Statistische Afdeling
van het
Mathematisch Centrum,
Amsterdam

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

Rapport S 93

Vergelijking van het histaminegehalte van het bloed van normale kinderen en kinderen lijdende aan allergische aandoeningen en aan tuberculose

door

Mej. C. van Eeden.

Augustus 1952

1. Inleiding:

Van een aantal normale kinderen en kinderen lijdende aan allergische aandoeningen en aan tuberculose werd enige malen, met tussenpozen van 20 seconden, het histaminegehalte van het bloed bepaald.

De kinderen lijdende aan t.b.c. werden gedeeltelijk verpleegd in Groesbeek, gedeeltelijk in het Onze Lieve Vrouwen Gasthuis (O.L.V.G.) te Amsterdam; de waarnemingen van de kinderen met allergische aandoeningen kwamen alle uit het O.L.V.G. De normale kinderen waren kinderen met lichte chirurgische afwijkingen uit het O.L.V.G.; de bepalingen werden hier verricht vlak vóór het ontslag uit het ziekenhuis.

De bepalingen werden bij de normale kinderen verricht in de maanden November 1949 t/m Augustus 1950. Bij de kinderen lijdende aan t.b.c. in de maanden November 1949 t/m Januari 1950 (O.L.V.G.) en in Februari 1950 (Groesbeek); bij de kinderen met allergische aandoeningen in de maanden November 1949 t/m April 1950.

Van alle kinderen was leeftijd en geslacht opgegeven.

2. Methode van onderzoek en resultaten:

Voor ieder der kinderen is het gemiddelde berekend van de waarnemingen op de verschillende tijdstippen. Deze gemiddelden zijn gebruikt om de groepen kinderen onderling te vergelijken. De gemiddelde gehalten van ieder der groepen kinderen in iedere maand zijn in grafiek 1 uitgezet[1].

2.1. Onderzoek naar een verschil tussen de histaminegehalten in de verschillende maanden:

Dit onderzoek is uitgevoerd voor ieder der drie groepen kinderen apart.

De waarnemingen zijn ingedeeld in groepen naar de maanden waarin de waarnemingen verricht zijn. Met de toets van KRUSKAL[2] is nu onderzocht of er een verschil is tussen de histaminegehalten in de verschillende maanden. De resultaten hiervan staan vermeld in tabel I:

---

1) Dit zijn de rekenkundige gemiddelden van de voor ieder kind berekende gemiddelden.

2) W.H. Kruskal, A nonparametric analogue based upon ranks of one-way analysis of variance, Ann. Math. Stat. 23 (1952), p. 140.

Tabel. I

Onderzoek naar een verschil tussen de histaminegehalten in de verschillende maanden:

	H	C-1	Overschrijdingskans
normaal	20,3	9	0,017
allergische aandoeningen	7,8	5	0,17
tuberculose	10,7	3	0,014

Wij vinden dus, zoals grafiek 1 reeds doet vermoeden, zowel bij de normale kinderen als bij de kinderen lijdende aan tuberculose een verschil tussen de histaminegehalten in de verschillende maanden. Bij de kinderen met allergische aandoeningen vinden wij geen verschil tussen de maanden. Het aantal waarnemingen is hier echter klein: er zijn twee maanden bij met ieder slechts één waarneming.

2.2 Onderlinge vergelijking der drie groepen kinderen:

Daar wij een verschil vinden tussen de verschillende maanden kunnen wij de drie groepen kinderen slechts onderling vergelijken in ieder der maanden apart. De resultaten der verschillende maanden worden daarna gecombineerd.

Dit is als volgt gebeurd:

Van twee groepen kinderen zijn de histaminegehalten, in ieder der maanden apart, vergeleken met behulp van de toets van WILCOXON (zie bijlage S 47 (M 7)). Wij krijgen dan voor iedere maand:

1. een waarde voor de toetsingsgrootheid $\underline{U}$,
2. een waarde voor $\mu_{\underline{U}}$,
3. een waarde voor $\sigma^2_{\underline{U}}$.

