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 175

Onderzoek naar de diagnostische waarde van de circulatietijd van het bloed.

door

Ir A.R. Bloemena

en

Ir Doralien Wabeke

Juli
1955

1. Beschrijving van het onderzoek.

Het onderzoek had tot doel na te gaan welke invloeden bepaalde ziekten hebben op de circulatietijd (CT) van het bloed bij de menselijke bloedsomloop.

Het onderzoek strekte zich uit van 1 Mei 1952 - 1 Mei 1953 en het had betrekking op ongeveer 500 proefpersonen, alle klinische of poliklinische patiënten.

Het meten van de CT geschiedde door de tijd te bepalen, die verliep vanaf het moment dat de proefpersoon een methylblauw injectie kreeg tot het ogenblik waarop de kleurstof het voorhoofd bereikte. Het concentratieverloop van de kleurstof in de bloedvaten van het voorhoofd is met behulp van een foto-electrisch apparaat, aangebracht op het voorhoofd, gemeten en fotografisch vastgelegd. Op deze wijze verkreeg men krommen van de gedaante als in figuur 1 getekend is.

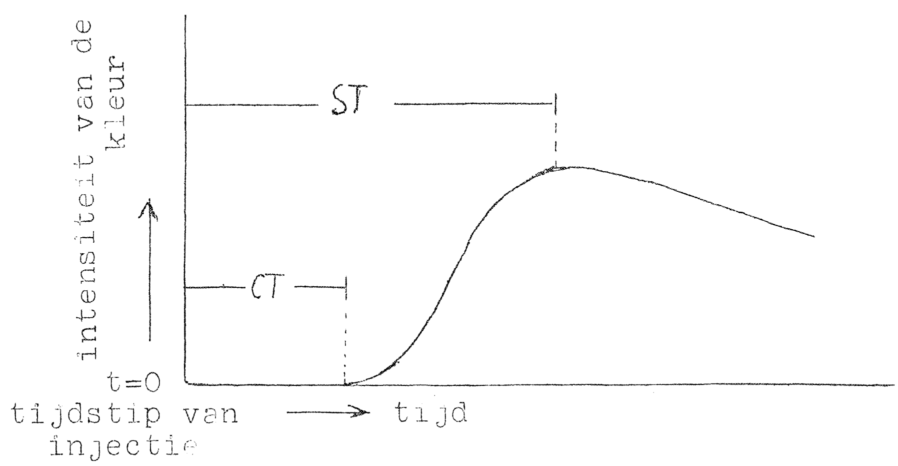

Fig. 1 Verloop van de intensiteit van de kleur met de tijd.

Behalve de grootheid, die als CT is gedefinieerd, is in deze figuur ook aangegeven een tweede grootheid, die bij het onderzoek gemeten werd: de stijgtijd (ST). Deze ST werd echter op verzoek van de opdrachtgever bij dit onderzoek verder buiten beschouwing gelaten.

Het onderzoek is verricht ten einde na te gaan of personen, die lijden aan een van de hierna volgende ziekten een andere CT hebben dan personen met een normale circulatie:

A anaemie
B angina pectoris
C asthma bronchiale
D emphyseem
E hypertensie
F hyperthyreoidie
G locale longafwijkingen, behalve longcarcinoom
H long-carcinoom.

Geen enkele patiënt leed aan meer dan één ziekte.

Een aantal proefpersonen leed aan een ziekte, waarvan men mocht veronderstellen dat deze de normale bloedcirculatie niet zou beïnvloeden, b.v. maagzweren of oogafwijkingen. Indien dan ook, op grond van andere waarnemingen dan die van de CT, geen afwijking van de normale circulatie werd gevonden, dan beschouwde men deze personen als personen met een normale circulatie. De circulatietijden van deze groep werden bij de statistische bewerking van het materiaal gebruikt als vergelijkingsbasis.

Bij de anaemiepatiënten werd tevens het haemoglobine-gehalte van het bloed bepaald, bij patiënten die leden aan hypertensie werd de diastolische en de systolische bloedruk gemeten en bij diegenen, die aan hypertensie leden, werd het percentage basaalmetabolisme (BM) bepaald. Deze gegevens zullen wij in dit rapport aanduiden met de term "extra gegevens".

Naar aanleiding van dit onderzoek hebben wij de volgende vragen beantwoord:

a) is er verband tussen leeftijd en CT?
b) is er verband tussen de CT en de extra gegevens en
c) zijn er verschillen tussen de CT-en van de personen met normale circulatie en die van patiënten, die aan één van bovengenoemde ziekten lijden?

Naar aanleiding van de beantwoording van deze vragen zijn in par. 4 enkele beschouwingen over de diagnostische waarde van de CT opgenomen.

2. Statistische bewerking van het materiaal.

(Alle toetsingen zijn voor mannen en vrouwen apart uitgevoerd)

2.1. Het verband tussen leeftijd en CT is getoetst met behulp van de rangcorrelatie-methode van KENDALL, waarvan een beschrijving wordt gegeven in Memorandum S 47 (M 13a), dat als bijlage aan dit rapport is toegevoegd. De gevonden overschrijdingskansen zijn vermeld in Tabel I, waarbij een plus-teken achter een overschrijdingskans aangeeft, dat in de waarnemingen een (min of meer) overeenstemmend verloop is gevonden tussen de beide gegevens, m.a.w. dat als de één stijgt, ook de ander vaker in grootte toe dan afneemt.

