

Statistische Afdeling
van het
Mathematisch Centrum,
Amsterdam.

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

Rapport S 89

De invloed van beenarbeid op het
histaminegehalte van het bloed .

door

Mej.C. van Eeden
en
Dr J.Hemelrijk Jr.

Juni 1952.

1. Inleiding:

Om de invloed van beenarbeid op het histaminegehalte van het bloed na te gaan werd bij 25 personen dit gehalte enige malen bepaald vóór, tijdens en na de arbeid.

Deze bepalingen werden verricht met gelijke tussenpozen van enkele minuten. Het aantal waarnemingen vóór de arbeid bij ieder der proefpersonen was gemiddeld ongeveer 3, tijdens arbeid gemiddeld ongeveer 5 en er na gemiddeld tussen de 2 en 3. Een reeks bepalingen bij één persoon ziet er dus b.v. als volgt uit:

$$g_1, g_2, g_3 \mid g_4, g_5, g_6, g_7 \mid g_8, g_9, g_{10}. \quad (1)$$

De arbeid is hier begonnen direct na de 3^e waarneming en opgehouden vlak vóór de achtste (niet vlak na de zevende) waarneming, die wij daarom ook nog tot de arbeidsperiode zullen rekenen; g_1 , g_2 en g_3 zijn dus de waarnemingen vóór de arbeid; g_4 t/m g_8 (niet t/m g_7) noemen wij de waarnemingen tijdens en g_9 en g_{10} de waarnemingen na arbeid.

Onder de 25 personen, waarbij waarnemingen verricht zijn, waren 15 mannen en 7 vrouwen. Van de overige drie personen was ons niet opgegeven of dit mannen of vrouwen waren.

2. Het waarnemingsmateriaal; een voorbehoud.

Van de 7 vrouwen, waarbij waarnemingen van het histaminegehalte verricht zijn, vertonen er 3 een sterk afwijkend gedrag, nl.: een hoog histaminegehalte vóór arbeid, terwijl het histaminegehalte tijdens arbeid lager is dan vóór arbeid, hetgeen bij de overige vrouwen niet het geval is. Daar de kleine aantallen geen definitieve conclusie omtrent dit verschijnsel toelaten, hebben wij de vrouwen verder buiten beschouwing gelaten, evenals de drie personen, waarvan het geslacht ons onbekend was.

Het waarnemingsmateriaal der mannen is in tabel I samengevat (zie blz.2).

De vraag, waarom het bij dit onderzoek gaat, is, of de beenarbeid het histaminegehalte beïnvloedt. Nu zal in het volgende blijken, dat het geven van een rechtstreeks antwoord op deze vraag door het optreden van een neveneffect niet mogelijk is. De oorzaak van dit neveneffect is onbekend; men kan deze zoeken in een psychologisch effect van de proef, maar ook de mogelijkheid van een afhankelijkheid van onbekende aard van de waarnemingen kan niet uitgesloten worden. De waarnemingstechniek vertoont nl., zoals in een vroeger rapport (S 75) besproken is, enkele zwakke punten, die het moeilijk maken de betrouwbaarheid der eindconclusies

te beoordelen. Dientengevolge kunnen wij deze conclusies slechts onder sterk voorbehoud geven. Het is niet uitgesloten, dat er bij dit onderzoek door één of meer aan de aandacht ontsnapte factoren schijn-effecten ontstaan zijn.

Tabel I

Waarnemingsmateriaal.

Persoon Nr.	Histaminegehalten
1	4,0;3,0 3,0;3,0;3,5;3,5;3,9;4,0;5,0;6,3;6,0 5,0;4,5.
5	6,1;6,4 5,1;7,7;6,7;12,7;7,7 6,5;7,0;6,4;7,5;7,6.
6	3,6;3,8;3,8 4,5;4,8;5,0;5,6 4,8;4,3;5,0.
8	4,0;5,2 4,5;3,5;3,0;7,7;4,0;4,1;3,8 5,0;5,6.
10	3,0;4,0;4,0;4,1 5,4;3,6;4,8;4,0 3,6;3,0;3,0.
11	3,0;2,5;2,5 2,5;3,0;4,0;4,1 5,1;4,6;5,0.
12	5,2;3,6;3,6;6,3 5,0;6,4;3,0;4,0 3,0;5,1;4,6;5,0.
13	5,5;5,5;4,0 6,2;6,6;6,4;3,2 3,4;3,0;3,5.
14	6,8;8,7;7,8;8,2;8,0 7,1;7,8;8,8;8,2 7,8;7,0;6,8.
15	6,0;7,8 8,3;9,9;10,2;11,6 10,0;10,3;9,4.
21	3,3;4,1;5,8 5,3;5,5 4,8;3,0.
22	4,3;4,5 3,9;4,1;4,5;4,2;4,1 3,6.
23	6,7;8,7;7,6;8,1 8,0;7,1;7,8;8,8;8,2 7,8;7,0;6,8.
24	5,0;5,0 3,0;4,1;5,1 4,3.
25	3,8;4,0;4,0 4,5;3,8;4,3;4,5 4,0;4,0;1,0.

Daar echter het verzamelen van het onderhavige waarnemingsmateriaal, hoe klein dit ook zij, veel tijd en moeite heeft gevergd, hebben wij toch getracht de statistische analyse zo uitgebreid mogelijk uit te voeren, om een zo groot mogelijk aantal aanwijzingen te verkrijgen, die bij een voortzetting van het onderzoek van belang kunnen zijn. Het uitvoeren van een zo uitgebreide analyse als de onderstaande op een waarnemingscomplex van kleine uitgebreidheid is in het algemeen niet aan te raden. De resultaten van toetsen, toegepast op verschillende onderdelen van hetzelfde materiaal zijn in het algemeen sterk van elkaar afhankelijk en van de aard van deze afhankelijkheid valt meestal weinig te zeggen. Het is dan ook niet onze bedoeling, dat de in dit rapport geformuleerde conclusies als eindconclusies worden beschouwd. Zij hebben bevestiging door verdere experimenten van node en geven aanwijzingen voor de weg, die bij later onderzoek kan worden gevolgd. Als zodanig echter kunnen zij van betekenis zijn.