We noemen deze resp. U_i, μ_i en σ_i^2.

Indien de te toetsen hypothese H_0, dat er geen verschil is tussen de histaminegehalten van de twee groepen kinderen, juist is, is $U = \sum_i \underline{U}_i$ bij benadering normaal verdeeld met gemiddelde $\mu = \sum_i \mu_i$ en variantie $\sigma^2 = \sum_i \sigma_i^2$. Is H_0 onjuist, dan zal de bij het experiment gevonden waarde van $\frac{U-\mu}{\sigma}$ sterk van nul afwijken. De kritieke zône bestaat dus uit grote waarden van $\left|\frac{U-\mu}{\sigma}\right|$.

De resultaten staan vermeld in tabel II:

Tabel II[3)]

Onderlinge vergelijking van de histaminegehalten der drie groepen kinderen:

	$U-\mu$	σ	$\frac{U-\mu}{\sigma}$	Overschrijdingskans
normaal-tuberculose	30	24,5	1,22	0,22 -
normaal-allergisch	1	9,6	0,10	0,92 -
tuberculose-allergisch	-8	11,4	-0,70	0,48 +

Uit tabel II zien we dus, dat er geen reden is om aan te nemen, dat er een verschil is tussen de histaminegehalten van de drie groepen kinderen.

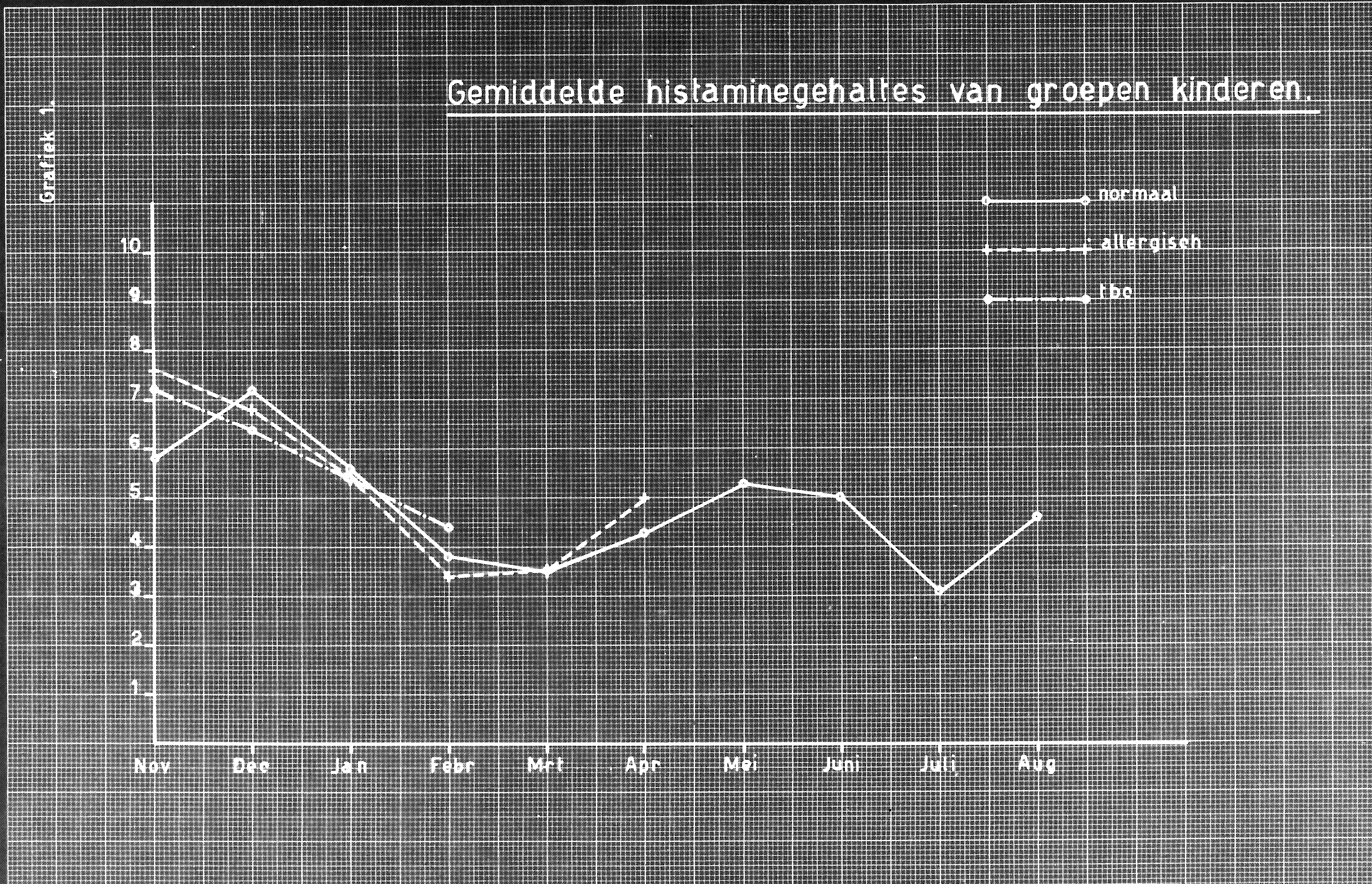
Opmerkingen:

