MATHEMATICSCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
STATISTISCHE AFDELING

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

Rapport S 150

Statistische beoordeling van celtellingen in vaginaal-uitstrijkjes

door

Jkvr. H.D.Sandberg
en
A.Benard

1954

1. Inleiding

Bij de studie van hormonale verhoudingen bij de vrouw maakt men onder andere vaak gebruik van het vaginaal-uitstrijkje. Hierbij wordt met een spatel of een pipet wat schedevocht, dat steeds met afschilference (plaveisel-epitheel) cellen beladen is, op een objectglaasje uitgestreken (gefixeerd), en gekleurd volgens de methode van HARRIS-SHORR of volgens die van PAPANICOLAOU.

Hierbij is het dan mogelijk een onderscheid te maken, enerzijds tussen eosinophiele en basophiele cellen en anderzijds tussen karyopycnotische en niet-karyopycnotische cellen.

1.a. Het waarnemingsmateriaal

Om de betrouwbaarheid van deze tellingen te onderzoeken werden de aantallen eosinophiele en karyopycnotische epitheelcellen bepaald bij herhaalde microscopische tellingen, van 100 cellen elk, uitgevoerd op uitstrijkjes, op verschillende data bij éénzelfde patiente genomen, en gekleurd volgens twee verschillende technieken, nl. volgens Harris-Shorr en volgens Papanicolaou.

De numerieke uitkomsten staan vermeld in tabel I (blz. 2).

1.b. De vragen.

De ons hierbij voorgelegde vragen zijn:

1) In hoeverre behoren de verschillende tellingen (van 100 cellen) tot éénzelfde verdeling en zo ja tot welke,
   1.1) op éénzelfde uitstrijkje, waarbij de nodige voorzorgen, om nooit opnieuw dezelfde cellen te tellen in acht genomen worden;
   1.2) op twee of meer verschillende uitstrijkjes, die op hetzelfde ogenblik bij dezelfde patiente werden gepreleveerd;
   1.3) op verschillende uitstrijkjes, op verschillende momenten van de cyclus bij dezelfde patiente gepreleveerd;
   1.4) op verschillende uitstrijkjes van verschillende patienten.
2) Welke betekenis kan er gehecht worden aan het "percentage" gevonden na
   2.1) één telling, zoals in de practijk het geval is;
   2.2) meerdere tellingen.
3) Is er correlatie tussen de aantallen eosinophiele en karyopycnotische cellen.
4) Is de techniek van het uitstrijken en kleuren van invloed op de aard der verdeling en op de spreiding.

N.B. Het waarnemingsmateriaal geeft slechts de mogelijkheid om een gedeelte der gestelde vragen te onderzoeken. Aanwijzingen voor verder onderzoek zullen aan het einde van dit rapport (par. 6.3) worden gegeven.

Tabel I 1)

Het waarnemingsmateriaal

4.3.54	Telling nr		1	2	3	4	5	6	7	8	9	10
	1) H-S:	eos	12	16	11	9						
		kar	1	3	0	0						
	2) Pap:	eos	16	15	12	10	10	12	13	15	16	17
		kar	3	2	1	2	2	0	0	2	2	1
23.3.54	Telling nr		1	2	3	4	5	6	7	8	9	10
	1) H-S:	eos	6	9	5	5	6	6	5	4	6	5
		kar	1	2	1	0	2	1	1	1	3	2
	2) Pap:	eos	8	7	10	10	8	8	9	16	14	12
		kar	3	3	4	4	1	2	1	1	8	9
3.4.54	Telling nr		1	2	3	4	5	6	7	8	9	10
	1) H-S:	eos	12	12	7	10	12	7	12	8	13	8
		kar	5	2	1	1	5	1	2	5	5	3
	2) Pap:	eos	20	12	13	17	10	14	7	15	14	16
		kar	3	4	5	6	3	1	3	7	4	6
6.4.54	Telling nr		1	2	3	4	5	6	7	8	9	10
	1) H-S:	eos	12	18	14	9	20	20	12	18	19	15
		kar	1	5	5	4	6	6	5	5	10	5
	2) Pap:	eos	18	16	12	15	16	21	16	19	17	10
		kar	3	3	5	1	2	3	3	3	1	2
10.4.54	Telling nr		1	2	3	4	5	6	7	8	9	10
	1) H-S:	eos	22	15	13	18	18	15	10	18	17	13
		kar	2	1	2	1	2	3	1	2	1	3
	2) Pap:	eos	7	5	10	5	7	3	2	4	5	5
		kar	0	0	1	0	0	1	0	0	0	0
13.4.54	Telling nr		1	2	3	4	5	6	7	8	9	10
	H-S:	eos	12	10	5	20	9	9	9	13	10	14
		kar	1	1	1	2	0	2	1	0	3	2
15.4.54	Telling nr		1	2	3	4	5	6	7	8	9	10
	1) H-S:	eos	12	17	17	15	15	18	18	15	17	17
		kar	1	1	1	2	2	2	0	2	2	2
	2) Pap:	eos	21	15	12	17	10	13	13	21	10	11
		kar	2	1	1	3	2	1	1	1	0	0