3. Onderzoek naar verloop ¹⁾.

Wij onderzoeken allereerst, of het histaminegehalte van het bloed in de drie perioden vóór, tijdens en na arbeid apart constant is, of dat het een systematische toe- of afname met de tijd vertoont. Hiertoe is een toets tegen verloop (Engels: trend) toegepast, die op rangcorrelatiemethoden berust (zie bijlage S 73 (M 13a)).

De toets is toegepast op:

1. de waarnemingen vóór arbeid,
2. de laatste waarneming vóór arbeid tezamen met de waarnemingen tijdens arbeid (dus in (1): $g_3, \dots, g_8$),
3. de laatste waarneming tijdens arbeid tezamen met de waarnemingen na arbeid (dus in (1): g_8, g_9, g_{10}).

De resultaten, alleen voor mannen, vinden we in tabel II:

Tabel II ²⁾:

Onderzoek naar een stijging of daling van het histaminegehalte vóór (A), tijdens (B) en na (C) arbeid.

	$S_{\text{tst.}}$	σ	$\frac{S_{\text{tot}}}{\sigma}$	Overschrijdingskansen
A	+16	8,19	+1,95	0,05 +
B	+74	20,52	+3,61	< 0,001+
C	-24	11,28	-2,13	0,03 -

Uit tabel II zien we dus dat het histaminegehalte tijdens arbeid stijgt en na arbeid daalt, terwijl er een aanwijzing is voor een stijging vóór de arbeid.

4. Vergelijking van de gehalten in de drie perioden.

Ter bevestiging van de gevonden resultaten vergelijken wij vervolgens de gehalten in de drie perioden. Daartoe toetsen we de hypothese H_0 , dat er geen verschil is in het histaminegehalte vóór, tijdens en na arbeid. Om deze hypothese te toetsen zijn, voor ieder der personen, de waarnemingen vóór, tijdens en na arbeid onderling vergeleken met behulp van de toets van Wilcoxon (zie bijlage S 47 (M 7)).

We krijgen dan voor ieder der personen:

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

¹⁾ Zie ook bijlage S 47 (M 6).

²⁾ Het teken + bij een overschrijdingskans betekent een stijging van het histaminegehalte, en - het tegenovergestelde.

We noemen deze resp.: U_i , μ_{u_i} en $\sigma_{u_i}^2$, waarbij de index i het nummer van de proefpersoon is.

Indien H_0 juist is zal $U = \sum_i U_i$ bij benadering normaal verdeeld zijn met gemiddelde $\mu = \sum_i \mu_{u_i}$ en variantie $\sigma^2 = \sum_i \sigma_{u_i}^2$; $\frac{U-\mu}{\sigma}$ zal dus bij benadering normaal verdeeld zijn met gemiddelde 0 en spreiding 1. H_0 zal verworpen worden als de bij het experiment gevonden waarde van $\frac{U-\mu}{\sigma}$ sterk van nul afwijkt. De kritieke zône bestaat dus uit grote waarden van $\left| \frac{U-\mu}{\sigma} \right|$.

De resultaten vinden we in tabel III:

Tabel III ³⁾

Onderlinge vergelijking van de histaminegehalten vóór (A), tijdens (B) en na (C) arbeid.

	$U - \mu$	σ	$\frac{U - \mu}{\sigma}$	Overschrijdingskans
B-A	-25,5	11,96	2,13	0,03 +
C-B	20	11,73	1,73	0,08 -
C-A	1	8,81	0,11	0,91

We vinden hier dus dat het histaminegehalte tijdens arbeid hoger is dan dat vóór arbeid, terwijl er slechts een zéér zwakke aanwijzing is voor een verschil tussen de histamineniveaux tijdens en na arbeid. Verder is er geen verschil te ontdekken tussen het histaminegehalte vóór en dat na arbeid.

Opmerking: Dat we in tabel II tijdens arbeid een veel kleinere overschrijdingskans vinden dan vóór arbeid behoeft niet te betekenen, dat de stijging tijdens arbeid sterker is dan die vóór arbeid. Dit kan nl. ook zeer wel een gevolg zijn van het grotere aantal waarnemingen tijdens arbeid. Er is dus reden, dit nader te onderzoeken. Daarom is in de volgende paragraaf de stijging vóór arbeid vergeleken met de stijging tijdens arbeid.

³⁾ Het teken + bij een overschrijdingskans betekent dat het eerstgenoemde histaminegehalte hoger is dan het tweede (i.c. een stijging).

5. Onderzoek naar een invloed van het begin van de arbeid.

Voor ieder der personen zijn de volgende verschillen berekend:

a. Het verschil tussen de tweede en de eerste waarneming vóór arbeid. Dit verschil is een maat voor de stijging van het histaminegehalte aan het begin van de proef.

b. Het verschil tussen de laatste en op één na laatste waarneming vóór arbeid. Dit verschil is een maat voor de stijging direct vóór het begin van de arbeid.

c. Het verschil tussen de eerste waarneming na en de laatste waarneming vóór het begin van de arbeid. Dit verschil is een maat voor de stijging bij het begin van de arbeid.

Van deze grootheden zijn de volgende paren vergeleken:

c-a: de stijging aan het begin van de proef en die bij het begin van de arbeid,

c-b: de stijging direct vóór en direct na het begin van de arbeid.