1. Dit resultaat betekent niet, dat er werkelijk geen verschil is, maar slechts, dat het eventuele verschil bij de huidige proefopzet niet gevonden is. Grafiek 1 doet echter vermoeden, dat een eventueel aanwezig verschil toch wel zeer gering zal zijn.

2. Een tweede door de onderzoekers gestelde vraag was, of er een verschil is tussen de histaminegehalten van kinderen en volwassenen. De vergelijking van deze histaminegehalten is slechts mogelijk op de in § 2.2 beschreven wijze voor de onderlinge vergelijking der groepen kinderen. Hiervoor is dus nodig een reeks waarnemingen van het histaminegehalte van volwassenen, die verricht zijn in dezelfde maanden als de waarnemingen bij de kinderen. Daar een dergelijke reeks waarnemingen ons niet ter beschikking staat, is het ons niet mogelijk het histaminegehalte van kinderen met dat van volwassenen te vergelijken.

----

3) Het teken + bij een overschrijdingskans betekent, dat het histaminegehalte van de eerstgenoemde groep kinderen hoger ligt, dan dat van de tweede.


Grafiek 1.
Gemiddelde histaminegehaltes van groepen kinderen.
normaal
allergisch
tbc
10
9
8
7
6
5
4
3
2
1
Nov
Dec
Jan
Febr
Mrt
Apr
Mei
Juni
Juli
Aug

MATHEMATISCH CENTRUM
Amsterdam
Statistische Afdeling.

S47(M6)

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

De toetsing van een hypothese $\mathcal{H}_0$ berust steeds op een aantal waarnemingen $x_1, x_2, \ldots, x_n$ van één of meer stochastische grootheden[2)], of op enige groepen van waarnemingen (bv. twee steekproeven).

Bij een toets behoort een toetsingsgrootheid $\underline{u}$ (soms meer dan één), die een functie is van bovengenoemde stochastische grootheden en die, voor de waargenomen waarden $x_1, x_2, \ldots, x_n$ 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 $\mathcal{H}_0$ juist is, de waarschijnlijkheidsverdeling van deze grootheid kan berekenen.

