STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

ZW 1952 - 022

Voordracht in de serie
Actualiteiten

# Een generalisatie van de methode van m rangschikkingen

Ph. van Elteren

29 november 1952

1952

Rapport Z.W. 1952-022

Voordracht door Ph.van Elteren in de serie
Actualiteiten op 29 November 1952.

# Een generalisatie van de methode van m rangschikkingen.

## 1. Inleiding.

De door ons te behandelen toets kan enerzijds beschouwd worden als een generalisatie van de methode van m rangschikkingen van FRIEDMAN (zie [3][1] en [5] hoofdstuk 6 en 7), anderzijds als een toepassing van de principes van de toets van WILCOXON (zie [8] en [11]).

### 1.1 Methode der m rangschikkingen van Friedman.

Wij beschouwen m waarnemers $P_1, \ldots, P_m$. Iedere waarnemer rangschikt n objecten $O_1, \ldots, O_n$ en drukt het resultaat uit in rangnummers. Zij $r_{\mu\nu}$ het rangnummer door $P_\mu$ toegekend aan $O_\nu$. Indien wij de mogelijkheid van gelijke waarnemingen voorlopig buiten beschouwing laten, zal de rij $r_{\mu 1}, \ldots, r_{\mu n}$ een permutatie zijn van de getallen $1, \ldots, n$.

Wij verenigen de m rangschikkingen in schema (1):

(1)

| | $O_1$ | $O_2$ | $\ldots$ | $O_n$ |
|---|---|---|---|---|
| $P_1$ | $r_{11}$ | $r_{12}$ | $\ldots$ | $r_{1n}$ |
| $P_2$ | $r_{21}$ | $r_{22}$ | $\ldots$ | $r_{2n}$ |
| $\vdots$ | $\vdots$ | | | |
| $P_m$ | $r_{m1}$ | $r_{m2}$ | $\ldots$ | $r_{mn}$ |
| | $s_1$ | $s_2$ | $\ldots$ | $s_n$ |

waarin $s_\nu \overset{\text{def}}{=\!=} \sum_{\mu=1}^{m} r_{\mu\nu}$.

De nulhypothese $H_0$ van de toets houdt in:

1°. Iedere rangschikking is een aselecte[2] keuze uit de verzameling van alle permutaties der getallen $1, \ldots, n$.

2°. De rangschikkingen zijn statistisch onafhankelijk.

De toetsingsgrootheid is:

$$\underline{S} \overset{\text{def}}{=\!=} \sum_{\nu=1}^{n} \left\{ \underline{s}_\nu - \tfrac{1}{2}m(n+1) \right\}^2 . \tag{2}$$

--------------------

1) Cijfers tussen teksthaken [ ] verwijzen naar de literatuurlijst.

2) Het woord "aselect" voor Engels "random" is ontleend aan de terminologie van Prof. Dr D.van Dantzig.

De verdeling van $\underline{S}$ onder de hypothese $H_o$ is voor kleine waarden van m en n exact bekend. Voor grote waarden van m kunnen wij gebruik maken van het feit dat:

(3) $$\chi^2 = \frac{12\underline{S}}{mn(n+1)},$$

voor $m \to \infty$ asymptotisch verdeeld is volgens de zogenaamde $\chi^2$-verdeling met n-1 vrijheidsgraden.

1.2 Toets van Wilcoxon.

Wij beschouwen twee steekproeven:

$$x_1, \ldots, x_{n_1} \quad \text{en} \quad y_1, \ldots, y_{n_2}$$

en wensen de hypothese $H_o$ te toetsen, dat dit twee reeksen van onafhankelijke waarnemingen uit eenzelfde verdeling zijn.

Wij rangschikken daartoe de grootheden $x_1, \ldots, x_{n_1}$, $y_1, \ldots, y_{n_2}$ in één rij naar opklimmende grootte. Wij tellen het aantal keren, dat er in deze rij een y-waarneming vóór een x-waarneming staat. Dit aantal wordt aangeduid met de letter U en is de toetsingsgrootheid van de toets van WILCOXON.

De verdeling van $\underline{U}$ onder de hypothese $H_o$ is voor kleine waarden van $n_1$ en $n_2$ exact bekend. Voor grote waarden van $n_1$ en $n_2$ kunnen wij gebruik maken van het feit dat:

(4) $$\frac{\underline{U} - \frac{1}{2}n_1 n_2}{\sqrt{\frac{1}{12}n_1 n_2(n_1+n_2+1)}}$$

asymptotisch normaal verdeeld is met gemiddelde 0 en spreiding 1, indien $n_1 \to \infty$, $n_2 \to \infty$ en $\frac{n_1}{n_2}$ en $\frac{n_2}{n_1}$ begrensd zijn.[3] Hierbij is aangenomen dat er in de rij $x_1, \ldots, x_{n_1}$, $y_1, \ldots, y_{n_2}$ geen gelijke waarnemingen voorkomen.