Voor ieder der proefpersonen werden nu de verschillen c-a en c-b berekend, terwijl vervolgens, met behulp van de symmetrietoets T_2 (zie bijlage S 47(M 10)), de hypothese werd getoetst dat de twee vergeleken stijgingen voor ieder der proefpersonen dezelfde waarschijnlijkheidsverdeling bezitten. Indien deze hypothese verworpen wordt besluit men tot een systematisch verschil tussen de vergeleken grootheden. De resultaten staan vermeld in tabel IV.

Tabel IV ⁴⁾

Vergelijking van de stijging vóór arbeid met de stijging tijdens arbeid.

	n	r	u	v	Overschrijdingskansen
c-a	15	8	1	6	0,02 -
c-b	14	7	1	4	> 0,10 -

Uit tabel IV zien we dus dat de stijging van het histaminegehalte aan het begin van de proef sterker is dan de stijging aan het begin van de arbeidsperiode, terwijl er geen reden is, om een verschil aan te nemen tussen het verloop vlak vóór en vlak na het begin van de arbeid.

⁴⁾ Het teken - bij een overschrijdingskans betekent dat de eerste stijging zwakker is dan de tweede (i.c. een afneming van de stijging).

Indien de arbeid een verhogende of een verlagende invloed op het histaminegehalte uitoefende, zouden wij bij het begin van de arbeid een knik in het verloop van dit gehalte mogen verwachten. Hiervan valt niets te bespeuren, zodat wij geen enkele aanwijzing vinden voor een invloed van het begin van de arbeid.

Om te onderzoeken of er in het begin van de proef en in het begin van de arbeidsperiode een op zichzelf aantoonbare stijging optreedt, is de symmetrietoets T_2 ook nog toegepast op de verschillen a en c zelf. We vinden dan:

Tabel V ⁵⁾

Onderzoek naar een stijging van het histaminegehalte in het begin van de proef en in het begin van de arbeidsperiode.

	n	r	u	v	Overschrij- dingskans
a	13	7	2	1	$>0,10 +$
c	13	6	2	3	$>0,10 -$

We zien dus dat we, noch aan het begin van de proef, noch bij het begin van de arbeid een stijging van het histaminegehalte kunnen aantonen. Hieruit kunnen we dus concluderen dat de stijging van het histaminegehalte gedurende de perioden vóór en tijdens arbeid, waartoe wij uit tabel II en III konden besluiten, geleidelijk tot stand komt.

Uit het bovenstaande blijkt dus, dat vóór en tijdens de arbeid een geleidelijke stijging van het histaminegehalte optreedt en dat de stijging vóór arbeid sterker is dan de stijging aan het begin van de arbeidsperiode. Bij het begin van de arbeidsperiode vinden we geen enkele aanwijzing voor een invloed van de arbeid op het histaminegehalte van het bloed.

6. Onderzoek naar een invloed van het einde der arbeid.

We zullen nu onderzoeken of het ophouden met de arbeid het histaminegehalte doet dalen.

We vonden reeds:

1. een daling van het histaminegehalte na arbeid (laatste regel tabel II),
2. geen verschil tussen de histamineniveaux vóór en na arbeid (laatste regel tabel III).

⁵⁾ Zie voetnoot 2.

Om de stijging aan het eind van de arbeidsperiode en de daling daarna nader te onderzoeken hebben wij voor ieder der personen nog de volgende verschillen berekend:

d: Het verschil tussen de laatste en op één na laatste waarneming tijdens arbeid. Dit verschil is een maat voor de stijging aan het eind van de arbeidsperiode.

e: Het verschil tussen de laatste waarneming tijdens arbeid en de eerste waarneming van de daaropvolgende rustperiode. Dit verschil is een maat voor de daling direct na het einde van de arbeid.

Deze twee grootheden hebben wij vergeleken door de symmetrietoets T_2 toe te passen op de verschillen d-e (dit is dus het verschil van de eerste waarneming in de rustperiode en de op één na laatste van de arbeidsperiode). Om de daling direct na de arbeid afzonderlijk te onderzoeken hebben wij de symmetrietoets T_2 ook nog toegepast op de verschillen e zelf. De resultaten hiervan staan vermeld in tabel VI:

Tabel VI

Vergelijking van de stijging aan het eind van de arbeidsperiode met de daling direct na arbeid en onderzoek naar deze daling.

	n	r	u	v	Overschrijdingskans
d-e	14	7	1	2	0,04 -
e	15	8	2	1	0,05 -

Uit de eerste regel van tabel VI blijkt, dat de stijging vlak voor het ophouden met de arbeid minder groot is dan de daling vlak daarna; de tweede regel geeft een aanwijzing, dat er inderdaad een daling optreedt na het ophouden met de arbeid. Dit doet vermoeden, dat de arbeid het histaminegehalte doet stijgen. Het is echter ook mogelijk, dat de stijging, die in het begin van de proef optreedt en die niet aan de arbeid geweten kan worden, na enige tijd in een daling overgaat (door een natuurlijk herstel van het niveau b.v.) en dat het tijdstip van deze daling toevallig ongeveer samenvalt met het ophouden met de arbeid. Het is daarom van belang te onderzoeken, of er vlak voor het ophouden met de arbeid nog een stijging van het gehalte is. Is deze er niet meer, dan is er dus ook hier geen sprake meer van een knik in de curven, zodat de mogelijkheid van een toevallig samenvallen van een daling met het ophouden met de arbeid aan aannemelijkheid wint.

Om dit nader te onderzoeken hebben wij de symmetrietoets T_2 toegepast op de verschillen d zelf. We vinden dan een overschrijdingskans $> 0,10.$, d.w.z. we kunnen aan het einde van de arbeidsperiode geen stijging meer aantonen.

