Statistische Afdeling
van het
Mathematisch Centrum,
Amsterdam

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

Rapport S 91

Palingvangst in de binnenwateren

door

Mej. C. van Eeden.

1. Inleiding:

Door het Rijksinstituut voor Visserij-Onderzoek werden ons de volgende gegevens verstrekt:

1. het aantal ponden paling dat ieder van een aantal vissers per dag gevangen heeft in de Friese binnenwateren. Deze dagvangsten werden ons opgegeven voor de jaren 1947 t/m 1951.
   Iedere visser heeft een vaste vis-plaats en het aantal netten dat hij uitzet, àls hij netten uitzet, mag als constant beschouwd worden. Het aantal vissers varieërt hier over de verschillende jaren van 14 tot 22.
2. het aantal ponden paling dat ieder van een aantal vissers per week gevangen heeft op het IJselmeer. Deze weekvangsten werden ons opgegeven voor de jaren 1950 en 1951. Het aantal dagen per week waarop gevist wordt, àls er gevist wordt, is constant.
   Een deel der vissers is afkomstig uit Wieringen, de anderen uit Makkum. Het aantal vissers is hier 11 à 13.

De vraag is nu:

a. of de stand van de maan invloed heeft op de palingvangst.

b. of de weersgesteldheid (en wel in het byzonder het passeren van een storingsfront van een depressie[1]) invloed heeft op de palingvangst.

2. Onderzoek naar de maaninvloed:

2.1. De palingvangst in de Friese binnenwateren:

Dit onderzoek is uitgevoerd voor de vangst van een viertal vissers met grote dagvangsten (n.l. de n$^{\underline{o}}$ 13, 18, 19 en 26) en voor de vangst van een viertal ~~met~~ kleinere dagvangsten (n.l. de n$^{\underline{o}}$ 3, 7, 14 en 21).

Het is hier niet nodig alle waarnemingen te gebruiken, daar mèt de vangsten van deze acht vissers de maaninvloed reeds zeer duidelijk aangetoond kan worden. Er zijn bovendien zoveel waarnemingen, dat gebruik van al deze waarnemingen een grote hoeveelheid onnodig werk zou veroorzaken.

----

1) Het passeren van zo'n front gaat gepaard met een sterke daling der barometerstand en een grote windsnelheid.

Het onderzoek is uitgevoerd met behulp van de methode der m rangschikkingen (zie bijlage S 47(M 14)). De rangschikkingen zijn als volgt gevormd:
Van ieder der bovengenoemde vissers wordt de vangstperiode van ieder der jaren verdeeld in perioden, lopend van eerste kwartier tot eerste kwartier. De waarnemingen van één zo'n maanperiode zetten we naast elkaar, de verschillende perioden onder elkaar. We krijgen dan een schema waarin de rijen niet alle even lang zijn, terwijl er bovendien open plaatsen optreden (dagen waarop niet gevist is).
Nu wordt in iedere rij het gemiddelde bepaald van de waarnemingen uit telkens drie opeenvolgende kolommen. We bepalen in iedere rij negen van deze gemiddelden; eventuele verdere waarnemingen worden weggelaten.
Uit het zo verkregen schema worden rijen, waarin nog open plaatsen voorkomen, weggelaten en op het dan verkregen schema wordt de methode der m rangschikkingen toegepast.

De resultaten vinden we in tabel I:

<u>Tabel I:</u>

<u>Kolomtotalen en overschrijdingskansen gevonden met de methode der m rangschikkingen.</u>

maan-stand	☽		○		☾			●		Over-schrij-dings-kans.
vissers n°	Kolomtotalen									
13,18 19 en 26	250,5	252,5	167,0	188,0	327,0	388,0	344,0	263,5	249,5	$\ll 10^{-4}$
3,7,14 en 21	2[illegible],5	186,0	112,0	142,0	230,0	315,0	300,0	230,0	217,5	$\ll 10^{-4}$

We vinden dus een duidelijke overeenstemming tussen de rijen en een maximum palingvangst tussen laatste kwartier en nieuwe maan, een minimum vangst omstreeks volle maan.