Achter elke overschrijdingskans is het aantal waarnemingen weergegeven waarop deze is gebaseerd. Kleine overschrijdingskansen (in de tabellen door onderstreping aangegeven) wijzen op de aanwezigheid van een systematisch verband van de beide vergeleken grootheden. Uit deze tabel blijkt, dat, althans in bepaalde groepen, de CT met de leeftijd stijgt.

2.2. De extra-gegevens zijn op dezelfde wijze bewerkt. Eerst is het verband onderzocht tussen deze gegevens en de leeftijd. Een systematisch verband werd echter in geen enkel geval gevonden. Daarna werd het verloop van de CT en dat van de extra-gegevens getoetst op analoge wijze als in 2.1. De overschrijdingskansen zijn vermeld in Tabel I.

2.3. Bij het vergelijken van de CT van patiënten, die aan een bepaalde ziekte lijden met die van personen met een ongestoorde circulatie, moet rekening gehouden worden met het feit dat de circulatietijd met de leeftijd stijgt. Om deze reden is de vergelijking op de volgende wijze uitgevoerd. De patiënten zijn alle ingedeeld in de leeftijdsklassen0-9, 10-19,...,80-89. Wij nemen nu aan dat de CT in deze leeftijdsklassen slechts zo weinig stijgt, dat dit effect te verwaarlozen is. Nu worden de CT's van patiënten lijdende aan een bepaalde ziekte voor elke leeftijdsklasse apart vergeleken met die van personen met normale circulatie met behulp van de toets van WILCOXON (zie bijlage M 7). Daarna zijn de resultaten van de verschillende leeftijdsklassen gecombineerd op de wijze, beschreven in het bijgevoegde memorandum M 17b, methode 1. De verkregen overschrijdingskansen zijn opgenomen in Tabel II, waarbij een plus-teken aangeeft, dat de CT van personen met normale circulatie bij dit onderzoek kleiner is dan van personen met een bepaalde ziekte. Kleine overschrijdingskansen wijzen op de aanwezigheid van een systematisch verschil van het vermelde teken tussen de CT'svande onderzochte groepen patiënten.

De toetsing is slechts uitgevoerd voor een aantal van de in par. 1 genoemde ziekten. Voor de andere ziekten waren de aantallen gegevens zo klein, dat statistisch onderzoek weinig zin had. Voorts zijn bij enkele ziekten bijna uitsluitend mannen of uitsluitend vrouwen onderzocht, hetgeen bepaalde vergelijkingen onmogelijk maakt.

3. Conclusies.

Bij de conclusies hebben wij alleen waarde gehecht aan overschrijdingskansen, die kleiner dan 0,05 zijn.

3.1. Uit Tabel 1 blijkt dat er voor personen met een normale circulatie een duidelijk positieve correlatie aanwezig is tussen de leeftijd en de circulatietijd.
Bij personen met een bepaalde ziekte komt deze ~~circulatie~~ correlatie niet altijd tot uiting in de gegevens van dit onderzoek. Deels kan dit te wijten zijn aan een te klein aantal waarnemingen in een bepaalde groep, voor het overige kan de invloed van de ziekte op de CT het verband tussen leeftijd en circulatietijd geheel of gedeeltelijk verstoren.

3.2. Ten aanzien van de "extra-gegevens" is er bij dit onderzoek geen enkele aanwijzing gevonden voor een verband tussen deze gegevens en de CT.

3.3. Uit tabel II volgt, dat vrouwen met hyperthyreoidie een kleinere circulatietijd hebben dan vrouwen met normale circulatie. Voor deze groep zijn de gegevens van de mannen niet statistisch onderzocht, daar het aantal daarvan te klein is. Voorts blijkt dat mannen die aan longcarcinoom lijden, een grotere CT hebben dan mannen met een normale circulatie. In deze groep kwamen slechts drie vrouwen voor, om welke reden geen uitspraak gedaan kan worden over de CT van vrouwen met longkanker.

4. De diagnostische waarde van de CT.

Hoewel het materiaal te gering van omvang is, om een gedetailleerd inzicht te kunnen verkrijgen in de diagnostische waarde van de CT, kunnen toch enkele algemene resultaten gegeven worden, die van belang zouden kunnen zijn bij een eventueel verder onderzoek naar de invloeden op de CT. Wij zijn er ons hierbij van bewust, dat de practische interpretatie van deze resultaten een medisch probleem is, dat niet onder onze competentie valt.

Bij de beoordeling van de diagnostische waarde van de CT moet uiteraard rekening worden gehouden met de mogelijkheid, dat ook andere factoren, waaronder niet onderzochte ziekten, invloed kunnen hebben op de CT.