De aanwijzingen voor een invloed van de arbeid zijn dus ook hier tot nu toe zeer gering. Alvorens nog een andere analyse toe te passen, om een verder inzicht te verwerven, vatten wij de tot nu toe behaalde resultaten samen. Deze zijn:

1. een stijging van het histaminegehalte vóór arbeid en gedurende een deel van de periode tijdens arbeid,
2. een daling van het histaminegehalte na arbeid.

Wat dus zou kunnen wijzen op een invloed van arbeid op het histaminegehalte van het bloed is de stijging tijdens arbeid en de daling daarna. Daar echter ook vóór de arbeid reeds een stijging geconstateerd is en het histaminegehalte niet meer stijgt aan het eind van de arbeidsperiode kunnen wij de stijging tijdens arbeid niet zonder nader onderzoek aan die arbeid toeschrijven.

Indien er een contrôlegroep aanwezig was geweest van mensen, die geen arbeid verrichtten maar overigens in dezelfde omstandigheden verkeerden, dan was het mogelijk geweest het verloop van het histaminegehalte van deze personen te vergelijken met dat van degenen die wél arbeid verrichtten. Het ontbreken van zo'n contrôlegroep maakt dit onmogelijk.

Wel is het mogelijk een vergelijking te maken tussen die personen waarbij veel waarnemingen zijn verricht (dus waarbij de proef lang heeft geduurd) en diegenen waarbij weinig waarnemingen zijn verricht.

Het aantal waarnemingen tellen we hier vanaf het begin van de proef, niet vanaf het begin van de arbeid. Bij het begin van de arbeid treedt nl. geen merkbare wijziging op in het verloop van het histaminegehalte en er is dus geen reden om dit punt als beginpunt te nemen.

Als twee groepen personen nemen we degenen waarbij vanaf het begin van de proef zeven of minder waarnemingen zijn verricht en degenen waarbij vanaf het begin van de proef meer dan zeven waarnemingen zijn verricht. Dit is nl. een punt waar velen ophouden met de arbeid (zie het overzicht van het waarnemingsmateriaal in § 2). Voor ieder der personen is nu nagegaan of het verschil tussen de laatste en op een na laatste waarneming tijdens arbeid positief of negatief is; dit is dus het in deze paragraaf reeds eerder genoemde verschil d. We vinden dan de volgende aantallen personen:

Tabel VII

Onderscheiding der proefpersonen naar korte of lange waarnemingsreeksen vanaf het begin der proef en naar een stijging of daling aan het eind der arbeid.

	+	-	Totaal
lang	1	5	6
kort	6	3	9
Totaal	7	8	15

Met de methode der dubbele dichotomie (zie bijlage S 53 (M 23)) is nu de hypothese getoetst dat de onderscheiding naar korte en lange waarnemingsreeksen onafhankelijk is van de onderscheiding naar een stijging of daling aan het eind van de arbeidsperiode. We willen deze hypothese toetsen tegen de alternatieve hypothese dat de kans op een daling aan het eind der arbeidsperiode bij de lange waarnemingsreeksen groter is dan bij korte waarnemingsreeksen. Daar het omgekeerde effect onaannemelijk is, toetsen wij dus ééNZijdig en wij vinden dan een overschrijdingskans 0,08. Dit is dus een zwakke aanwijzing, dat, als men lang met de arbeid doorgaat, de daling van het histaminegehalte optreedt vóór het ophouden met de arbeid, dus dat dit ophouden niet of niet uitsluitend de oorzaak van de daling is.

De vraag is nu nog hoe het verloop van het histaminegehalte is op het moment van het ophouden van de arbeid, bij de personen met weinig waarnemingen, in vergelijking met het verloop, op datzelfde moment bij de personen met veel waarnemingen. We onderscheiden de personen dus weer naar korte of lange waarnemingsreeksen zoals boven is aangegeven; de tweede onderscheiding is hier: een positief of negatief verschil tussen de 8^e en 7^e waarneming. Bij degenen met lange waarnemingsreeksen is dit dus het verschil tussen twee waarnemingen tijdens arbeid. Bij degenen met 7 waarnemingen voor en tijdens arbeid tezamen is dit het verschil tussen de eerste waarneming na en de laatste waarneming tijdens arbeid. Bij de overige personen (dus degenen met minder dan 7 waarnemingen) nemen we eveneens het verschil tussen eerste waarneming na en laatste tijdens arbeid, daar voor deze personen de arbeid bij de 7^e waarneming reeds is opgehouden.

We vinden dan de volgende aantallen personen:

Tabel VIII

Onderscheiding van de proefpersonen naar korte of lange waarnemingsreeksen vanaf het begin van de proef en naar een positief of negatief verschil tussen de 8^e en 7^e waarneming:

	+	-	Totaal
lang	5	1	6
kort	2	7	9
Totaal	7	8	15

Met de methode der dubbele dichotomie is weer de hypothese getoetst dat bovengenoemde onderscheidingen onafhankelijk zijn, tegenover de alternatieve hypothese dat de kans op een stijging bij degenen met lange waarnemingsreeksen groter is dan bij degenen met korte waarnemingsreeksen. We vinden dan een (eenzijdige) overschrijdingskans van 0,035.

Dit is dus een aanwijzing, dat degenen die na de 7^e waarneming nog met de arbeid doorgaan vaker nog een stijging van het histaminegehalte vertonen dan degenen die na de 7^e waarneming ophouden (of reeds eerder opgehouden zijn).

Dit is dus een aanwijzing vóór een verhogende invloed van arbeid op het histaminegehalte.

Indien wij de lengte der waarnemingsreeksen niet tellen vanaf het begin der proef, maar vanaf het begin van de arbeid en dan een scheiding maken tussen korte en lange waarnemingsreeksen dan vinden wij bij nogmaals toepassen van de methode der dubbele dichotomie resp. een (eenzijdige) overschrijdingskans van 0,02 en 0,76.