2.2. <u>De palingvangst op het IJselmeer:</u>
Hier is het onderzoek toegepast op de vangsten van alle vissers en van beide jaren.

We passen weer de methode der m rangschikkingen toe en een rangschikking bestaat hier uit de waarnemingen van vier achtereenvolgende weken, waarbij de eerste week steeds een week is waarin eerste kwartier valt.
Eventuele rijen met open plaatsen worden weggelaten.
De resultaten staan vermeld in tabel II:

Tabel II

Kolomtotalen en overschrijdingskansen gevonden met de methode der m rangschikkingen

☽	○	☾	●	Overschrijdingskans.
Kolomtotalen				
251	223	352	354	$\ll 10^{-4}$

Ook hier vinden we dus een duidelijke overeenstemming tussen de rijen en een maximum vangst in de weken van laatste kwartier en nieuwe maan en een minimum in de week van volle maan.

3. Onderzoek naar de invloed van een frontpassage:

3.1. De palingvangst in de Friese binnenwateren;

Om een eventuele invloed van een frontpassage op de palingvangst na te gaan hebben wij de vangst op dagen met frontpassage vergeleken met die op dagen zonder frontpassage [2].
Hierbij moeten we ervoor zorgen dat de andere invloeden (seizoen-maan-en jaarinvloed) zo veel mogelijk uitgeschakeld zijn. Daarom kiezen we de te vergelijken data als volgt:
We nemen twee dagen uit een zelfde jaar, waarbij op de ene dag wel en op de andere dag niet een front passeert. Uit een ander jaar kiezen we ook twee dergelijke dagen en wel zo dat:

a. een dag met een frontpassage uit het ene jaar dezelfde maanstand heeft als een dag zonder frontpassage uit het andere jaar. Hierdoor wordt de maaninvloed uitge schakeld.

---

2) Onder een dag wordt hier verstaan de periode van 24 uur tussen twee opeenvolgende keren dat de netten worden opgehaald. Onder de vangst van een bepaalde dag wordt dan verstaan de vangst bij de tweede keer dat de netten worden opgehaald.

b. de data alle vier zo dicht mogelijk bij elkaar liggen. Hierdoor wordt de seizoen invloed zoveel mogelijk uitgeschakeld.[3]

Verder nemen we deze vier dagen nog zo dat op de dagen ervoor en erna geen front passeert; dit om de invloed van andere fronten uit te schakelen.
We kunnen nu de jaren nog op verschillende manieren twee aan twee combineren en het blijkt dat de data zo dicht mogelijk bij elkaar liggen als we 1948 met 1949 en 1947 met 1950 combineren.
We kiezen nu uit de combinatie 1948-1949 en uit de combinatie 1947-1950 alle mogelijke viertallen dagen, die aan bovengestelde eisen voldoen en berekenen dan voor alle viertallen dagen apart en voor alle vissen apart de som van de vangsten op de twee dagen met een frontpassage verminderd met de som van de vangsten op de twee dagen zonder frontpassage. Door op deze wijze de verschillen te nemen tussen vangsten op dagen met een frontpassage en die op dagen zonder frontpassage is de jaarinvloed uitgeschakeld.
Op de op deze wijze verkregen verschillen wordt de symmetrietoets T_2(zie bijlage S 47(M 10)) toegepast. Wij onderzoeken dan dus of een frontpassage invloed heeft op de palingvangst direct na die frontpassage. De resultaten van de symmetrietoets vinden we in tabel III.