De volgende resulaten zijn bij het onderzoek verkregen:

1. Mannen met longcarcinoom hebben een hogere CT dan mannen met een ongestoorde circulatietijd. Ten einde een indruk te krijgen van de diagnostische waarde van dit verschil hebben wij de frequentieverdelingen getekend voor de CT. Daarbij was het wenselijk het verloop van de CT met de leeftijd zo goed mogelijk te elimineren door het toepassen van een correctie op de CT.[1] Voor dit geval hebben wij gebruik gemaakt van een vrij ruwe voorlopige correctie, omdat een meer verfijnde correctie-methode weinig zinvol was in verband met de, statistisch gezien, vrij geringe aantallen onderzochte personen. De hier toegepaste correcties zijn:

voor de leeftijd beneden 35 jaar:	$+\frac{1}{2}$
" " " tussen 35 en 50 jaar:	0
" " " boven 50 jaar:	$-\frac{1}{2}$

De frequentieverdelingen van de CT's zijn getekend in fig. 2. Uit deze figuur zien wij dat bij het onderzoek een gecorrigeerde CT groter dan 9 bijna alleen voorkomt bij longcarcinoom-patiënten en dat van de onderzochte longcarcinoom-patiënten 45% een CT had, die groter is dan 9. Deze cijfers zijn natuurlijk sterk afhankelijk van de (vrij willekeurige) correctiemethode, maar geven toch wel de indruk dat de CT enige diagnostische waarde bezit ten aanzien van longkanker.

Een complicatie wordt gevormd door de mannelijke hypertensie-patiënten, die wat CT betreft tussen de normale mannen en de longkanker-patiënten inliggen. Het aantal patiënten in de hypertensie-groep is echter te klein om de verschillende complicaties die zouden kunnen optreden te kunnen beoordelen uit de gegevens van het onderzoek.

2. Brengen wij ook voor de vrouwen een correctie aan op de CT en wel:

voor de leeftijd beneden 65 jaar:	0
" " " boven 65 jaar:	$-2\frac{1}{2}$,

dan zien wij uit de frequentieverdelingen (fig. 3) dat slechts

---

1) Vgl. T.J. TERPSTRA EN J. HEMELRIJK: Centrale veneuze druk bij mannen en vrouwen met een normale tractus circulatoris, aanhangsel bij J.A. MOLHUYSEN, De centrale veneuze druk, dissert. Amsterdam 1953.

zeer weinig vrouwen met normale circulatie een gecorrigeerde CT $\leqq$ 5 bezitten, terwijl een groot percentage der hyperthyreoïdie-patiënten (bij dit onderzoek ongeveer 35%) een CT $\leqq$ 5 heeft.

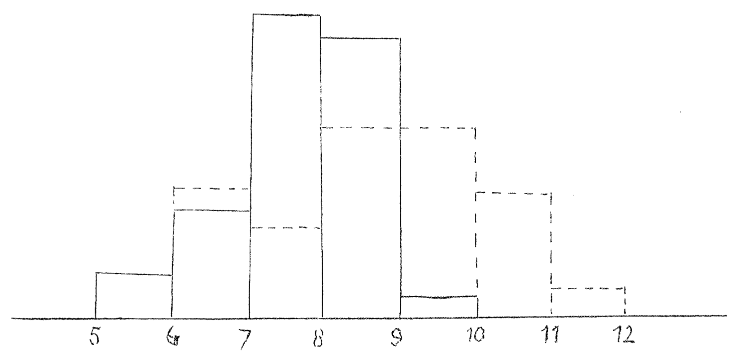

Fig. 2. Frequentieverdeling van de gecorrigeerde circulatietijden van mannen 1) met normale circulatie (———) 2) met longcarcinoom (------)

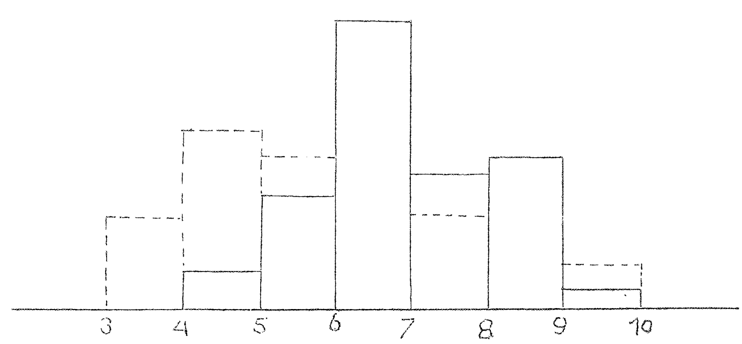

Fig. 3. Frequentieverdeling van de gecorrigeerde circulatietijden van vrouwen 1) met normale circulatietijden (————) 2) met hyperthyreoïdie (---------).

Hieruit kunnen wij afleiden dat de CT ook ten aanzien van de hyperthyreoïdie een diagnostische waarde bezit.

Resumé.

Uit het onderzoek is gebleken dat de CT hulp zou kunnen bieden bij het diagnostiseren van

1 longcarcinoom (bij mannen)

2 hyperthyreoidie (bij vrouwen).

Een nader onderzoek van meer patiënten en personen met normale circulatie is echter gewenst om de frequentieverdelingen en de leeftijdscorrectie nauwkeuriger te leren kennen.

Tabel I: Resultaten van het onderzoek naar de afhankelijkhei
en de extra - gegevens enerzijds en de circulatieti

De + en - tekens geven het teken van de uit de waarnemingen bereke

	normale circulatie		anaemie		angina pectoris		asthma		emphy-seem		
	k	n	k	n	k	n	k	n	k	n	
leeftijd m - CT m	0,006+	35	0,54-	4	1,00	3	0,54+	4	0,72-	13	
leeftijd v - CT v	<0,002+	39	0,86+	6	0,67+	3	0,34+	8	-	1	
											sy
leeftijd m- extra gegevens m			0,75+	4							O
" v - " " v			0,72-	6							O
CT m - " " m			0,21-	4							O
CT v - " " v			0,60+	6							O

Tabel II

Resultaten van de vergelijking van circulatietijden van groepen patiënten.