Vervolgens kiest men een verzameling Z van mogelijke uitkomsten van $\underline{u}$, en wel op zodanige wijze, dat de kans, dat $\underline{u}$ een in Z gelegen waarde aanneemt, onder de hypothese $\mathcal{H}_0$, gelijk is aan een gegeven getal α, zodat Z dus van α afhankelijk is[3)]. 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 verwerpt nu $\mathcal{H}_0$ op grond van de waarnemingen $x_1, x_2, \ldots, x_n$, indien de bij deze waarnemingen behorende waarde van $\underline{u}$ 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 $\mathcal{H}_0$ 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 <u>geen</u> mogelijkheid om tot <u>aanvaarding</u> 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 <u>nevenvoorwaarden</u> 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 <u>overschrijdingskans</u> 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 toetsingsgrootheid nog juist in de (bij α behorende) kritieke zône Z ligt. <u>Wordt dus de waarde k opgegeven en werkt men met onbetrouwbaarheidsdrempel α, dan wordt verworpen, indien $k \leqq \alpha$ is</u>.

Voor het onderscheid tussen één- en tweezijdige toetsing en de keuze tussen deze twee mogelijkheden vergelijke men bv. de tweede hieronder gegeven litteratuurplaats. 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.

<u>Litteratuur:</u>

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 <u>4</u> (1950) p.54-66.

Mathematisch Centrum,
2de Boerhaavestraat 49,
Amsterdam O.
Statistische Afdeling,
S47 (M7).

Maart, 1952.

De toets van Wilcoxon.[1]

Deze methode dient tot het toetsen van de hypothese H_o, inhoudende, dat twee steekproeven $x_1,\ldots,x_n$ en $y_1,\ldots,y_m$ afkomstig zijn uit één collectie (ook wel populatie of universum genaamd).

Voor het toetsen van de hypothese H_o wordt gebruik gemaakt van een toetsingsgrootheid $\underline{U}$ [2], die als volgt uit de waarnemingen berekend wordt. Onderstellen we, dat de waarnemingen $x_1,\ldots,x_n$ en $y_1,\ldots,y_m$ naar opklimmende grootte gerangschikt zijn, dan bepalen we eerst het aantal waarnemingen uit de tweede steekproef, dat kleiner is dan de kleinste waarneming x_1 uit de eerste steekproef (bij gelijkheid tellen wij $\frac{1}{2}$ in plaats van 1). Noem dit aantal V_1. Vervolgens wordt het aantal waarnemingen uit de tweede steekproef bepaald, dat kleiner is dan de op één na kleinste waarneming x_2 uit de eerste steekproef (bij gelijkheid wordt weer een $\frac{1}{2}$ in plaats van 1 geteld). Dit aantal noemen we V_2. Evenzo worden met betrekking tot $x_3,x_4,\ldots,x_n$ de aantallen $V_3,V_4,\ldots,V_n$ bepaald. De waarde U van de toetsingsgrootheid $\underline{U}$ wordt voor de twee steekproeven dan gegeven door

$$U=V_1+V_2+\ldots+V_n.$$

Wanneer onder de waarnemingen niet te veel gelijken voorkomen, kan bewezen worden, dat de toetsingsgrootheid $\underline{U}$ onder de hypothese H_o voor grote waarden van n en m (beide $\geqq$ 10) bij benadering een normale verdeling bezit. De waarnemingen $x_1,\ldots,x_n$ en $y_1,\ldots,y_m$ tezamen genomen vallen uiteen in een aantal groepen van gelijke waarnemingen. Noem het aantal van deze groepen k, dan is k minstens 1 (als alle waarnemingen gelijk zijn) en hoogstens m+n (als alle waarnemingen verschillend zijn).

---

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

2) Stochastische grootheden worden door onderstreping aangeduid.