---

1) Wij gebruiken de volgende afkortingen:
H-S voor Harris-Shorr
Pap voor Papanicolaou
eos voor eosinophiele cellen
kar voor karyopycnotische cellen

2. Vergelijking tussen verschillende tellingen op één uitstrijkje

2.1. Bij de eerste vraag, in hoeverre verschillende tellingen van de cellen van één uitstrijkje al dan niet tot een zelfde verdeling behoren, hebben wij de waarnemingen gerangschikt als aangegeven in tabel II.

Tabel II

Voorbeeld van een 2 x k-tabel voor vergelijking van k tellingen

	Telling nr	1	2	3	4	tot.
4.3.54	H-S: eos	12	16	11	9	48
	niet eos	88	84	89	91	352
	totaal	100	100	100	100	400

Hierop is de toets voor de hypothese $p_1 = p_2 = \dots = p_k$ met behulp van een 2 x k -tabel – zoals beschreven in het memorandum S 145 (M 23a), zie bijlage – uitgevoerd. Deze bewerking werd voor iedere regel herhaald. De resultaten zijn vermeld in de tabellen III en IV.

Tabel III

Vergelijking tussen de eos-tellingen der verschillende dagen

Datum	H-S			Pap		
	χ^2	ν 2)	k 2)	χ^2	ν 2)	k 2)
4.3.54	2,47	3	0,54	4,97	9	0,83
23.3.54	3,00	9	0,96	8,47	9	0,50
3.4.54	5,61	9	0,77	10,05	9	0,34
6.4.54	10,13	9	0,34	6,85	9	0,64
10.4.54	7,84	9	0,54	9,18	9	0,42
13.4.54	14,68	9	0,10			
15.4.54	2,29	9	0,99	12,57	9	0,18
totaal	46,02	57	0,85	52,09	54	0,50

De gecombineerde overschrijdingskansen werden verkregen door het optellen van de toetsingsgrootheden χ^2. De som van deze grootheden heeft dan onder de nul-hypothese bij benadering weer een χ^2-verdeling, met als aantal vrijheidsgraden de som van de afzonderlijke aantallen vrijheidsgraden.

De aantallen bij de kar-cellen waren te klein om een toets

---

2) ν is het aantal vrijheidsgraden
k is de overschrijdingskans

als bij de eos-cellen toe te passen. Hier is alleen een gecombineerde toets toegepast: de verwachting en de variantie van de toetsingsgrootheden zijn opgeteld en de overschrijdingskans is met de normale benadering bepaald.

Tabel IV

Vergelijking tussen de kar-tellingen der verschillende dagen

	H-S	Pap
$\Sigma\chi^2$	42,38	53,30
$\mathcal{E}\Sigma\chi^2$	57,06	54,05
d	- 14,68	- 0,75
$\Sigma\, var\,(\chi^2)$	106,29	94,29
σ	10,31	9,71
d/σ	- 1,42	- 0,08
k	0,92	0,53

Conclusie

Wij concluderen dat er noch bij de H-S, noch bij de Papkleuring enige aanleiding bestaat om aan te nemen, dat de kans om een eosinophiele, resp. karyopycnotische, cel te vinden bij de tellingen, voorkomende op dezelfde regel, verschillen. Dit betekent, dat men niet kan verwerpen dat de op één regel voorkomende aantallen eos resp. kar uit eenzelfde binomiale waarschijnlijkheidsverdeling komen; bij de kar-tellingen kan men deze verdeling, gezien de geringe aantallen, ook als een Poissonverdeling beschouwen.

2.2. Bij de toets in par. 2.1 wordt tegen alle mogelijke alternatieven getoetst. We doen het nu nog eens tegen het speciale alternatief, dat er een verloop in de tijd bestaat bij het tellen van cellen. Een dergelijk verloop zou kunnen ontstaan doordat men bij meerdere tellingen van één uitstrijkje steeds meer cellen gaat tellen, of ten gevolge van vermoeidheid van het oog steeds minder. Wij hebben dit onderzocht met een trendtoets — S 73 (M 13a), zie bijlage — omdat deze toets, wanneer het alternatief stijging of daling is, tot een duidelijker resultaat leidt.

De resultaten staan in tabel V vermeld.