Tabel III:

Onderzoek naar de invloed van een frontpassage op de palingvangst direct na die frontpassage.

n	r	u	v	Overschrijdingskans
48	23	11	13	> 0,10

---

3) We zouden de vier dagen ook zo kunnen kiezen dat een dag met een frontpassage uit het ene jaar dezelfde datum heeft als een dag zonder frontpassage uit het andere jaar en ze dan nog zo te nemen dat de maanstand op de vier dagen zoveel mogelijk dezelfde is. Daar echter de maaninvloed belangrijker is dan de seizoeninvloed hebben wij de maanstanden gelijk genomen.

Op analoge wijze kunnen we onderzoeken of een frontpassage de palingvangst vlak vóór die frontpassage beïnvloedt. We vinden dan:

Tabel IV:

Onderzoek naar de invloed van een frontpassage op de palingvangst vlak vóór die frontpassage:

n	r	u	v	Overschrijdingskans
38	20	8	9	> 0,10

Uit de tabellen III en IV zien we dus dat er geen enkele reden is om aan te nemen dat een frontpassage invloed heeft op de palingvangst direct na of op die direct vóór die frontpassage.

3.2. Palingvangst op het IJselmeer:

De invloed van een frontpassage zou hier op analoge wijze onderzocht kunnen worden als in § 3.1. beschreven is voor de vangst in de Friese binnenwateren. We kiezen dan i.p.v. viertallen dagen viertallen weken.
Het onderzoek is echter niet uitgevoerd daar het aantal waarnemingen te klein is, om er enig succes van te verwachten.

MATHEMATISCH CENTRUM
Amsterdam
Statistische Afdeling.
S47(M6)

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

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

Bij een toets behoort een toetsingsgrootheid $\underline{u}$ (soms meer dan één), die een functie is van bovengenoemde stochastische grootheden en die, voor de waargenomen waarden $x_1, x_2, \ldots, x_n$ een waarde aanneemt, die berekend kan worden (bv.: het gemiddelde der waarnemingen, of de spreiding, of het verschil van de gemiddelden van twee waarnemingen).

De toetsingsgrootheid wordt steeds zo gekozen, dat men, op grond van de onderstelling, dat $\mathcal{H}_0$ juist is, de waarschijnlijkheidsverdeling van deze grootheid kan berekenen.

Vervolgens kiest men een verzameling Z van mogelijke uitkomsten van $\underline{u}$, en wel op zodanige wijze, dat de kans, dat $\underline{u}$ een in Z gelegen waarde aanneemt, onder de hypothese $\mathcal{H}_0$, gelijk is aan een gegeven getal α, zodat Z dus van α afhankelijk is[3]. Z heet de kritieke zône van de toets, α de onbetrouwbaarheidsdrempel (Engels: level of significance). Voor α neemt men veelal de waarde 0,05 of 0,01.

Men verwerpt nu $\mathcal{H}_0$ op grond van de waarnemingen $x_1, x_2, \ldots, x_n$, indien de bij deze waarnemingen behorende waarde van $\underline{u}$ in Z ligt. Dit wordt vaak uitgedrukt door te zeggen, dat het resultaat van het experiment "significant" is. De waarde van α moet dan echter worden vermeld. De kans, dat dit zal gebeuren, is, indien $\mathcal{H}_0$ juist is, gelijk aan α. Derhalve is α de kans op ten onrechte verwerping van de juiste hypothese, ook de kans op een fout van de eerste soort genoemd. Indien men deze methode toepast, met 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, álvorens ze bij de verdere bewerking van het materiaal te gebruiken. Worden zij dan op grond van de toets niet verworpen, dan houdt dit in zo verre een rechtvaardiging van die onderstellingen in, dat een grote afwijking door de toets veelal wel zou zijn ontdekt. Indien men dan verder de onderstellingen gebruikt, verwaarloost men eventueel aanwezige afwijkingen van onbekende grootte, die echter niet zo groot zijn, dat zij door de toets zijn ontdekt.