Indien er wel gelijke waarnemingen voorkomen, levert de definitie van U moeilijkheden als bv. $x_i = y_j$. Wij kennen aan U voor ieder dergelijk paar een bijdrage $\frac{1}{2}$ toe. Tevens moet voor dit geval de nulhypothese iets gewijzigd worden. Wij zullen deze gewijzigde nulhypothese hieronder bespreken. De verdeling van $\underline{U}$ onder $H_o$ is, noch voor kleine $n_1$ en $n_2$, noch asymptotisch voor grote waarden van $n_1$ en $n_2$ precies bekend. Doch indien de groepen gelijke waarnemingen niet te groot zijn, mag men aannemen dat de verdeling van $\underline{U}$ niet veel zal afwijken van de verdeling van de overeenkomstige $\underline{U}$ zonder gelijken.

Wel exact bekend is de "variantie" $\sigma^2$ van $\underline{U}$ onder $H_o$, als in de rij $x_1, \ldots, x_{n_1}$, $y_1, \ldots, y_{n_2}$ g groepen gelijken voorkomen met $t_1, t_2, \ldots, t_g$ waarnemingen. Door $t_\gamma = 1$ ook toe te laten bereiken wij dat: $\sum_{\gamma=1}^{g} t_\gamma = n_1 + n_2$.

---

3) Deze laatste voorwaarde is overbodig (A. Benard).

Uit een resultaat van J.Hemelrijk (zie [4]) kan gemakkelijk worden afgeleid dat geldt:

$$\sigma^2 = \frac{n_1 n_2 \{(n_1+n_2)^3 - \sum_{\gamma=1}^{g} t_\gamma^3\}}{12(n_1+n_2)(n_1+n_2-1)} \quad (5)$$

Opm.: Men kan de gereduceerde waarde $\tilde{U} = U - \frac{1}{2} n_1 n_2$ van de toetsingsgrootheid van Wilcoxon ook berekenen door aan de waarnemingen $x_1, \ldots, x_{n_1}$, $y_1, \ldots, y_{n_2}$ naar opklimmende grootte rangnummers toe te kennen en deze rangnummers te verminderen met hun gemiddelde $\frac{1}{2}(n_1+n_2+1)$. $\tilde{U}$ is dan gelijk aan de som van de aldus gereduceerde rangnummers der x-waarnemingen. Dit geldt ook als er gelijke rangnummers voorkomen.

1.3 Parametervrije k steekproeventoets.

De toets van Wilcoxon is door Kruskal [7], Rijkoort [9], Terpstra [10] en Krishna Iyer [6] gegeneraliseerd tot een toets voor k steekproeven:

$$\begin{array}{l} x_{11}, \ldots, x_{1n_1}; \\ x_{21}, \ldots, x_{2n_2}; \\ \vdots \\ x_{k1}, \ldots, x_{k,n_k}. \end{array}$$

Wij kennen nu aan al deze waarnemingen rangnummers toe naar opklimmende grootte, en reduceren deze rangnummers (d.w.z. wij verminderen de rangnummers met $\frac{1}{2}(N+1)$ als $N = \sum_{\varkappa=1}^{k} n_\varkappa$). Zij nu $\tilde{U}_\varkappa$ de som van de gereduceerde rangnummers van de steekproef $\varkappa$, dan is de toetsingsgrootheid van de k-steekproeventoets

$$\chi^2 = \frac{12}{N(N+1)} \sum_{\varkappa=1}^{k} \frac{\tilde{U}_\varkappa^2}{n_\varkappa}. \quad (6)$$

Indien de hypothese $H_0$, inhoudende dat de k steekproeven uit dezelfde verdeling komen, vervuld is, zal bovengenoemde toetsingsgrootheid asymptotisch verdeeld zijn volgens de $\chi^2$-verdeling met k-1 vrijheidsgraden, als de aantallen $n_1, \ldots, n_k$ naar oneindig gaan.