Een + teken geeft aan, dat de tweede groep een kleinere circulatietijd heeft dan de eerste groep.

	k	aantallen
Emphyseem m - normale m	0,83 +	11-22
Hypertensie m - normale m	0,62 +	10-22
Hypertensie v - normale v	0,77 +	22-34
Hyperthyreoïdie v - normale v	<0,003 -	16-39
Locale longafwijking m - normale m	0,92 -	33-33
Locale longafwijking v - normale v	0,34 -	24-29
Longcarcinoom m - normale m	0,016 +	22-22
Hypertensie m - longcarcinoom m	0,14 -	10-22
Hyperthyreoïdie v - locale longafwijking v	0,013 -	13-24

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, ..., 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 α = 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 $\mathcal{H}_0$ 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 waarnemingén 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 $\mathcal{H}_0$ 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 $k \leq \alpha$ 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 $\mathcal{H}_0$ 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_0, inhoudende, dat twee steekproeven $x_1,\dots,x_n$ en $y_1,\dots,y_m$ afkomstig zijn uit één collectie (ook wel populatie of universum genaamd).

Voor het toetsen van de hypothese H_0 wordt gebruik gemaakt van een toetsingsgrootheid $\underline{U}$ [2]), die als volgt uit de waarnemingen berekend wordt. Onderstellen we, dat de waarnemingen $x_1,\dots,x_n$ en $y_1,\dots,y_m$ 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 $x_3,x_4,\dots,x_n$ de aantallen $V_3,V_4,\dots,V_n$ bepaald. De waarde U van de toetsingsgrootheid $\underline{U}$ wordt voor de twee steekproeven dan gegeven door

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

Wanneer onder de waarnemingen niet te veel gelijken voorkomen, kan bewezen worden, dat de toetsingsgrootheid $\underline{U}$ onder de hypothese H_0 voor grote waarden van n en m (beide $\geq$ 10) bij benadering een normale verdeling bezit. De waarnemingen $x_1,\dots,x_n$ en $y_1,\dots,y_m$ 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 $t_1,\ldots,t_k$ 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

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

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 eenttabel 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 $\underline{k}$ voor de uit de steekproef bepaalde waarde U van $\underline{U}$ (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 $\underline{U}$ 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. 494-500.
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.

MATHEMATISCH CENTRUM
2de Boerhaavestr. 49,
A m s t e r d a m -O.
Statistische Afdeling
S47(M13)

Rangcorrelatie[1]

1. Beschrijving van de methode.

De door M.G. Kendall ontwikkelde methode der rangcorrelatie is toepasbaar op de volgende situatie:

De stochastische grootheden $\underline{x}$ en $\underline{y}$ bezitten een simultane verdeling. Over deze verdeling zelf behoeft niets ondersteld te worden.

(x_i, y_i) $(i = 1, \ldots, n)$, zijn onafhankelijke waarnemingsparen van deze stochastische grootheden

Voorbeeld:

i =	1	2	3	4	5	6
x_i	0,11	0,12	0,10	0,11	0,15	0,13
y_i	3,4	3,0	3,2	3,5	3,5	3,5

Wij zeggen dat de waarnemingsparen (x_i, y_i) en (x_j, y_j) positief gecorreleerd zijn, als de volgorde van x_i en x_j hetzelfde is als die van y_i en y_j (bv. $x_i < x_j$ en $y_i < y_j$); zij zijn negatief gecorreleerd als de volgorde van x_i en x_j tegengesteld is aan de volgorde van y_i en y_j (bv. $x_i > x_j$ en $y_i < y_j$) en zij zijn niet gecorreleerd als $x_i = x_j$ of $y_i = y_j$.

In tabel 1 hebben wij van alle tweetallen (x_i, y_i) en (x_j, y_j) uit ons voorbeeld nagegaan of zij positief, negatief dan wel niet gecorreleerd zijn. Een positieve correlatie is aangeduid met +1, een negatieve met -1, terwijl het ontbreken van correlatie wordt aangegeven door een 0.

De toetsingsgrootheid van de methode van rangcorrelatie is nu het aantal positief gecorreleerde tweetallen verminderd met het aantal negatief gecorreleerde, of wel de som van de getallen, die in tabel 1 in de kolom "correlatie" voorkomen.

De verdeling van $\underline{S}$ voor het geval dat $\underline{x}$ en $\underline{y}$ onafhankelijk zijn is bekend (zie § 2). De hypothese dat $\underline{x}$ en $\underline{y}$ onafhankelijk

---

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

Tabel 1

Berekening van S voor het voorbeeld

i	j	Correlatie
1	2	-1
1	3	+1
1	4	0
1	5	+1
1	6	+1
2	3	-1
2	4	-1
2	5	+1
2	6	+1
3	4	+1
3	5	+1
3	6	+1
4	5	0
4	6	0
5	6	0

$S = +5$

zijn, kan dus getoetst worden.