Vele toetsen gelden zelf alleen onder bepaalde onderstellingen omtrent de waarschijnlijkheidsverdelingen der stochastische grootheden, waarvan waarnemingen zijn verricht. Deze <u>nevenvoorwaarden</u> dienen steeds uitdrukkelijk te worden vermeld en, zo mogelijk, zelf te worden getoetst.

In plaats van de onbetrouwbaarheidsdrempel α wordt vaak bij de uitslag van een toetsing de <u>overschrijdingskans</u> k opgegeven; dit is de kleinste waarde van α, waarbij in het betrokken geval, nog tot verwerping van $\mathcal{H}_0$ 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 Boerhaavestr. 49,
Amsterdam-O.
Statistische Afdeling.
S 47 (M 10)

Symmetrietoets[1].

Hypothese H_0: de waarnemingen $z_1, \ldots, z_{n'}$ zijn afkomstig van n' onafhankelijk verdeelde stochastische grootheden, die alle symmetrisch ten opzichte van 0 verdeeld zijn[2]. Van deze toets bestaan meerdere versies $T_1, \ldots, T_2''$. We bespreken eerst T_1 en T_2.

Toetsingsgrootheden. Deze worden als volgt uit $z_1, \ldots, z_{n'}$ afgeleid:

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

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

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

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

5e. De groep waarden $w_1, \ldots, w_n$ wordt verdeeld in twee groepen $w_1, w_2, \ldots, w_r$ en $w_{r+1}, \ldots, w_n$, 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 $w_1, \ldots, w_r$ 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 B[3].

---

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 $x_1 - a, \ldots, x_{n'} - a$.

3) In de oorspronkelijke publicaties over deze toets (zie de litteratuurverwijzingen aan het einde van dit memorandum) is een enigszins minder algemene definitie van r gegeven. Alle stellingen blijven echter gelden, indien de hier gegeven definitie gebruikt wordt.

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

De toetsingsgrootheden zijn $\underline{n_1}$ en $\underline{u}$, r is een hulpgrootheid.

<u>V.B.</u> 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/-6,3/-7,0/-7,9/-8,0/-8,7.

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

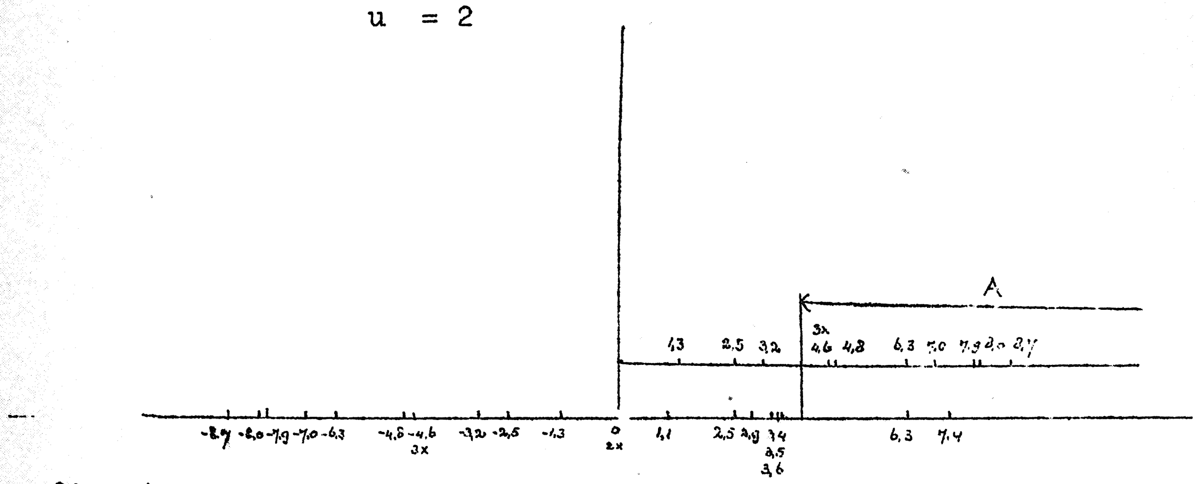

fig. 1

<u>Kritieke zônes</u>. 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 <u>positieve</u> (resp, negatieve) waarden zijn, die <u>verder</u> van 0 verwijderd liggen dan de <u>negatieve</u> (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.