## 2. Principe van de gegeneraliseerde methode van m rangschikkingen.

### 2.1 Schema der waarnemingen.

Wij kunnen ons schema (1) gebaseerd denken op een schema van waarnemingen, waarbij waarnemer $P_\mu$ één waarneming $x_{\mu\nu}$ gedaan heeft van object $O_\nu$. In het gegeneraliseerde schema wordt nu toegestaan, dat op een of meer plaatsen $(\mu,\nu)$ de waarneming ontbreekt, óf dat er in plaats van $x_{\mu\nu}$

een steekproef van waarnemingen $x_{\mu\nu\varkappa}$ komt. Het aantal waarnemingen corresponderend met plaats $(\mu,\nu)$ duiden wij aan met $k_{\mu\nu}$.

Iedere rij van het schema bestaat nu uit n of minder steekproeven. Wij gaan nu in ieder van de rijen te werk als beschreven in 1.3. De som van de gereduceerde rangnummers corresponderend met de waarnemingen op plaats $(\mu,\nu)$ duiden wij aan met $\tilde{u}_{\mu\nu}$. Zijn er geen waarnemingen op plaats $(\mu,\nu)$ dan is per definitie $\tilde{u}_{\mu\nu} = 0$.

De resultaten verenigen wij in het volgende schema:

(7)

| | $O_1$ | $O_2$ | ..... | $O_n$ |
|---|---|---|---|---|
| $P_1$ | $\tilde{u}_{11}$ | $\tilde{u}_{12}$ | .... | $\tilde{u}_{1n}$ |
| $P_2$ | $\tilde{u}_{21}$ | $\tilde{u}_{22}$ | .... | $\tilde{u}_{2n}$ |
| ⋮ | | | | |
| $P_m$ | $\tilde{u}_{m1}$ | $\tilde{u}_{m2}$ | .... | $\tilde{u}_{mn}$ |
| | $\tilde{u}_1$ | $\tilde{u}_2$ | .... | $\tilde{u}_n$ |

,

waarin $\tilde{u}_{\nu} \stackrel{\text{def}}{=} \sum_{\mu=1}^{m} \tilde{u}_{\mu\nu}$.

## 2.2 Nulhypothese.

De hypothese $H_0$, die we wensen te toetsen, komt hierop neer, dat in ieder van de rijen de steekproeven uit dezelfde verdeling komen (deze verdeling behoeft niet voor alle rijen identiek te zijn) en de rijen onderling onafhankelijk zijn.

De hypothese $H_0$ kan als volgt nader gespecificeerd worden: met de waarnemingen van $P_\mu$ corresponderen $k_\mu = \sum_{\nu=1}^{n} k_{\mu\nu}$ gereduceerde rangnummers.

De verzameling van deze rangnummers is op een of andere wijze verdeeld in klassen corresponderend met de objecten $O_1,\dots,O_n$.

De hypothese $H_0$ houdt nu in, dat alle verdelingen van de rangnummers over de klassen even waarschijnlijk zijn. Verder moeten de verdelingen van de rangnummers over de klassen in de verschillende rijen onderling volledig onafhankelijk zijn.

## 2.3 Varianties en covarianties der kolomtotalen.

Het is op vrij elementaire wijze na te gaan dat onder $H_0$ geldt:

(8) $$\mathcal{E}\, \underline{\tilde{u}}^2_{\mu\nu} = - k_{\mu\nu}(k_\mu - k_{\mu\nu})\, K_\mu$$

en

(9) $$\mathcal{E}\, \underline{\tilde{u}}_{\mu\nu}\, \underline{\tilde{u}}_{\mu\nu'} = - k_{\mu\nu} k_{\mu\nu'}\, K_\mu \qquad (\nu \neq \nu') ,$$

als

(10) $$K_\mu \stackrel{\text{def}}{=} \frac{k_\mu^3 - \sum_{\gamma=1}^{g_\mu} t^3_{\mu\gamma}}{12\, k_\mu(k_\mu - 1)} ,$$

waarbij

$$k_\mu = \sum_{\nu=1}^{n} k_{\mu\nu},$$

$t_{\mu\gamma}$ = grootte van een groep gelijke rangnummers in rij $\mu$,

$g_\mu$ = aantal groepen gelijke rangnummers in rij $\mu$.