De eerste overschrijdingskans versterkt de bovengegeven conclusie, dat bij lange waarnemingsreeksen de daling reeds optreedt vóór het einde van de arbeid. In het tweede geval (overschrijdingskans 0,76) vinden we wellicht een grote overschrijdingskans, omdat door het verschuiven van de grens tussen kort en lang, deze grens zo dicht bij het einde der arbeid komt te liggen dat hij praktisch alle personen reeds een daling optreedt.

7. Het verloop na het ophouden met de arbeid.

We willen tenslotte nog het verloop van de daling na arbeid nader onderzoeken. Hiertoe hebben wij de in § 3 genoemde toets tegen verloop nog toegepast op de waarnemingen na arbeid (dus zonder de laatste waarneming tijdens arbeid). We vinden dan:

Tabel IX ⁶⁾

Onderzoek naar de daling van het histaminegehalte na arbeid
(zonder de laatste waarneming tijdens arbeid):

S_{tot}	σ	$\frac{S_{\text{tot}}}{\sigma}$	Overschrijdingskans
-3	7,60	0,39	0,70 -

We vinden dus, door het weglaten van de laatste waarneming tijdens arbeid, een stijging van de overschrijdingskans van 0,03 tot 0,70. Dit wijst er op, dat de daling van het histaminegehalte na arbeid grotendeels plaats vindt direct na het ophouden met de arbeid. Hierbij moet nog wel opgemerkt worden dat een deel van de stijging van de overschrijdingskans een gevolg is van de vermindering van het aantal waarnemingen. Daarom hebben wij nog onderzocht of er een verschil is tussen de eerste en de laatste waarnemingen na de arbeid. Passen wij op het verschil tussen deze twee waarnemingen de symmetrietoets toe dan vinden wij een overschrijdingskans $> 0,10$ (-). We vinden hier dus geen daling meer, hetgeen de bovengegeven conclusie versterkt.

8. Samenvatting der conclusies.

We kunnen de conclusies als volgt samenvatten:

1. Het histaminegehalte vertoont vóóren tijdens arbeid een stijging.
2. Er is geen aanwijzing, dat het begin van de arbeid de reeds vóór de arbeid bestaande stijging beïnvloedt.
3. De duur van de proef beïnvloedt het gedrag van het histaminegehalte aan het einde van de arbeidsperiode. Bij de lange waarnemingsreeksen is de kans, dat er reeds voor het einde van de arbeid een daling optreedt groter dan bij de korte.
4. Er is ook een aanwijzing, dat het ophouden met de arbeid de kans op daling van het histaminegehalte vergroot.
5. Er is een aanwijzing, dat de daling, die bij het ophouden van de arbeid optreedt, een steile daling is, d.w.z. een daling, die in korte tijd vrijwel geheel tot stand komt. In hoeverre deze daling een gevolg van dit ophouden met de arbeid is, valt uit het experiment niet op te maken.

⁶⁾ Zie voetnoot 2.

Opmerkingen: Onder het aan het eind van § 2 genoemde voorbehoud deze conclusies beschouwende, zouden wij deze als volgt kunnen interpreteren. Er is een storende invloed geweest, waarvan de aard onbekend is, die een eventueel aanwezige verhogende werking van de arbeid op het histaminegehalte vrijwel geheel heeft overschaduwd. De proefopzet maakte het ons onmogelijk deze storende invloed te elimineren. In het bijzonder vormt het ontbreken van een contrôlegroep van proefpersonen, die geen arbeid verrichten, een ernstige moeilijkheid. Daar er echter toch wel enige aanwijzing gevonden is voor een invloed van de arbeid, is het niet onmogelijk, dat deze bij verder experimenteel onderzoek aangetoond zou kunnen worden. Het is echter wel zeer gewenst, dat de oorzaak van het storende effect opgespoord wordt, of dat dit effect door een andere proefopzet wordt geëlimineerd, indien dit althans mogelijk blijkt. Het aantal mogelijkheden van een gewijzigde proefopzet is te groot, om hier uitvoerig besproken te worden. Wij wijzen daarom slechts op enkele punten. Ten eerste valt het gebruik van de in rapport S 75 beschreven gewijzigde methode te overwegen. Ten tweede is een voorafgaand onderzoek van een contrôlegroep gewenst. Ten derde is het wenselijk de lengten der proefperiodes minder ongelijk te doen zijn, dan bij het onderhavige materiaal het geval was. Een splitsing van de groep van proefpersonen in twee of drie groepen, met ieder een vaste lengte van de verschillende periodes, zou de statistische analyse zeker ten goede komen.

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 ³⁾. 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 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.

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 toetsingsgrootte $\underline{U}$ 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 $\frac{1}{2}$ 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 toetsingsgrootte $\underline{U}$ 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 toetsingsgrootte $\underline{U}$ onder de hypothese H_0 voor grote waarden van n en m (beide ≥ 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 $t_1, \dots, t_k$ de aantallen waarnemingen in deze groepen van gelijken, dan worden het gemiddelde μ en de variantie σ^2 van de toetsingsgrootte $\underline{U}$ gegeven door

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

en

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

De grootte $\mu(\underline{U})$ is dus onafhankelijk van de waarden vast. Indien de hypothese H_0 niet vervuld is, zal de grootte $\underline{U}$ 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 verworpt indien de gevonden waarde $\underline{U}$ van $\underline{U}$ te sterk van μ afwijkt, d.w.z. als

$$\frac{|\underline{U} - \mu|}{\sigma} > \xi_{\alpha} \quad 2)$$

waarin α de onbetrouwbaarheidsdrempel is en ξ_{α} volgt uit

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

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

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

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

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 overschrijdskans 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 overschrijdskansen, 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.

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
A m s t e r d a m 0.
Statistische Afdeling.