Tabel V 3)

Onderzoek naar trend

datum	overschrijdingskansen			
	H-S		Pap	
	eos	kar	eos	kar
4.3.54	0,17 (+)	0,45 (+)	0,24 (-)	0,21 (+)
23.3.54	0,21 (+)	0,33 (-)	0,06 (-)	0,71 (-)
3.4.54	0,93 (+)	0,51 (-)	1.	0,36 (-)
6.4.54	0,47 (-)	0,12 (-)	1.	0,28 (+)
10.4.54	0,27 (+)	0,56 (-)	0,19 (+)	0,60 (+)
13.4.54	0,52 (-)	0,40 (-)		
15.4.54	0,30 (-)	0,12 (-)	0,20 (+)	0,03 (+)
gecombineerde overschrijdingskans	0,88 (+)	0,02 (-)	0,83 (-)	0,14 (+)

In deze tabel zien wij dat de gecombineerde overschrijdingskansen slechts in één geval een significant resultaat opleveren en wel een aanwijzing voor dalende trend. Dit is bij de H-S karceltellingen het geval. De andere overschrijdingskansen bevestigen echter deze dalende trend geenszins.

3. Vergelijking van tellingen op verschillende dagen verricht

Vervolgens hebben wij de tellingen van één regel samen genomen en met behulp van een 2 x k-tabel onderzocht of de eos, resp. kar, gekleurd volgens de H-S, resp. Pap-methode, voor de verschillende dagen overeenkomstige resultaten opleveren. Dit bleek niet het geval te zijn. De uitkomsten van dit onderzoek zijn vermeld in tabel VI.

Tabel VI

Vergelijking van de tellingen van verschillende dagen

methode	cellen	ν	χ^2	k
H-S	eos	5	81,46	$< 10^{-6}$
	kar	5	50,09	$< 10^{-6}$
Pap	eos	5	70,00	$< 10^{-6}$
	kar	5	53,56	$< 10^{-6}$

Opmerking:
Bij de berekening zijn de H-S-tellingen van 4.3.54 niet gebruikt.

Hieruit concluderen wij dat er voor beide kleuringsmethoden systematische verschillen bestaan zowel voor de fracties eos als de

---

3) Een (+) resp. (-) teken achter de overschrijdingskans geeft aan, dat de gevonden trend stijgend resp. dalend is.

fracties kar-cellen, gevonden bij de tellingen van de verschillende dagen van het onderzoek. Zie voor een verdere toelichting op deze conclusie par. 6.2.

4. Vergelijking van de beide kleuringsmethoden

Wij hebben vervolgens de hypothese getoetst, dat de kans om een eos, resp. kar-cel, te vinden bij gebruik van de H-S-methode gelijk is aan de overeenkomstige kans bij gebruik van de Pap-methode.

Voor de eos is dit onderzocht met behulp van de toets voor de hypothese $p_1 = p_2$ met behulp van een 2 x 2-tabel — zie memorandum S 53 (M 23) — voor iedere datum afzonderlijk.

De totalen per dag hebben wij hiertoe voor ieder der dagen als volgt gerangschikt:

datum	methode	eos	niet-eos	totaal
23.3.54	H-S Pap	57 102	943 898	1000 1000
	totaal	159	1841	2000

De resultaten zijn in tabel VII opgenomen

Tabel VII [4]

Vergelijking van kleuringsmethoden

eos	datum	23.3	3.4	6.4	10.4	15.4
	methode					
	H-S Pap	57 102	101 138	157 160	159 53	161 143
		0,0002 (-)	0,01 (-)	0,86 (-)	$< 10^{-6}$ (+)	0,26 (+)

Daar de aantallen van de kar-cellen veel kleiner zijn, is het hiervoor onderzocht met behulp van de toets voor gelijkheid van kleine kansen, beschreven in rapport S 131 van het Mathematisch Centrum [5] voor iedere datum afzonderlijk, op de totalen per dag.

----

4) Een (+) resp. (-) teken geeft aan dat de H-S-telling een groter, resp. kleiner, aantal geeft dan de Pap-telling.

5) Rapport S 131; Gerda Klerk-Grobben en H.J.Prins, Toets voor de gelijkheid van twee kleine kansen met behulp van even grote steekproeven en het onderscheidingsvermogen van deze toets.

De resultaten zijn in tabel VIII opgenomen.

Tabel VIII [4]

Vergelijking van kleuringsmethoden

kar	datum / methode	23.3	3.4	6.4	10.4	15.4
	H-S	14	30	57	18	15
	Pap	36	42	26	2	12
		0,003 (-)	0,195 (-)	0,0009 (+)	0,0004 (+)	0,702 (+)

De hier opgegeven overschrijdingskansen zijn bepaald door de tekentoets – memorandum S 53 (M 22) – toe te passen op de aantallen karyopycnotische cellen geteld met beide kleuringsmethoden. Deze methode ligt tevens ten grondslag aan de toets beschreven in rapport S 131 [5]; het blijkt daar echter dat het onderscheidingsvermogen kan worden vergroot door een kleine wijziging aan te brengen in de keuze van het kritieke gebied. Het is echter minder eenvoudig om deze wijziging ook aan te brengen bij de bepaling van overschrijdingskansen. Indien dit mogelijk was, zou men iets kleinere overschrijdingskansen vinden dan hier zijn opgegeven.