Is de hypothese van onafhankelijkheid niet vervuld, dan is de waarschijnlijkheid van grote positieve of grote negatieve waarden van S groter, dan wanneer dit wel het geval is. De kritieke zône is daarom van de vorm $|S| \geq S_0$, en bij éénzijdige toetsing van de vorm $S \geq S_0'$ (rechtszijdige toetsing) of $S \leq S_0''$ (linkszijdige toetsing).

2. Verdeling van S als x en y onafhankelijk zijn.

Als er noch bij de x_i noch bij de y_j gelijke waarden voorkomen kunnen wij gebruik te maken van exacte tabellen, die voorkomen in [1] pg 141 ($n = 4$ t/m,10) en in [2] (tables I and II, $n = 4$ t/m 40). Bovendien vindt men in [2] table III de kleinste waarden van S, waarvan de overschrijdingskansen onder de hypothese van onafhankelijkheid hoogstens gelijk zijn aan α voor $\alpha = 0{,}005$; $0{,}01$; $0{,}025$; $0{,}05$ en $0{,}10$ en $n = 4,5,6,\ldots,40$.

Als er bij de x_i óf bij de y_i, doch niet bij beide tweetallen of drietallen gelijken voorkomen, kan men voor $n \leq 10$ gebruik maken van de tabel van Sillitto [4].

Voor grote waarden van n is de verdeling van $\frac{S}{\sigma_S}$ (waarin σ_S de spreiding van S is, die uit een hieronder op te geven formule berekend kan worden) bij benadering normaal met gemiddelde 0 en spreiding 1. Hiervan kunnen we gebruik maken om de hypothese van onafhankelijkheid te toetsen in de gevallen waar de exacte verdeling niet getabelleerd is. Dit geschiedt dan, door in een tabel van de normale verdeling de overschrijdingskans op te zoeken, die behoort bij de gevonden waarde van $\frac{\underline{S}}{\sigma_{\underline{S}}}$.

Om $\sigma_{\underline{S}}$ te berekenen, nemen wij in de rij der waarnemingen x_i de gelijke waarnemingen in groepen bij elkaar. De aantallen waarnemingen in die groepen duiden wij aan met t_h, waarin $h = 1,2,\ldots,k_1$. Evenzo doet men in de rij der waarnemingen y_j, waar we de overeenkomstige aantallen aanduiden met u_ℓ waarin $\ell = 1,2,\ldots,k_2$. $\sigma_{\underline{S}}$ kan dan gevonden worden uit de volgende formule:

$$(1)\quad \sigma^2_{\underline{S}} = \frac{1}{18}\Big\{n(n-1)(2n+5) - \sum_{h=1}^{k_1} t_h(t_h-1)(2t_h+5) -$$

$$- \sum_{\ell=1}^{k_2} u_\ell(u_\ell-1)(2u_\ell+5)\Big\} +$$

$$+ \frac{1}{9n(n-1)(n-2)}\sum_{h=1}^{k_1} t_h(t_h-1)(t_h-2)\sum_{\ell=1}^{k_2} u_\ell(u_\ell-1)(u_\ell-2) +$$

$$+ \frac{1}{2n(n-1)} \sum_{h=1}^{k_1} t_h(t_h-1)\sum_{\ell=1}^{k_2} u_\ell(u_\ell-1).$$

In ons voorbeeld van § 1 komt in de rij x_i één tweetal gelijken(dus $k_1=1$ en $t_1=2$) en in de rij y_j één drietal gelijken ($k_2= 1$, $u_1 = 3$) voor. Dus geldt:

$$t_1(t_1- 1)(2t_1+ 5) = 2.1.9 = 18$$
$$u_1(u_1- 1)(2u_1+ 5) = 3.2.11 = 66$$
$$t_1(t_1- 1)(t_1- 2) = 0,\ u_1(u_1- 1)(u_1- 2) = 0$$
$$t_1(t_1- 1) = 2.1 = 2$$
$$u_1(u_1- 1) = 3.2 = 6$$
$$n(n-1)(2n+5) = 6.5.17 = 510$$
$$n(n-1) = 6.5 = 30$$

zodat:

$$\sigma^2_{\underline{S}} = \frac{1}{18}\{510-18-66\} + \frac{1}{60} \times 2 \times 6 = 23,87$$

en $\sigma_{\underline{S}} = 4,89$ is.

Als alle t_h en alle u_ℓ gelijk zijn aan 1 en er dus in geen van beide rijen gelijken voorkomen, gaat formule (2) over in:

$$(2)\qquad \sigma_{\underline{S}} = \sqrt{\frac{1}{18}\, n(n-1)(2n+5)}$$

Een tabel van deze functie voor $n = 40,41,\ldots,100$ vindt men in [2] (table IV).

3. Rangcorrelatiecoëfficiënt τ

Als maat voor de correlatie in de rij van waarnemingsparen $(x_1,y_1),\dots,(x_n,y_n)$ heeft Kendall de coëfficiënt τ gedefinieerd, die + 1 is als de volgorden der waarnemingen in beide rijen $x_1,\dots,x_n$ en $y_1,\dots,y_n$ volledig overeenstemmen en -1 is, als deze volgorden volkomen tegengesteld zijn. De definitie van τ is:

$$(3)\quad \tau = \frac{2S}{\left\{n(n-1)-\sum_{h=1}^{k_1} t_h(t_h-1)\right\}^{\frac{1}{2}}\left\{n(n-1)-\sum_{l=1}^{k_2} u_l(u_l-1)\right\}^{\frac{1}{2}}}.$$

Als er in geen van beide rijen gelijke waarnemingen voorkomen wordt deze formule:

$$(4)\quad \tau = \frac{2S}{n(n-1)} .$$

Literatuur:

[1] M.G. Kendall. Rank correlation Methods London 1948, Hoofdstuk 1.