S 47 (M 10)

Symmetrietoets¹⁾.

Hypothese H_0 : de waarnemingen $z_1, \dots, z_n$ zijn afkomstig van n onafhankelijk verdeelde stochastische grootheden, die alle symmetrisch ten opzichte van 0 verdeeld zijn²⁾. Van deze toets bestaan meerdere versies $T_1, \dots, T_2$. We bespreken eerst T_1 en T_2 .

Toetsingsgrootheden. Deze worden als volgt uit $z_1, \dots, z_n$ afgeleid:

1e. de waarnemingen, die gelijk aan 0 zijn worden weggelaten.
Stel er blijven over: $z_1, \dots, z_n$.

2e. Hieruit worden de positieve waarnemingen gezocht. Stel dit zijn $x_1, \dots, x_{n_1}$, dus n_1 in aantal.

3e. De overblijvende negatieve waarnemingen worden van teken veranderd, zodat zij ook positief worden. Stel dit zijn dan $y_1, \dots, y_{n_2}$.

4e. De grootheden $x_1, \dots, x_{n_1}$ en $y_1, \dots, y_{n_2}$ worden, door elkaar, in afdalende grootte-volgorde opgeschreven. Stel dit geeft: $w_1, \dots, w_n$. (Komen er gelijken voor, dan worden deze in willekeurige volgorde geplaatst.)

5e. De groep waarden $w_1, \dots, w_n$ wordt verdeeld in twee groepen $w_1, w_2, \dots, w_r$ en $w_{r+1}, \dots, w_n$, waarbij $w_r \neq w_{r+1}$ is en r zo dicht mogelijk bij de waarde $\frac{1}{2}n$ genomen wordt. Is n even, dan wordt $r = \frac{1}{2}n$, indien althans $w_{\frac{1}{2}n} \neq w_{\frac{1}{2}n+1}$ is. Zijn er twee mogelijk keuzen voor r , beide op gelijke afstand van $\frac{1}{2}n$, dan nemen wij $r > \frac{1}{2}n$. Is b.v. n oneven en $w_{\frac{1}{2}n-\frac{1}{2}} \neq w_{\frac{1}{2}n+\frac{1}{2}}$, dan nemen wij $r = \frac{1}{2}n + \frac{1}{2}$. Wij geven de waarden $w_1, \dots, w_r$ aan als groep A (die dus r elementen bevat) en de overigen als groep B. Alle elementen van A zijn dus groter dan ieder element van B³⁾.

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

2) Zetten wij hier a in plaats van 0, dan geldt H_0 voor $x_1 - a, \dots, x_n - a$.

3) In de oorspronkelijke publicaties over deze toets (zie de literatuurverwijzingen aan het einde van dit memorandum) is een enigzins minder algemene definitie van r gegeven.

Alle stellingen blijven echter gelden, indien de hier gegeven definitie gebruikt wordt.

6e. Het aantal waarden van $x_1, \dots, x_n$ die in A voorkomen noemen wij u .

De toetsingsgrootheden zijn n_1 en u , r is een hulpgrootheid.

V.B. 22 waarden z_i : 7,4/6,3/3,6/3,5/3,4/2,9/2,5/1,1/0/0/-1,3/-2,5/-3,2/-4,6/-4,6/-4,6/-4,8/-5,3/-7,0/-7,9/-8,0/-8,7.