Zijn $t_1,\dots,t_k$ de aantallen waarnemingen in deze groepen van gelijken, dan worden het gemiddelde μ en de variantie σ^2 van de toetsingsgrootheid $\underline{U}$ gegeven door

$$\mu(\underline{U})=\tfrac{1}{2}nm,$$

en

$$\sigma^2=\mathrm{Var}\ (\underline{U})=\frac{1}{12}\frac{nm}{(n+m)(n+m-1)}\left\{(n+m)^3+\right.$$
$$\left.-(t_1{}^3+t_2{}^3+\dots+t_k{}^3)\right\} \qquad {}^{1})$$

De grootheid $\mu(\underline{U})$ is dus onafhankelijk van de waarden vast. Indien de hypothese H_0 niet vervuld is, zal de grootheid $\underline{U}$ grote of kleine waarden bezitten, al naar gelang $\underline{y}$ systematisch kleiner of groter is dan $\underline{x}$.

De (tweezijdige) toets bestaat nu daarin, dat men H_0 verwerpt indien de gevonden waarde U van $\underline{U}$ te sterk van μ afwijkt, d.w.z. als

$$\frac{|U-\mu|}{\sigma}>\xi_\alpha\ , \qquad {}^{2})$$

waarin α de onbetrouwbaarheidsdrempel is en ξ_α volgt uit

$$\frac{1}{\sqrt{2\pi}}\int_{\xi_\alpha}^{\infty}e^{-\frac{1}{2}x^2}dx=\tfrac{1}{2}\alpha\ ,$$

en in een tabel van de normale verdeling kan worden opgezocht.

De (tweezijdige) overschrijdingskans k , behorende bij T, is gedefiniëerd als

$$k=\frac{2}{\sqrt{2\pi}}\int_{\left|\frac{U-\mu}{\sigma}\right|}^{\infty}e^{-\frac{1}{2}x^2}dx \qquad {}^{2})$$

en kan ook in een tabel van de normale verdeling worden gevonden.

Bij eenzijdige toetsing wordt α door 2α vervangen, resp. k gehalveerd.

Een bijzonder geval van het bovenstaande is, dat onder de waarnemingen voor $\underline{x}$ en $\underline{y}$ in 't geheel geen gelijken voorkomen. In dat geval kan de uitdrukking voor de variantie herleid worden tot

$$\sigma^2=\frac{1}{12}\,nm(n+m+1).$$

---

[1]) Deze formule is een door T.J.Terpstra gegeven vereenvoudiging van de door J.Hemelrijk([5] en [7]) afgeleide formule. De afleiding van deze vereenvoudigde formule zal nog gepubliceerd worden.

[2]) Deze formules berusten op de normale benadering van de verdeling van $\underline{U}$.

Indien n en m kleiner zijn dan 10, zijn tabellen beschikbaar voor het berekenen van de overschrijdingskans k voor de uit de steekproef bepaalde waarde U van U (zie [2] en [4]). Dergelijke tabellen bestaan echter niet voor het geval, dat gelijke waarnemingen optreden.

Opmerking. Men kan gemakkelijk bewijzen, dat de variantie van U door het optreden van gelijke waarnemingen vermindert. Het verschil, dat door deze gelijken optreedt, is echter in het algemeen gering. Men kan daarom in eerste instantie deze correctie op σ^2 verwaarlozen. De overschrijdingskansen, die men dan vindt, zijn iets te groot.

Litteratuur:

1. F.Wilcoxon, Individual comparisons by ranking methods, Biometrics 1 (1945), p.80-83.
2. H.B.Mann and D.R.Whitney, On a test of whether one of two random variables is stochastically larger than the other, Amer.Math.Stat. 18 (1947),p. 50-60.
3. 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.
4. H.R.van der Vaart, Gebruiksaanwijzing voor de toets van Wilcoxon, met tabellen voor n en $m \leq 10$, Rapport S32 (M4) (1950).
5. H.R.van der Vaart, De toets van Wilcoxon voor het probleem van twee steekproeven. (Cursus "Parametervrije Methoden", 1951-'52).
6. D.van Dantzig, Kadercursus Mathematische Statistiek, Math. Centrum, Amsterdam (1947-'50), hoofdst. 6, § 3.
7. J.Hemelrijk, Note on Wilcoxon's two sample test, when ties are present, Ann.Math.Stat. 23 (1952) no. 2.