[2] L. Kaarsemaker en A. van Wijngaarden. Tables for use in rank correlation . (1952) Report R 73 of the Computation Department of the Mathematical Centre.

[3] J. Hemelrijk. Kendall's rangcorrelatie-coëfficiënt . Hoofdstuk I der cursus "Parametervrije Methoden" Rapport S 59 (1951) Mathematisch Centrum, blz. 1-17.

[4] G.P. Sillitto. "The Distribution of Kendall's coefficient of rankcorrelation in rankings containing ties. Biometrica 34 (1947) p. 36-40.

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
A m s t e r d a m - O.

Statistische Afdeling
Rapport S 102 (M 17b)

Het combineren van onafhankelijke toetsen (aanvulling) [1].

In memorandum S 73 (M 17a) wordt een methode voor combinatie van onafhankelijke toetsen behandeld, waarbij het nodig is de overschrijdingskans van iedere toets te bepalen. In vele gevallen kan men de combinatie ook direct op de afzonderlijke toetsingsgrootheden baseren en dit verdient zelfs de voorkeur.

Wij beschouwen hier het geval, dat een bepaalde toets moet worden toegepast op een heterogeen materiaal. Dit materiaal wordt dan eerst verdeeld in h homogeen geachte groepen. Het aantal waarnemingen van de i^e groep zij n_i en de toetsingsgrootheid $\underline{t}_i$ [2]. Laat verder gegeven zijn, dat de verdeling van $\underline{t}_i$ onder de getoetste hypothese (voor de i^e groep aangeduid door H_i) voor grote n_i asymptotisch normaal [3] is, met bekende verwachting μ_i en bekende spreiding σ_i. Aan deze voorwaarden is o.a. voldaan, indien wij te doen hebben met toetsen van WILCOXON, rangcorrelatietoetsen van KENDALL of SPEARMAN, tekentoetsen enz.

Wij toetsen met al de hier te behandelen gecombineerde methoden de hypothese H, dat voor iedere groep de desbetreffende hypothese H_i geldt, terwijl de groepen onderling onafhankelijk zijn. De toetsen verschillen echter ten aanzien van de alternatieve (van H afwijkende) hypothesen waarvoor zij gevoelig [4]) zijn.

De meest gebruikelijke toetsingsgrootheden van gecombineerde toetsen zijn van de gedaante:

$$\underline{T} = \sum_{i=1}^{h} c_i(\underline{t}_i - \mu_i)$$

waarin de letters c_i ($i = 1, 2, \ldots, h$) constanten voorstellen, die voor ieder van de combinatiemethoden op een bepaalde wijze

---

1) Dit memorandum dient slechts ter oriëntatie en streeft niet naar volledigheid of volledige exactheid. Het is bedoeld als een aanvulling op Rapport S 73 (M 17a).

2) De onderstreping geeft aan dat een toetsingsgrootheid stochastisch is, d.w.z. een waarschijnlijkheidsverdeling bezit.

3) Dit houdt in dat $\underline{t}_i$ een waarschijnlijkheidsverdeling heeft, die als n_i toeneemt, steeds minder van een normale verdeling (verdeling van Gauss) afwijkt.

4) Een toets van hypothese H is gevoelig ten opzichte van een alternatieve hypothese H', als de kans dat H verworpen wordt, indien H' juist is, groot is.

gekozen worden. Onder de hypothese H zal $\underline{T}$ asymptotisch (voor grote h en/of grote n_i) normaal verdeeld zijn met verwachting 0 en spreiding $\sigma = \sqrt{\sum_{i=1}^{h} c_i^2 \sigma_i^2}$. De dubbele overschrijdingskans van een gevonden waarde T van $\underline{T}$ is dus bij benadering gelijk aan:

$$\frac{2}{\sqrt{2\pi}} \int_{\left|\frac{T}{\sigma}\right|}^{\infty} e^{-\frac{1}{2}x^2}\, dx$$

en kan bepaald worden met behulp van een tabel van de normale verdeling. Indien de dubbele overschrijdingskans kleiner is dan de onbetrouwbaarheidsdrempel α, zal men H verwerpen.