Wij constateren op grond van de gevonden overschrijdingskansen bij de eos en bij de kar-cellen op drie van de vijf waarnemingsdagen systematische verschillen tussen de H-S en de Pap-tellingen, waarbij in sommige gevallen met de H-S-methode en in andere met de Pap-methode meer cellen gevonden werden, zowel bij de eos als bij de kar.

Dit resultaat kan misschien te wijten zijn aan een minder goed gelukken van de kleuring, nu eens bij de ene, dan weer bij de andere kleuringsmethode. Voor een nadere toelichting hierop verwijzen wij naar par. 6.2.

5. Correlatie tussen de aantallen eosinophiele en karyopycnotische cellen

Verder hebben wij nog onderzocht of er een correlatie bestaat tussen de aantallen eos en kar-cellen. Dit onderzoek is uitgevoerd volgens de methode der rangcorrelatie – zie bijlage S 47 (M 13) – voor de beide kleuringsmethoden en voor iedere dag afzonderlijk. Daarna zijn de resultaten gecombineerd door optelling van de toetsingsgrootheden – zie bijlage S 73 (M 13a) De resultaten zijn in tabel IX samengevat.

Tabel IX

Onderzoek naar correlatie tussen de aantallen eos en kar-cellen

overschrijdingskansen		
datum	H-S	Pap
4.3.54	0,07 (+)	0,62 (+)
23.3.54	0,13 (+)	0,30 (+)
3.4.54	0,07 (+)	0,23 (+)
6.4.54	0,008 (+)	0,62 (+)
10.4.54	1.	0,77 (+)
13.4.54	0,50 (+)	
15.4.54	0,50 (-)	0,14 (+)
gecombineerde overschrijdingskans	0,013 (+)	0,036 (+)

Een (+) resp. (-) teken achter de overschrijdingskans geeft aan, dat de gevonden correlatiecoëfficiënt positief resp. negatief is.

Wij concluderen hier uit, dat er een positieve correlatie bestaat tussen de eos en de kar-celtellingen, hetgeen betekent, dat men bij tellingen, waarbij veel resp. weinig eos gevonden worden, meestal ook veel resp. weinig kar-cellen vindt.

Dit verband is voor de beide methoden van kleuren gevonden.

Gaan wij nu na in hoeverre deze conclusie gehandhaafd kan blijven in verband met de resultaten van tabel V, dan blijkt dat zowel bij de H-S-methode als bij de Pap-methode de eos en de kar-celtellingen een tegengesteld verloop in de tijd hebben. Daar dit op een negatieve correlatie tussen eos en kar-celtellingen zou wijzen, kan dit niet de oorzaak van de gevonden positieve correlatie zijn.

6. Samenvatting, toelichting, aanwijzingen voor verder onderzoek

6.1. Samenvatting van de conclusies

1. Wij hebben geen verschil gevonden tussen de tellingen die werden uitgevoerd op éénzelfde uitstrijkje. Men kan de hypothese niet verwerpen, dat de aantallen eos, resp. kar-cellen op één dag bij dezelfde patiënte gevonden, afkomstig zijn uit dezelfde binomiale verdeling. Bij de kar-cellen zal deze verdeling ook door een Poisson-verdeling kunnen worden benaderd.

2. Er werd bij beide gebruikte kleuringsmethoden een systematisch verschil gevonden tussen de aantallen eos en tussen de aantallen kar-cellen, geteld op de verschillende proefdagen bij dezelfde patiënte (zie toelichting in par. 6.2).

3. Het blijkt niet, dat de tellingen van eos, resp. kar-cellen na kleuring volgens de methode van Harris-Shorr over het geheel genomen systematisch hogere of lagere uitkomsten geeft dan na kleuring volgens de methode van Papanicolaou. Op verschillende proefdagen afzonderlijk werd echter wel een duidelijk verschil gevonden, soms werden echter met de ene, soms met de andere methode meer eos, resp. kar-cellen geteld.

4. Er werd een systematische positieve correlatie gevonden tussen de tellingen van de eos en van de kar-cellen na kleuring volgens ieder der beide methoden.

6.2. Toelichting van de conclusies

Ad 1. Uit conclusie 1. volgt, dat geen aanwijzingen gevonden zijn voor klonteringen van de eos of de kar-cellen in éénzelfde uitstrijkje en evenmin voor systematische verschillen in de tellingen verricht op eenzelfde uitstrijkje.

Ad 2. Het hiet geconstateerde verschil kan een gevolg zijn:

a. van een werkelijk dageffect, dus van systematische verschillen tussen de concentraties van eos en kar in het vocht op verschillende dagen;

b. van verschillen tussen kleuringen (ook bij gebruik van eenzelfde methode); deze kunnen weer een gevolg zijn van verschillen in de gebruikte kleurstof of van verschillen in de wijze van aanbrengen van de kleurstof of beide;

c. van verschillen in de plaats waar het vocht geprelevеerd is;

d. van klontering van de cellen in het vocht.
Voor deze mogelijkheid zijn bij 1. geen aanwijzingen gevonden, hoewel een grondiger onderzoek naar de trend gewenst is.