$$\therefore r = 11, \quad n_1 = 8 \\ u = 2$$

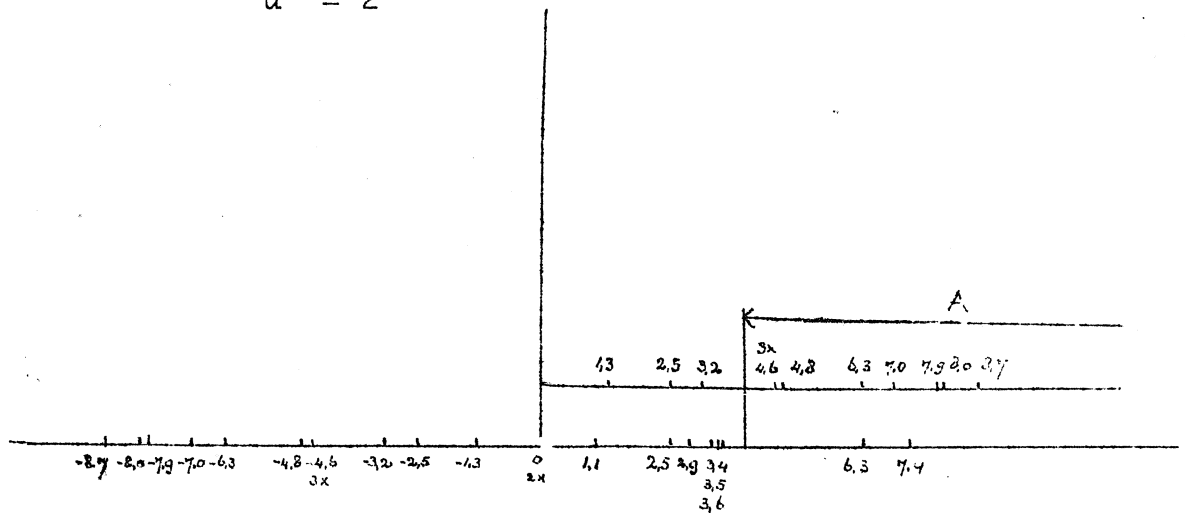


fig. 1

Kritieke zônes. Waarden van n_1 , die dicht bij 0 of dicht bij n liggen, zullen, als H_0 juist is weinig, maar als H_0 onjuist is vaker, voorkomen. Grote resp. kleine waarden van u zullen eveneens, als H_0 juist is weinig voorkomen. Hierop berust de keuze van de bij T_1 en T_2 behorende kritieke zônes Z_1 resp. Z_2 . Z_1 bevat grote en kleine waarden van n_1 en grote en kleine waarden van u , terwijl Z_2 bij grote waarden van n_1 in hoofdzaak grote waarden van u en bij kleine waarden van n_1 in hoofdzaak kleine waarden van u bevat. T_1 leidt bij voldoende grote n vrijwel steeds tot verwerping als de hypothese niet is vervuld. T_2 leidt echter alleen tot verwerping van H_0 als er veel positieve (resp, negatieve) waarden zijn, die verder van 0 verwijderd liggen dan de negatieve (resp. positieve). In sommige gevallen is het juist van belang om deze laatste afwijkingen van H_0 te ontdekken. In dat geval gebruikt men T_2 liever dan T_1 . In fig. 2 is een schematisch voorbeeld gegeven van een serie waarnemingen, waarbij het aantal positieve groter is dan het aantal negatieve, terwijl deze positieve dichter bij 0 liggen dan de negatieve, zodat T_2 niet tot verwerping leidt.



fig. 2

Van T_1 en T_2 bestaan ook éézijdige versies, waarvan de beschrijving te ver zou voeren.

T'_1 en T'_2 .

Toetsingsgrootheden.

1e, 2e en 3e als boven. Eerste toetsingsgrootheid: n_1 .

4e: op $x_1, \dots, x_{n_1}$ en $y_1, \dots, y_{n_2}$ wordt de toets van Wilcoxon toegepast (vgl. S 47 (M 8)). De toetsingsgrootheden zijn n_1 en de U van deze toets van Wilcoxon.

Kritieke zônes. Overwegingen analoog aan die voor T_1 en T_2 (met U in plaats van u) leiden tot analoge kritieke zônes Z'_1 en Z'_2 , behorend bij T'_1 en T'_2 .

Opmerkingen: T_1 en T_2 zijn bijzonder geschikt voor een niet te groot aantal waarnemingen. Zij gelden ook voor niet continue verdelingen. T'_1 en T'_2 zijn alleen geschikt, als er geen of weinig paren (x_i, y_j) met $x_i = y_j$ zijn. Voor grote aantallen zijn T'_1 en T'_2 geschikter dan T_1 en T_2 . Er is ook een versie voor grote aantallen (T''_1 en T''_2), die geheel analoog is met T'_1 en T'_2 met dien verstande, dat u in plaats van U wordt gebruikt (vgl. b.v. [2], blz. 77, § 6.4.5).

Litteratuur:

- [1] J. Hemelrijk, A family of parameterfree tests for symmetry with respect to a given point, I, II. Proceedings van de Kon. Ned. Ak. v. Wet. 53 (1950), p. 945-955. Indagationes Mathematicae 12 (1950), p. 340-350.
- [2] - " - , Symmetrietoetsen, Diss., Den Haag 1950, Excelsior.

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

Statistische Afdeling.

S 73 (M 13a)

Trend-toets met behulp van rangcorrelatie¹⁾

Men kan de methode der rangcorrelatie toepassen om na te gaan of er in een reeks waarnemingen van een stochastische grootte y een trend, d.w.z. een stijgend of dalend verloop aanwezig is. Men maakt daarbij gebruik van het feit, dat bij een systeem van n waarnemingsparen (x_i, y_i) de grootte van $S^{(2)}$ en de verdeling van S onder de hypothese van onafhankelijkheid niet verandert als men de volgorde van de paren wijzigt. Men kan dus bij de berekening van S de volgorde der paren (x_i, y_i) zodanig kiezen, dat de rangnummers van bijvoorbeeld de x_i de rij 1, 2, 3, ..., n vormen. De hypothese van onafhankelijkheid komt dan overeen met de hypothese H_0 , dat alle mogelijke permutaties van de rangnummers der y_i even waarschijnlijk zijn.

Indien wij nu willen onderzoeken of er een trend aanwezig is in de rij waarnemingen $y_1, \dots, y_n$ van de stochastische grootte y , dan voegen wij de rij 1, 2, ..., n toe en bepalen vervolgens S zoals aangegeven in memorandum S 47 (M 13) en toetsen bovengenoemde hypothese H_0 . Onder deze hypothese heeft S dezelfde verdeling als de overeenkomstige grootte der rangcorrelatie-toets onder de hypothese van onafhankelijkheid; wij gebruiken hier ook dezelfde kritieke zônes. Indien S een positieve (resp. negatieve) waarde heeft behorende tot de kritieke zône spreken wij van een significant stijgende (resp. dalende) trend.

Voor literatuur over deze toets zie men [1].

Deze toets kan worden uitgebreid tot meerdere onafhankelijke reeksen waarnemingen, die niet alle even groot behoeven te zijn. Bij voorbeeld:

$y_{1,1}, y_{1,2}, \dots, y_{1,n_1}$

$y_{2,1}, y_{2,2}, \dots, y_{2,n_2}$

$y_{3,1}, y_{3,2}, \dots, y_{3,n_3}$

$\vdots$

$y_{m,1}, y_{m,2}, \dots, y_{m,n_m}$

1) Dit memorandum is een aanvulling op memorandum S 47 (M 13). Het is slechts bedoeld ter orientatie en streeft niet naar volledigheid of volledige exactheid.

2) Zie voor de definitie van S memorandum S 47 (M 13).

Wij bepalen nu de S voor het stelsel paren $(1, y_{11})(2, y_{12}) \dots (n_1, y_{1n_1})$ en noemen deze S_1 . Evenzo behandelen wij de tweede rij en verkrijgen zo S_2 etc.

Wij bepalen dan de som

$$S_{\text{tot}} = S_1 + S_2 + \dots + S_m.$$

Eveneens berekenen wij bij de eerste rij de spreiding σ_1 , volgens formule (1) of formule (2) van memorandum S 47 (M 13). Evenzo $\sigma_2, \sigma_3, \dots, \sigma_m$.