Vanwege de onafhankelijkheid der rijen kennen wij dan ook onmiddellijk de varianties en covarianties van de kolomtotalen:

(11) $$\sigma_\nu^2 \overset{\text{def}}{=\!=} \mathcal{E}\,\underline{\tilde{u}}^2 = \sum_{\mu=1}^{m} \mathcal{E}\underline{\tilde{u}}_\nu^2 \ ,$$

(12) $$\sigma_{\nu\nu'} \overset{\text{def}}{=\!=} \mathcal{E}\,\underline{\tilde{u}}_\nu\,\underline{\tilde{u}}_{\nu'} = \sum_{\mu=1}^{m} \mathcal{E}\,\underline{\tilde{u}}_{\mu\nu}\,\underline{\tilde{u}}_{\mu\nu'} \ .$$

2.4 Toetsingsgrootheid.

De toetsingsgrootheid kan gedefinieerd worden als de absolute waarde van een breuk, waarvan de teller wordt afgeleid van de determinant

$$\begin{vmatrix} \sigma_{11} & \cdots & \sigma_{1n} & \underline{\tilde{u}}_1 \\ \vdots & & \vdots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} & \underline{\tilde{u}}_n \\ \underline{\tilde{u}}_1 & \cdots & \underline{\tilde{u}}_n & 0 \end{vmatrix} ,$$

door een willekeurige rij en een willekeurige kolom (mits niet de laatste rij of de laatste kolom) weg te laten en de noemer wordt afgeleid van de determinant

$$\begin{vmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{vmatrix} ,$$

door een willekeurige rij en een willekeurige kolom weg te laten.

Het kan gemakkelijk aangetoond worden dat alle aldus verkregen grootheden gelijk zijn.

Noodzakelijk voor de existentie van deze toetsingsgrootheid is, dat de rang van de matrix

$$S = \begin{Vmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{Vmatrix}$$

gelijk is aan n-1.

Het kan aangetoond worden dat, zo de rang van S gelijk is aan $n-r < n-1$ men de rij getallen $I(1,\dots,n)$ kan verdelen in juist r deelrijen $I_1,\dots,I_r$ zó dat indien $\nu \in I_\rho$ en $\nu' \in I_{\rho'}$ en $\rho \neq \rho'$ geldt $\sigma_{\nu\nu'} = 0$.

Indien b v. r=2, kan matrix S door de volgorde van rijen en kolommen geschikt te kiezen in de volgende gedaante worden gebracht:

$$\begin{Vmatrix} \sigma_{11} & \dots & \sigma_{1k} & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & S_1 & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ \sigma_{k1} & \dots & \sigma_{kk} & 0 & \dots & 0 \\ 0 & \dots & 0 & \sigma_{k+1,k+1} & \dots & \sigma_{k+1\,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & S_2 & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \sigma_{n\,k+1} & \dots & \sigma_{nn} \end{Vmatrix}$$

In dit geval kan de toets worden toegepast op ieder der schema's corresponderend met de matrices $S_1$ en $S_2$ afzonderlijk.

## 2.5 Centrale limietstelling voor stochastische vectoren.

Indien de rang van S gelijk is aan n-1 en aan nog enkele andere vrij algemene voorwaarden voldaan is, is de toetsingsgrootheid onder de hypothese $H_0$ voor grote waarden van m asymptotisch verdeeld volgens een $\chi^2$-verdeling met n-1 vrijheidsgraden.

Dit kan worden bewezen met behulp van de centrale limietstelling voor de som van stochastische vectoren.

Wij willen ons hier beperken tot de vorm waarin de stelling door Cramèr (zie [1]) is gegeven.

Zij $\underline{X}_1$, $\underline{X}_2, \dots$ een rij onafhankelijke stochastische vectoren in $R_l$, zodat iedere $X_k$ de verdelingsfunctie $P_k$ bezit met gemiddelden 0 en begrensde varianties en covarianties $\mu^{(k)}_{\lambda\lambda'}$. Indien nu de volgende voorwaarden vervuld zijn:

$$\lim_{k\to\infty} \frac{1}{k}\sum_{\varkappa=1}^{k} \mu^{(\varkappa)}_{\lambda\lambda'} = \mu_{\lambda\lambda'}, \tag{13}$$

waarbij niet alle $\mu_{\lambda\lambda'}$ gelijk zijn aan nul en

$$\lim_{k\to\infty} \frac{1}{k}\sum_{\varkappa=1}^{k} \int_{|X|\geqq \varepsilon\sqrt{k}} |X|^2 dP_\varkappa = 0, \tag{14}$$

waarbij

$$X = (x_1,\dots,x_l),\quad |X| = \sqrt{\sum_{\lambda=1}^{l} x_\lambda^2}.$$

Dan convergeert de verdelingsfunctie van de variabele $\frac{1}{\sqrt{k}} \sum_{\varkappa=1}^{k} \underline{X}_{\varkappa}$ naar de normale verdelingsfunctie met gemiddelden 0 en varianties en covarianties $\mu_{\lambda\lambda'}$. Het aantal dimensies van deze verdeling is gelijk aan de rang van de matrix $M = \| \mu_{\lambda\lambda'} \|$.