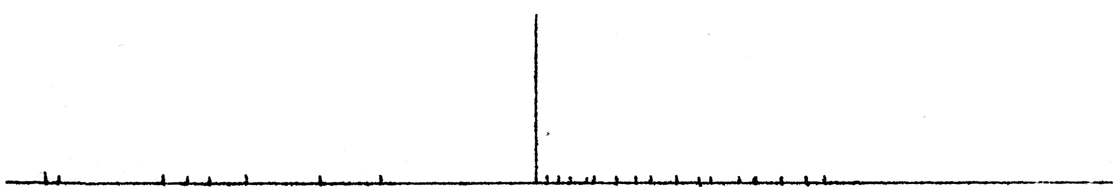

fig. 2

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

<u>T_1' en T_2'</u>.

<u>Toetsingsgrootheden</u>.

1e, 2e en 3e als boven. Eerste toetsingsgrootheid: $\underline{n}_1$.

4e: op $x_1,\ldots,x_{n_1}$ en $y_1,\ldots,y_{n_2}$ wordt de toets van Wilcoxon toegepast (vgl. S 47 (M 8)). De toetsingsgrootheden zijn $\underline{n}_1$ en de $\underline{U}$ van deze toets van Wilcoxon.

<u>Kritieke zônes</u>. 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'.

<u>Opmerkingen</u>: 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 $\underline{u}$ in plaats van $\underline{U}$ wordt gebruikt (vgl. b.v. [2], blz. 77, § 6.4.5).

<u>Litteratuur</u>:

[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. <u>53</u> (1950), p.945-955. Indagationes Mathematicae <u>12</u> (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
S47 (M14)

Methode der m rangschikkingen [1])

Een duidelijk voorstelling van deze toetsingsmethode verkrijgt men door n elementen te beschouwen, die een bepaald kenmerk, eventueel in verschillende mate, bezitten. dit kenmerk wordt door m waarnemers beoordeeld en ieder van deze waarnemers rangschikt deze n elementen volgens zijn beoordeling naar opklimmende waardering. Op deze wijze ontstaan m rijen van rangschikkingen. We willen nu een maat aangeven voor de overeenstemming tussen deze rangschikkingen, m.a.w. een maat voor de overeenstemming tussen de m beoordelingen. De hypothese $\mathcal{H}_0$, die met deze methode getoetst kan worden, houdt in dat er geen overeenstemming tussen de waarnemers bestaat; precieser gezegd, dat alle rangschikkingen onafhankelijk van elkaar op toevallige wijze zijn ontstaan. Dit is b.v. het geval, als het betrokken kenmerk in werkelijkheid voor alle elementen dezelfde waarde bezit.

We kunnen de afleiding voor de maat van overeenstemming het eenvoudigst geven aan de hand van een voorbeeld.

elementen		A	B	C	D	E	F
rangnummers toegekend door waarnemer	a	5	4	1	6	3	2
	b	2	3	1	5	6	4
	c	4	1	6	3	2	5
	d	4	3	2	5	1	6
		15	11	10	19	12	17

De som van alle rangnummers is $\frac{1}{2} n m(n+1)$. Onder de hypothese $\mathcal{H}_0$ is het theoretische gemiddelde van iedere kolom: $\frac{1}{2} m(n+1)$

We beschouwen nu de afwijkingen van dit gemiddelde. In ons voorbeeld is het theoretisch kolomgemiddelde gelijk aan 14. De afwijkingen daarvan zijn

$$1 \quad -3 \quad -4 \quad 5 \quad -2 \quad 3$$

---

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

De som der kwadraten van deze afwijkingen noemen wij S.
In ons voorbeeld is S = 64.
Als alle m rangschikkingen gelijk zijn wordt het maximum van S bereikt.
Dit maximum is $\frac{1}{12} m^2(n^3-n)$.
We definiëren nu als coëfficiënt van overeenstemming

$$W = \frac{12\ S}{m^2(n^3-n)}$$

In ons voorbeeld is $W = \frac{12 \times 64}{16 \times 210} = 0,229$.
W varieert dus tussen 0 en 1.