Wij geven hier 3 combinatiemethoden van dit type:

Methode 1: $c_1 = c_2 = \cdots\cdots = c_h = 1$

dus: $\underline{T} = \sum_{i=1}^{h} \underline{t}_i - \sum_{i=1}^{h} \mu_i$; $\sigma = \sqrt{\sum_{i=1}^{h} \sigma_i^2}$

Methode 2: $c_1 = \frac{1}{\sigma_1}$, $c_2 = \frac{1}{\sigma_2}$, ..., $c_h = \frac{1}{\sigma_h}$

dus: $\underline{T} = \sum_{i=1}^{h} \frac{\underline{t}_i - \mu_i}{\sigma_i}$; $\sigma = \sqrt{h}$

Methode 3: $c_1 = \frac{1}{n_1}$, $c_2 = \frac{1}{n_2}$, ..., $c_h = \frac{1}{n_h}$

dus: $\underline{T} = \sum_{i=1}^{h} \frac{\underline{t}_i - \mu_i}{n_i}$; $\sigma = \sqrt{\sum_{i=1}^{h} \left(\frac{\sigma_i}{n_i}\right)^2}$

Deze methoden zijn alleen gevoelig ten aanzien van alternatieve hypothesen volgens welke de grootheden $\underline{t}_i$ verdelingen hebben die, voor zover zij afwijken van de verdelingen onder de corresponderende hypothesen H_i, dit over het algemeen in dezelfde richting doen. Men zal dan methode 1 bij voorkeur toepassen als men aan de $\underline{t}_i$ met een kleine spreiding (in de regel zullen dat de $\underline{t}_i$ van kleine groepen zijn) een geringer gewicht wil toekennen dan aan de $\underline{t}_i$ met een grote spreiding. De methoden 2 en 3 zijn te gebruiken als men aan de verschillende groepen waarnemingen, ongeacht hun grootte, een ongeveer gelijke invloed op het resultaat wil toekennen. De keuze tussen deze twee methoden hangt verder van hier niet te behandelen theoretische overwegingen af (zie literatuur [1]).

Indien men verwacht dat mogelijke verschuivingen van de verdelingen der $\underline{t}_i$ in beide richtingen kunnen liggen, verdient het de voorkeur om gebruik te maken van de volgende toetsingsgrootheid:

$$\sum_{i=1}^{h} \left(\frac{\underline{t}_i - \mu_i}{\sigma_i}\right)^2 \qquad \text{(methode 4)}$$

Deze grootheid is onder de hypothese H asymptotisch verdeeld volgens een χ^2-verdeling met h vrijheidsgraden. De overschrijdingskans van een gevonden waarde van deze grootheid kan dus met behulp van een tabel van de χ^2-verdeling bepaald worden.

De toets, behandeld in memorandum S 73 (M 17a) par. 1, waarbij men het product van linkszijdige en product van alle rechtszijdige overschrijdingskansen bepaalt en het kleinste van deze twee producten gebruikt, heeft betrekking op dezelfde gevallen als de hier behandelde methoden 2 of 3, terwijl de methode, behandeld in S 73 (M 17a) par. 2, berustend op het product van de tweezijdige overschrijdingskansen, meer overeenkomt met methode 4. Men mag echter verwachten, dat, zo aan de asymptotische normaliteit der $\underline{t}_i$ voldaan is, de in dit memorandum behandelde methoden scherper zijn dan de toetsen behandeld in S 73 (M 17a).

Literatuur:

1 C.van Eeden, Methoden voor het vergelijken, toetsen en schatten van onbekende kansen, Rapport S 115 (M 45) van het Mathematisch Centrum (1953).

2 -----------, Trendtoets met behulp van rangcorrelatie, Memorandum S 73 (M 13a). (Voorbeeld van toepassing van methode 1.)

3 Dr J.Hemelrijk, Het combineren van onafhankelijke toetsen, Memorandum S 73 (M 17a).

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
Amsterdam-O.

Statistische Afdeling
Rapport S 102 (M 17b)

Het combineren van onafhankelijke toetsen (aanvulling) [1].

In memorandum S 73 (M 17a) wordt een methode voor combinatie van onafhankelijke toetsen behandeld, waarbij het nodig is de overschrijdingskans van iedere toets te bepalen. In vele gevallen kan men de combinatie ook direct op de afzonderlijke toetsingsgrootheden baseren en dit verdient zelfs de voorkeur.

Wij beschouwen hier het geval, dat een bepaalde toets moet worden toegepast op een heterogeen materiaal. Dit materiaal wordt dan eerst verdeeld in h homogeen geachte groepen. Het aantal waarnemingen van de i^e groep zij n_i en de toetsingsgrootheid $\underline{t}_i$ [2]. Laat verder gegeven zijn, dat de verdeling van $\underline{t}_i$ onder de getoetste hypothese (voor de i^e groep aangeduid door H_i) voor grote n_i asymptotisch normaal [3] is, met bekende verwachting μ_i en bekende spreiding σ_i. Aan deze voorwaarden is o.a. voldaan, indien wij te doen hebben met toetsen van WILCOXON, rangcorrelatietoetsen van KENDALL of SPEARMAN, tekentoetsen enz.

Wij toetsen met al de hier te behandelen gecombineerde methoden de hypothese H, dat voor iedere groep de desbetreffende hypothese H_i geldt, terwijl de groepen onderling onafhankelijk zijn. De toetsen verschillen echter ten aanzien van de alternatieve (van H afwijkende) hypothesen waarvoor zij gevoelig [4] zijn.

De meest gebruikelijke toetsingsgrootheden van gecombineerde toetsen zijn van de gedaante:

$$\underline{T} = \sum_{i=1}^{h} c_i(\underline{t}_i - \mu_i)$$

waarin de letters c_i ($i = 1, 2, \ldots, h$) constanten voorstellen, die voor ieder van de combinatiemethoden op een bepaalde wijze

---

1) Dit memorandum dient slechts ter oriëntatie en streeft niet naar volledigheid of volledige exactheid. Het is bedoeld als een aanvulling op Rapport S 73 (M 17a).

