance). Voor α neemt men veelal de waarde 0,05 of 0,01.

Men verwerpt 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 verwerping 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 verwerping 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.

De toets van Wilcoxon. ¹⁾

Deze methode dient tot het toetsen van de hypothese H_0 , inhoudend dat twee steekproeven $x_1, \dots, x_n$ en $y_1, \dots, y_m$ afkomstig zijn uit één collectie (ook populatie of universum genaamd). Zij is strict genomen, toepasbaar onder de voorwaarde, dat er geen enkel paar waarden (x_i, y_j) is met $x_i = y_j$. Verdere voorwaarden zijn voor de toepassing niet nodig, terwijl ^{ook} de zojuist genoemde, indien er niet teveel dergelijke paren zijn, de toepassing van de toets weinig hindert.

De toetsingsgrootte U is het aantal paren (x_i, y_j) waarvoor $x_i > y_j$ is (het aantal "inversies"). Daar er $n \cdot m$ dergelijke paren zijn, kan U alle gehele waarden van 0 tot en met $n \cdot m$ aannemen. Is U groot, dan liggen er veel waarden x_i verder naar rechts dan waarden y_j , is U klein, dan juist weinig.

De kritieke zone K neemt nu daarom de kleine en de grote waarden van U en wel van beide zoveel, dat de gekozen onbetrouwbaarheidsdrempel α niet overschreden wordt.

Voor éénzijdige toetsing, te onderscheiden in linker- en rechter-éénzijdige toetsing, gebruikt men kritieke zones K'_1 , resp. K'_2 , die geheel bestaan uit kleine, resp. grote waarden van U .

Verwerping van H_0 ten gevolge van het vinden van een grote (resp. kleine) waarde van U wijst erop, dat $x_1, \dots, x_n$ en $y_1, \dots, y_m$ steekproeven uit verschillende collecties zijn, waarbij de op de x -collectie aangenomen waarden systematisch groter (resp. kleiner) dan de op de y -collectie aangenomen waarden zijn.

Litteratuur:

- F. Wilcoxon, Individual comparisons by ranking methods, Biometrics 1 (1945), p.80-83.
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), p.50-60. Bevat tabellen voor n en $m \leq 8$.
H.R. van der Vaart, Some remarks on the power function of Wilcoxon's test for the problem of two samples, Proceedings van de Kon.Ned.Ak.v.Wet., 53 (1950), p.494-520.
H.R. van der Vaart, Gebruiksaanwijzing voor de toets van Wilcoxon, met tabellen voor n en $m \leq 10$, Rapport S32 (M4)(1950).
D. van Dantzig, Kadercursus Mathematische Statistiek, Math. Centrum, Amsterdam(1947-50), hoofdstuk 6, 3.

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