Ad 3. De hier geconstateerde verschillen kunnen berusten op één der onder ad 2. b, c en d genoemde oorzaken.

Ad 4. De gevonden correlatie is vermoedelijk reëel, daar geen der ad 2. genoemde punten het verschijnsel veroorzaakt kan hebben. Het onderzoek is nl. per uitstrijkje (dus ook voor iedere kleuring) afzonderlijk verricht en pas daarna zijn de resultaten afzonderlijk gecombineerd. Een nauwkeuriger onderzoek zou daarin bestaan, dat de cellen in 4 soorten onderscheiden worden (eos-kar, bas-kar, eos-niet-kar en bas-niet-kar) en dat de aantallen van deze vier soorten genoteerd worden.

6.3. Aanwijzingen voor verder onderzoek

Aan de hand van het waarnemingsmateriaal hebben wij de vragen 1.1 (conclusie 1); 1.3 (conclusie 2); 3 (conclusie 4) en de vraag 4 (conclusie 3) gedeeltelijk kunnen beantwoorden.

Wegens het ontbreken van de desbetreffende gegevens, was dit niet mogelijk voor de vragen 1.2 en 1.4.

Om een afdoende antwoord te kunnen geven op vraag 2 is het gewenst, dat de ad 2. genoemde punten nader onderzocht worden. Om de reproduceerbaarheid van de kleuring te onderzoeken kan men een aantal uitstrijkjes afzonderlijk, uit dezelfde pipet of spatel, dus op hetzelfde moment bij dezelfde patiënte gepreleveerd, volgens één van beide kleuringsmethoden met hetzelfde flesje kleurstof al of niet op verschillende dagen kleuren. Men kan dan nagaan of de kleurstof aan verandering onderhevig is.

Een andere groep uitstrijkjes uit dezelfde pipet of spatel kan op dezelfde dag met kleurstof uit verschillende flesjes gekleurd worden om na te gaan of er geen variaties in de kleurende werking van de kleurstof bestaan.

Als de kleuring voldoende onderzocht is kan men nagaan of de plaats waar het vocht wordt afgenomen enige invloed heeft, door uitstrijkjes van het vocht, tegelijkertijd op verschillende plaatsen bij dezelfde patiënte afgenomen, te vergelijken.

Als de mogelijke nevenoorzaken zijn onderzocht kan pas worden vastgesteld of er een werkelijk dageffect is en of er enig verschil bestaat tussen de beide kleuringsmethoden, waarna het wellicht mogelijk is de vraag betreffende de betekenis van het percentage, gevonden na één of na meerdere tellingen, te beantwoorden.

MATHEMATISCH CENTRUM
Amsterdam
Statistische Afdeling.

S47(M6)

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

De toetsing van een hypothese 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 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 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 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 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 $\mathcal{H}_0$ 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 waarnemingen zijn verricht. Deze nevenvoorwaarden dienen steeds uitdrukkelijk te worden vermeld en, zo mogelijk, zelf te worden getoetst.

In plaats van de onbetrouwbaarheidsdrempel α wordt vaak bij de uitslag van een toetsing de overschrijdingskans k opgegeven; dit is de kleinste waarde van α, waarbij in het betrokken geval, nog tot verwerping van $\mathcal{H}_0$ 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. 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 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.

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.

MATHEMATICSCH CENTRUM,
2de Boerhaavestr. 49,
Amsterdam-O.
Statistische Afdeling
S 145 (M 23a)

Toets voor de hypothese $p_1 = p_2 = \ldots = p_k$ met behulp van een $2 \times k$ tabel [1)]

Wij beschouwen k reeksen $R_1, R_2, \ldots, R_k$ van onafhankelijke waarnemingen, waarbij iedere waarneming als resultaat het kenmerk A of het kenmerk $\bar{A}$ (non-A) kan geven. De kans op A is binnen ieder der reeksen constant en wel gelijk aan p_i voor de waarnemingen van reeks R_i $(i = 1,2,\ldots,k)$.
Laat het aantal waarnemingen van reeks R_i $(i = 1,2,\ldots,k)$ gelijk zijn aan n_i en laat hieronder het aantal met kenmerk A, m_i zijn.
Gevraagd wordt dan de hypothese H_0: $p_1=p_2= \ldots =p_k$
te toetsen op grond van deze gegevens.
De gegevens kunnen in een $2 \times k$-tabel worden samengevat;