De verdeling van S onder de hypothese H_0 is exact berekend voor een aantal waarden van n en m [1] , terwijl voor grote m en n benaderingen bekend zijn.

De meest gebruikelijke benaderingen zijn de volgende.

1°. De χ^2-benadering:

$\chi_r^2 = m(n-1)\underline{W} = \frac{12\ \underline{S}}{mn(n+1)}$ heeft voor $m \to \infty$ een χ^2-verdeling met n-1 vrijheidsgraden ([1] pg. 84 [2] pg. 36-37).

2°. De z-benadering:

$\underline{V} = (m-1)\frac{\underline{W}}{1-\underline{W}}$ is bij benadering verdeeld als $\underline{F} = e^{2\underline{z}}$

($\underline{F}$ is de $\underline{F}$ van Snedecor, $\underline{z}$ de $\underline{z}$ van Fisher) met

$\nu_1 = n-1-\frac{2}{m}$

$\nu_2 = (m-1)\nu_1$ vrijheidsgraden ([1] pg. 84 [2] pg.33-36).

Met behulp van de verdelingen van S of W onder de hypothese H_0, kan deze hypothese getoetst worden, waarbij H_0 verworpen wordt als W waarden dichtbij 1 (resp. S dichtbij $\frac{1}{12} m^2(n^3-n)$) aanneemt, de kritieke zône is dus van de vorm $W \geqq W_0$ (resp. $S \geqq S_0$).

Het kan voorkomen dat de waarnemers geen onderscheid ontdekken in de mate waarin verschillende elementen het kenmerk bezitten. Ze geven deze elementen dan gelijke rangnummers.

Veronderstel, dat door een waarnemer geen onderscheid wordt gemaakt tussen de elementen, die de rangnummers 3 t/m 6 moeten dragen. Dan wordt als rangnummer van ieder van deze elementen het gemiddelde van de rangnummers $\frac{1}{4}(3 + 4 + 5 + 6) = 4\frac{1}{2}$ gebruikt.

Daar het maximum van S nu verandert, moeten wij een correctie op de formule voor W toepassen. Deze vindt men in [1] (pg.82) en [2] (pg. 28-30). Eveneens veranderen dan de formules voor de χ^2-benadering ([1] pg. 86, [2] pg. 37) en voor de z-benadering ([1] pg. 86 [2] pg. 34), doch deze correcties zijn van weinig betekenis, tenzij het aantal gelijken groot is.

Literatuur: [1] M.G.Kendall, Rank correlation methods, London 1948, Hoofdstuk 6, pag. 80.

Tabel van de verdelingsfunctie van S voor:

n = 3	m = 2 t/m 10
n = 4	m = 2 t/m 6
n = 5	m = 3

op pag. 146-149.

Tabel van de waarden van S, waarvan de overschrijdingskansen onder de hypothese H_0 gelijk zijn aan 0,05 of 0,01, berekend met behulp van de z-benadering voor:

n = 3	m = 8,9,10,12,14,15,16,18,20
n = 4	m = 4,5,6,8,10,15,20
n = 5 t/m 7	m = 3,4,5,6,8,10,15,20

op pag. 150.

[2] Ph.van Elteren, Methode der m rangschikkingen, Cursus "Parametervrije Methoden", Hoofdstuk II, Rapport S 59, Mathematisch Centrum (1951), Blz. 18-45.