Wij weten nu, dat onder de hypothese H_0 voor rij 1 $\underline{S}_1$ bij benadering normaal verdeeld is met gemiddelde 0 en spreiding σ_1 , indien H_0 geldt voor rij 2 is $\underline{S}_2$ bij benadering normaal verdeeld met gemiddelde 0 en spreiding σ_2 etc. Indien hypothese H_0 geldt voor alle rijen, dus indien in alle rijen alle permutaties der rangnummers van de waarnemingen even waarschijnlijk zijn, en indien deze rijen bovendien onafhankelijk zijn, dan geldt dat $\underline{S}_{\text{tot}}$ bij benadering normaal verdeeld is met gemiddelde 0 en spreiding:

$$\sigma_{\text{tot}} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_m^2}.$$

Wij kiezen nu als kritieke z one weer $|S_{\text{tot}}| \geq S_0$ bij tweezijdige toetsing, $S_{\text{tot}} \geq S'_0$ bij rechtszijdige en $S_{\text{tot}} \leq S''_0$ bij linkszijdige toetsing. Voor literatuur over deze toets zie men [2].

Literatuur: [1] H.B.Mann: Non parametric tests against trend, *Econometrica* 13 (1945) blz. 245-259.

[2] G.Elfvig en J.H.Whitlock: A simple trend-test with application to erythrocyte data, *Biometrics* 6 (1950) Blz. 282-288.

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
A m s t e r d a m - 0.

Statistische Afdeling.

S 53 (M 23).

Toetsing van de hypothese $p_1 = p_2$ met behulp
van een 2 x 2-tabel¹⁾.

Wij beschouwen twee reeksen van onafhankelijke experimenten, waarbij ieder experiment van de ene reeks één van de twee resultaten A of $\bar{A}$ (non-A) heeft en ieder experiment van de tweede reeks één van de beide resultaten B of $\bar{B}$ (hierbij kan $A=B$ zijn). Daarbij wordt ondersteld, dat bij ieder der experimenten van de ene reeks de kans op A gelijk aan p_1 (en dus de kans op $\bar{A}$ gelijk aan $1-p_1$) is en bij ieder der experimenten van de tweede reeks de kans op B gelijk aan p_2 (en dus de kans op $\bar{B}$ gelijk aan $1-p_2$). De te toetsen hypothese luidt nu:

$$H_0: p_1 = p_2.$$

Indien de eerste reeks uit n en de tweede reeks uit m waarnemingen bestaat, waaronder n_1 (resp. m_1) maal A (resp. B) voorkomt, kunnen deze gegevens in de volgende 2 x 2-tabel worden samengevat:

	A resp. B	$\bar{A}$ resp. $\bar{B}$	totaal
eerste reeks	n_1	$n-n_1$	n
tweede reeks	m_1	$m-m_1$	m
totaal	r	$n+m-r$	$n+m$

Als toetsingsgrootheid wordt n_1 , het aantal malen A in de eerste reeks waarnemingen, gebruikt. Indien H_0 juist is bezit deze grootheid onder de voorwaarde, dat r de bij het experiment gevonden waarde aanneemt, de volgende waarschijnlijkheidsverdeling: de kans, dat een bepaalde waarde n_1 aangenomen wordt, is gelijk aan:

$$\frac{\binom{n}{n_1} \binom{m}{m_1}}{\binom{n+m}{r}}$$

Als kritieke zône worden de waarden van n_1 met de kleinste waarschijnlijkheden bijeengezocht, tot de gekozen betrouwbaarheidsdrempel het toevoegen van een nieuwe waarde verhindert (bij éénzijdige toetsing bestaat de kritieke zône uitsluitend uit grote of uitsluitend uit kleine waarden van n_1).

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

De overschrijdingskans, behorende bij de gevonden waarde van n_1 , is gedefiniëerd als de som van alle waarschijnlijkheden van bovenstaande verdeling, die hoogstens gelijk aan de waarschijnlijkheid van de gevonden waarde zijn (bij ééNZzijdige toetsing echter gelijk aan de som van de waarschijnlijkheden van alle waarden die groter of gelijk aan de gevondene, of van alle waarden, die kleiner of gelijk aan de gevondene zijn). Deze exacte toetsingsmethode voor H_0 is afkomstig van R.A. Fisher.

Indien n en m zo groot zijn, dat deze exacte berekening te omslachtig wordt, maakt men gebruik van de volgende benadering:

Gemiddelde en spreiding van de grootheid n_1 zijn (indien H_0 juist is):

$$\frac{nr}{n+m} \quad \text{resp.} \quad \sqrt{\frac{n m r s}{(n+m)^2 (n+m-1)}} \quad (s = n + m - r)$$

Men gebruikt dan in plaats van de exacte waarschijnlijkheidsverdeling van n_1 de normale verdeling met hetzelfde gemiddelde en dezelfde spreiding en in plaats van de gevonden waarde van n_1 neemt men het getal, dat $\frac{1}{2}$ dichter bij het gemiddelde ligt dan deze gevonden waarde (dit laatste is de z.g. "continuïteitscorrectie", die bij toenemende n en m weldra verwaarloosd kan worden). Met behulp van de benadering gaat men dan verder te werk als boven beschreven, daarbij gebruik makende van een tabel van de normale verdeling.

Litteratuur:

R.A.Fisher, Statistical Methods for Research Workers, London 1948, p. 96. Opmerking: Fisher gebruikt hier de ééNZzijdige overschrijdingskans.

J.Hemelrijk, Waarschijnlijkheidsrekening en Statistiek, Vacantiecursus Mathematisch Centrum, Amsterdam 1950, § 4.