	R_1	R_2		R_i		R_k	totaal
A	m_1	m_2	...	m_i	...	m_k	m
$\bar{A}$	n_1-m_1	n_2-m_2	...	n_i-m_i	...	n_k-m_k	$n-m$
totaal	n_1	n_2		n_i		n_k	n

waarin dus $m_1 + m_2 + \ldots + m_k = m$
en $n_1 + n_2 + \ldots + n_k = n$.
De hypothese H_0 wordt getoetst met de grootheid

$$\chi_c^2 = \sum_i \frac{\left(m_i - \frac{m\, n_i}{n}\right)^2}{\frac{m\, n_i}{n}} + \sum_i \frac{\left(n_i - m_i - \frac{(n-m)n_i}{n}\right)^2}{\frac{(n-m)\, n_i}{n}} =$$

$$= \sum_i \frac{(n m_i - m n_i)^2}{m(n-m)n_i} = \frac{n^2}{m(n-m)} \sum_i \frac{m_i^2}{n_i} - \frac{nm}{n-m}$$

Deze grootheid $\underline{\chi}_c^2$ [2)] heeft onder de hypothese H_0 bij benadering een χ^2-verdeling met $k-1$ vrijheidsgraden (zie b.v. [1] p. 445 e.v.).

---

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

2) Als wij grootheden als stochastische grootheden (dit zijn grootheden met een waarschijnlijkheidsverdeling) beschouwen geven wij dit door onderstreping aan. Niet onderstreepte letters geven waarden aan, die door de stochastische grootheden worden aangenomen.

Deze benadering is goed, indien $m\frac{n_i}{n} \geqq 5$ voor iedere i (zie [2]). Indien H_0 onjuist is, dus als er bij verschillende reeksen verschillende kansen op A zijn, zal $\underline{\chi}_c^2$ gewoonlijk grotere waarden aannemen, dan wanneer H_0 juist is.

De kritieke zone bestaat uit die waarden van $\underline{\chi}_c^2$, waarvoor geldt $\underline{\chi}_c^2 \geqq \chi_\alpha^2$. Hierin is χ_α^2 die waarde van $\underline{\chi}^2$, die voldoet aan

$$P[\underline{\chi}^2 \geqq \chi_\alpha^2] = \alpha$$

met α als van te voren vastgelegde onbetrouwbaarheid.

De overschrijdingskans behorende bij een bepaalde gevonden waarde χ_c^2 van $\underline{\chi}_c^2$ is gedefinieerd als

$$P[\underline{\chi}_c^2 \geqq \chi_c^2 | H_0]$$

waarin "$|H_0$" aangeeft, dat deze kans berekend wordt op grond van H_0.

χ_α^2 en de overschrijdingskans kunnen in tabellen of nomogrammen worden opgezicht (zie [3]).

Opmerking. Indien niet voldaan is aan de voorwaarde $m\frac{n_i}{n} \geqq 5$ voor iedere i, kan men een (meer bewerkelijke) exacte toets baseren op de voorwaardelijke waarschijnlijkheidsverdeling van de grootheden $\underline{m}_i$ ($i = 1,\ldots,k$), onder de voorwaarde, dat hun som de waarde m aanneemt:

$$P[\underline{m}_1 = m_1, \underline{m}_2 = m_2, \ldots, \underline{m}_k = m_k | \underline{m}_1 + \underline{m}_2 + \cdots + \underline{m}_k = m ; H_0] =$$

$$= \binom{n_1}{m_1}\binom{n_2}{m_2}\cdots\binom{n_k}{m_k} \Big/ \binom{n}{m}$$

De geldigheid van deze formule volgt direct uit de waarschijnlijkheidsverdelingen van de $\underline{m}_i$ en van $\underline{m}$ (onder H_0) en uit de definitie van een voorwaardelijke waarschijnlijkheid.

In dit geval definiëren wij de overschrijdingskans behorend bij een gevonden resultaat $(m_1, m_2, \ldots, m_k)$ met $m_1 + m_2 + \ldots + m_k = m$ als de som van alle waarschijnlijkheden van bovengenoemde verdeling (met de gevonden waarde van m), die hoogstens gelijk zijn aan de waarschijnlijkheid van het gevonden resultaat.

Literatuur.

[1] H.Cramér, Mathematical methods of statistics, Princeton University Press, 1946.

[2] P.G.Hoel, On indices of dispersion, Ann. Math. Stat. 14 (1943), p. 155-163.

[3] Tabellen en nomogrammen van de -verdeling.
M.G.Kendall, The advanced theory of statistics, I, 1947, p. 444-446.
H.Cramér, Mathematical methods of statistics, Princeton University Press, 1946, p. 559.
Statistica 1 (1946), p. 109.

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

Statistische Afdeling.