2) De onderstreping geeft aan dat een toetsingsgrootheid stochastisch is, d.w.z. een waarschijnlijkheidsverdeling bezit.

3) Dit houdt in dat $\underline{t}_i$ een waarschijnlijkheidsverdeling heeft, die als n_i toeneemt, steeds minder van een normale verdeling (verdeling van Gauss) afwijkt.

4) Een toets van hypothese H is gevoelig ten opzichte van een alternatieve hypothese H', als de kans dat H verworpen wordt, indien H' juist is, groot is.

gekozen worden. Onder de hypothese H zal $\underline{T}$ asymptotisch (voor grote h en/of grote n_i) normaal verdeeld zijn met verwachting 0 en spreiding $\sigma = \sqrt{\sum_{i=1}^{h} c_i^2 \sigma_i^2}$. De dubbele overschrijdingskans van een gevonden waarde T van $\underline{T}$ is dus bij benadering gelijk aan:

$$\frac{2}{\sqrt{2\pi}} \int_{\left|\frac{T}{\sigma}\right|}^{\infty} e^{-\frac{1}{2}x^2}\, dx$$

en kan bepaald worden met behulp van een tabel van de normale verdeling. Indien de dubbele overschrijdingskans kleiner is dan de onbetrouwbaarheidsdrempel α, zal men H verwerpen.

Wij geven hier 3 combinatiemethoden van dit type:

Methode 1: $c_1 = c_2 = \cdots\cdots = c_h = 1$

dus: $\underline{T} = \sum_{i=1}^{h} \underline{t}_i - \sum_{i=1}^{h} \mu_i$; $\sigma = \sqrt{\sum_{i=1}^{h} \sigma_i^2}$

Methode 2: $c_1 = \frac{1}{\sigma_1}, c_2 = \frac{1}{\sigma_2}, \ldots\ldots, c_h = \frac{1}{\sigma_h}$

dus: $\underline{T} = \sum_{i=1}^{h} \frac{\underline{t}_i - \mu_i}{\sigma_i}$; $\sigma = \sqrt{h}$

Methode 3: $c_1 = \frac{1}{n_1}, c_2 = \frac{1}{n_2}, \ldots\ldots, c_h = \frac{1}{n_h}$

dus: $\underline{T} = \sum_{i=1}^{h} \frac{\underline{t}_i - \mu_i}{n_i}$; $\sigma = \sqrt{\sum_{i=1}^{h} \left(\frac{\sigma_i}{n_i}\right)^2}$

Deze methoden zijn alleen gevoelig ten aanzien van alternatieve hypothesen volgens welke de grootheden $\underline{t}_i$ verdelingen hebben die, voor zover zij afwijken van de verdelingen onder de corresponderende hypothesen H_i, dit over het algemeen in dezelfde richting doen. Men zal dan methode 1 bij voorkeur toepassen als men aan de $\underline{t}_i$ met een kleine spreiding (in de regel zullen dat de $\underline{t}_i$ van kleine groepen zijn) een geringer gewicht wil toekennen dan aan de $\underline{t}_i$ met een grote spreiding. De methoden 2 en 3 zijn te gebruiken als men aan de verschillende groepen waarnemingen, ongeacht hun grootte, een ongeveer gelijke invloed op het resultaat wil toekennen. De keuze tussen deze twee methoden hangt verder van hier niet te behandelen theoretische overwegingen af (zie literatuur [1]).

Indien men verwacht dat mogelijke verschuivingen van de verdelingen der $\underline{t}_i$ in beide richtingen kunnen liggen, verdient het de voorkeur om gebruik te maken van de volgende toetsingsgrootheid:

$$\sum_{i=1}^{h} \left(\frac{\underline{t}_i - \mu_i}{\sigma_i}\right)^2 \qquad \text{(methode 4)}$$

Deze grootheid is onder de hypothese H asymptotisch verdeeld volgens een χ^2-verdeling met h vrijheidsgraden. De overschrijdingskans van een gevonden waarde van deze grootheid kan dus met behulp van een tabel van de χ^2-verdeling bepaald worden.

De toets, behandeld in memorandum S 73 (M 17a) par. 1, waarbij men het product van linkszijdige en product van alle rechtszijdige overschrijdingskansen bepaalt en het kleinste van deze twee producten gebruikt, heeft betrekking op dezelfde gevallen als de hier behandelde methoden 2 of 3, terwijl de methode, behandeld in S 73 (M 17a) par. 2, berustend op het product van de tweezijdige overschrijdingskansen, meer overeenkomt met methode 4. Men mag echter verwachten, dat, zo aan de asymptotische normaliteit der $\underline{t}_i$ voldaan is, de in dit memorandum behandelde methodenscherper zijn dan de toetsen behandeld in S 73 (M 17a).

Literatuur:

1 C.van Eeden, Methoden voor het vergelijken, toetsen en schatten van onbekende kansen, Rapport S 115 (M 45) van het Mathematisch Centrum (1953).

2 -----------, Trendtoets met behulp van rangcorrelatie, Memorandum S 73 (M 13a). (Voorbeeld van toepassing van methode 1.)

3 Dr J.Hemelrijk, Het combineren van onafhankelijke toetsen, Memorandum S 73 (M 17a).