S 73 (M 13a)

Trend-toets met behulp van rangcorrelatie[1]

Men kan de methode der rangcorrelatie toepassen om na te gaan of er in een reeks waarnemingen van een stochastische grootheid $\underline{y}$ 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 $\underline{S}$ 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_o, 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 $y_1, \ldots, y_n$ van de stochastische grootheid $\underline{y}$, 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_o. Onder deze hypothese heeft S dezelfde verdeling als de overeenkomstige grootheid der rangcorrelatietoets 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:

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

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

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

$$\vdots$$

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

---

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 $(1,y_{11})(2,y_{12})\ldots$ $\ldots(n_1,y_{1n_1})$ en noemen deze S_1. Evenzo behandelen wij de tweede rij en verkrijgen zo S_2 etc.

Wij bepalen dan de som

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

Eveneens berekenen wij bij de eerste rij de spreiding σ_1, volgens formule (1) of formule (2) van memorandum S 47 (M 13). Evenzo σ_2, $\sigma_3, \ldots, \sigma_m$.

Wij weten nu, dat onder de hypothese H_o voor rij 1 $\underline{S}_1$ bij benadering normaal verdeeld is met gemiddelde 0 en spreiding σ_1, indien H_o geldt voor rij 2 is $\underline{S}_2$ bij benadering normaal verdeeld met gemiddelde 0 en spreiding σ_2 etc. Indien hypothese H_o 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 $\underline{S}_{tot}$ bij benadering normaal verdeeld is met gemiddelde 0 en spreiding:

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

Wij kiezen nu als kritieke zône weer $|S_{tot}| \geqq S_o$ bij tweezijdige toetsing, $S_{tot} \geqq S_o'$ bij rechtszijdige en $S_{tot} \leqq S_o''$ 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.Elfving en J.H.Whitlock: A simple trend-test with application to erythrocyte data, Biometrics 6 (1950) Blz. 282-288.

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

Statistische Afdeling
S 53 (M 23).

Toetsing van de hypothese $p_1 = p_2$ met behulp van een 2 x 2-tabel [1])

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:

$$H_o : p_1 = p_2.$$

Indien de eerste reeks uit n en de tweede reeks uit m waarnemingen bestaat, waaronder a (resp. b) 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	a	c	n
tweede reeks	b	d	m
totaal	r	s	N

Als toetsingsgrootheid wordt a, het aantal malen A in de eerste reeks waarnemingen, gebruikt. Indien H_o, 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 a aangenomen wordt, is gelijk aan:

$$\frac{\binom{n}{a}\binom{m}{b}}{\binom{N}{r}}$$

Als kritieke zone worden de waarden van a met de kleinste waarschijnlijkheden bijeengezocht, tot de gekozen onbetrouwbaarheidsdrempel het toevoegen van een nieuwe waarde verhindert (bij éénzijdige toetsing bestaat de kritieke zone uitsluitend uit grote of uitsluitend uit kleine waarden van a). De overschrijdingskans, behorende bij de gevonden waarde van a, is gedefinieerd als de som van alle waarschijnlijkheden van bovenstaande verdeling, die hoogstens gelijk aan de waarschijn-

---

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

lijkheid van de gevonden waarde zijn (bij éénzijdige 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 a zijn (indien H_0 juist is):

$$\frac{nr}{N} \text{ resp. } \sqrt{\frac{nmrs}{N^2(N-1)}} .$$

Men gebruikt dan in plaats van de exacte waarschijnlijkheidsverdeling van a de normale verdeling met hetzelfde gemiddelde en dezelfde spreiding en in plaats van de gevonden waarde van a 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). Bij positieve $a - \frac{nr}{N}$ berekent men dus:

$$a^* = \frac{a - \frac{nr}{N} - \frac{1}{2}}{\sqrt{\frac{nmrs}{N^2(N-1)}}} = \frac{aN - nr - \frac{1}{2}N}{\sqrt{\frac{nmrs}{N-1}}} = \frac{ad - bc - \frac{1}{2}N}{\sqrt{\frac{nmrs}{N-1}}}$$

en bij negatieve $a - \frac{nr}{N}$ berekent men:

$$a^* = \frac{a - \frac{nr}{N} + \frac{1}{2}}{\sqrt{\frac{nmrs}{N^2(N-1)}}} = \frac{ad - bc + \frac{1}{2}N}{\sqrt{\frac{nmrs}{N-1}}} .$$

De overschrijdingskans wordt nu opgezocht in een tabel der normale verdeling met gemiddelde 0 en spreiding 1. De rechts-éénzijdige (resp. links-éénzijdige) overschrijdingskans is het oppervlak rechts (resp. links) gelegen van a^*. De tweezijdige overschrijdingskans is twee maal het oppervlak der normale verdeling dat rechts van $\frac{|ad - bc| - \frac{1}{2}N}{\sqrt{\frac{nmrs}{N-1}}}$ ligt.

<u>Literatuur</u>.

R.A.Fisher, Statistical Methods for Research Workers, London 1948, p. 96. <u>Opmerking</u>: Fisher gebruikt hier de éénzijdige overschrijdingskans.

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

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

Statistische Afdeling
S 53 (M 22)

Tekentoets[1]

Deze toets dient voor het toetsen van de hypothese H_0, dat een aantal grootheden $\underline{z}_1, \dots, \underline{z}_n$ alle nul tot mediaan hebben, d.w.z. dat

$$P[\underline{z}_i > 0] = P[\underline{z}_i < 0] \qquad i = 1, \dots, n,$$

is. De toets geldt zonder enige verdere beperking dan de eis, dat de grootheden $\underline{z}_i$ onderling onafhankelijk verdeeld zijn; zij behoeven niet dezelfde waarschijnlijkheidsverdeling te bezitten.

De toets berust op één waarneming van ieder der grootheden $\underline{z}_i$, dus op n waarnemingen $z_1, \dots, z_n$. De waarnemingen, die de waarde 0 bezitten, laten wij buiten beschouwing[2]. Als toetsingsgrootheid gebruiken wij nu $\underline{n}_1$, het aantal positieve waarnemingen. Zijn er m waarnemingen $\neq 0$, dan bezit $\underline{n}_1$ een binomiale verdeling, onderstellende, dat H_0 juist is:

$$P[\underline{n}_1 = n_1 \mid H_0] = \binom{m}{n_1} 2^{-m}.$$

Als kritieke zône worden de grote en kleine waarden van n_1 genomen. De kritieke zône is, voor onbetrouwbaarheidsdrempels 0,01; 0,05; 0,10 en 0,25 en m =1 tot 100 getabelleerd door

W.J.Dixon and A.M.Mood, The statistical sign test, Jrn.Am.Stat.Ass. 41 (1946) p. 556-566.

Voor een groter aantal waarnemingen gebruikt men als benadering van de binomiale verdeling de aangepaste normale verdeling.

Opmerking: De toets wordt vaak gebruikt, indien men een aantal grootheden twee maal heeft waargenomen, voor en na een bepaalde gebeurtenis, om na te gaan of deze gebeurtenis invloed op de grootheden heeft uitgeoefend. Noemen wij de waarnemingen vóór het optreden der gebeurtenis $\underline{x}_i$ (i=1,...,n) en erna $\underline{y}_i$, dan hebben de grootheden $\underline{x}_i - \underline{y}_i$ alle o als mediaan, indien $\underline{x}_i$ dezelfde verdeling bezit als $\underline{y}_i$ (dus als de gebeurtenis geen invloed heeft gehad). De toets wordt nu toegepast op $\underline{z}_i = \underline{x}_i - \underline{y}_i$ (i=1,...,n).

---

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

2) In tegenstelling tot de gewoonte deze waarnemingen voor de helft bij de positieve en voor de helft bij de negatieve te tellen; de door ons gebruikte methode geeft de toets een groter onderscheidingsvermogen.

MATHEMATISCH CENTRUM
2de Boerhaavestr. 49,
Amsterdam -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 $|\underline{S}| \geq S_0$, en bij éénzijdige toetsing van de vorm $\underline{S} \geq S_0'$ (rechtszijdige toetsing) of $\underline{S} \leq S_0''$ (linkszijdige toetsing).

2. Verdeling van $\underline{S}$ als $\underline{x}$ en $\underline{y}$ 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 $\underline{S}$, waarvan de overschrijdingskansen onder de hypothese van onafhankelijkheid hoogstens gelijk zijn aan α voor α = 0,005; 0,01; 0,025; 0,05 en 0,10 en n = 4,5,6,...,40.

Als er bij de $\underline{x}_i$ óf bij de $\underline{y}_j$, 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 $\frac{\underline{S}}{\sigma_{\underline{S}}}$ (waarin $\sigma_{\underline{S}}$ de spreiding van $\underline{S}$ 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 $h = 1,2,\dots,k_1$. Evenzo doet men in de rij der waarnemingen y_j, waar we de overeenkomstige aantallen aanduiden met u_ℓ waarin $\ell = 1,2,\dots,k_2$. $\sigma_{\underline{S}}$ kan dan gevonden worden uit de volgende formule:

$$(1)\quad \sigma^2_{\underline{S}} = \frac{1}{18}\left\{ n(n-1)(2n+5) - \sum_{h=1}^{k_1} t_h(t_h-1)(2t_h+5) - \sum_{\ell=1}^{k_2} u_\ell(u_\ell-1)(2u_\ell+5)\right\} + \frac{1}{9n(n-1)(n-2)}\sum_{h=1}^{k_1} t_h(t_h-1)(t_h-2)\sum_{\ell=1}^{k_2} u_\ell(u_\ell-1)(u_\ell-2) + \frac{1}{2n(n-1)}\sum_{h=1}^{k_1} t_h(t_h-1)\sum_{\ell=1}^{k_2} u_\ell(u_\ell-1).$$

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:

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

zodat:

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

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:

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

Een tabel van deze functie voor $n = 40,41,\dots,100$ 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 τ. 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. Biometrika 34 (1947) p. 36-40.