Uit de stelling van Cramèr volgt direct dat de grootheid $\underline{X}'M^{-1}\underline{X}$ ($\underline{X} = \lim_{k \to \infty} \frac{1}{\sqrt{k}} \sum_{\varkappa=1}^{k} \underline{X}_{\varkappa}$) een $\chi^2$-verdeling heeft met v vrijheidsgraden, waarbij v de rang van M is.

2.6 Asymptotische verdeling van de toetsingsgrootheid.

Wij passen de centrale limietstelling op ons geval toe door k=m, l=n te nemen en $\underline{X}_{\varkappa}$ te vervangen door $\underline{\tilde{U}}_{\mu} = (\underline{\tilde{u}}_{\mu 1}, \ldots, \underline{\tilde{u}}_{\mu n})$. Er geldt $\mathcal{E}\underline{\tilde{u}}_{\mu\nu}=0$ en $\mathcal{E}\underline{\tilde{u}}_{\mu\nu}\underline{\tilde{u}}_{\mu\nu'}$ is begrensd voor alle $\mu$, $\nu$ en $\nu'$ als alle $k_{\mu\nu}$ begrensd zijn. In dat geval zijn de stochastische vectoren $\underline{\tilde{U}}_{\mu}$ alle begrensd, waaruit zeer gemakkelijk volgt, dat voldaan is aan (14). De eis (13) is de onaangenaamste; alleen bij schema's die een bepaalde regelmaat vertonen zal men deze kunnen verifiëren. Is hieraan voldaan dan geldt dat de in 2.4 gedefinieerde toetsingsgrootheid asymptotisch overeenkomt met $\underline{X}'M^{-1}\underline{X}$, mits de rang van matrix M (hier: $\mu_{\nu\nu'} = \lim_{m \to \infty} \frac{1}{m} \sigma_{\nu\nu'}$) gelijk is aan n-1.

Is ook aan de laatste eis voldaan, dan is de in 2.4 gedefinieerde toetsingsgrootheid (in het vervolg aan te duiden met $\underline{\chi}^2$) asymptotisch verdeeld volgens een $\chi^2$-verdeling met n-1 vrijheidsgraden.

3. Bijzondere gevallen.

In zijn meest algemene vorm is de toets dermate bewerkelijk dat zij voor practische toepassing niet zeer geschikt is. Er worden hier enige gevallen genoemd, waarbij de toetsingsgrootheid minder gecompliceerd is.

3.1 De toetsingsgrootheid $\underline{\chi}^2$ is dan en slechts dan te schrijven als een lineaire combinatie van de kwadraten $\underline{\tilde{u}}_1^2, \ldots, \underline{\tilde{u}}_n^2$, indien er positieve getallen $c_\nu$ bestaan zodat:

(15) $\sigma_{\nu\nu'} = -c_\nu c_{\nu'}$ ($\nu = 1, \ldots, n$; $\nu' = 1, \ldots, n$; $\nu' \neq \nu$)

De toetsingsgrootheid wordt dan

$$\chi^2 = \frac{1}{c} \sum_{\nu=1}^{n} \frac{\underline{\tilde{u}}_\nu^2}{c_\nu}, \text{ waar } c = \sum_{\nu=1}^{n} c_\nu .$$

Voorbeelden:

In het algemene schema met n = 3 is aan voorwaarde (15) altijd voldaan ($c_1 = -\sqrt{\frac{-\sigma_{12}\sigma_{31}}{\sigma_{23}}}$, enz.); de toetsingsgrootheid wordt dan:

(16) $$\underline{\chi}^2 = -\frac{\sigma_{23}\underline{\tilde{u}}_1^2 + \sigma_{31}\underline{\tilde{u}}_2^2 + \sigma_{12}\underline{\tilde{u}}_3^2}{\sigma_{12}\sigma_{23} + \sigma_{31}\sigma_{12} + \sigma_{23}\sigma_{31}} .$$

b) $k_{\mu\nu} = a_\mu b_\nu$ ($\mu = 1,\ldots,m$; $\nu = 1,\ldots,n$)

dan geldt:

(17) $$\underline{\chi}^2 = \frac{\sum_{\nu=1}^{n} \frac{\underline{\tilde{u}}_\nu^2}{b_\nu}}{b \sum_{\mu=1}^{m} a_\mu^2 K_\mu} \qquad \left\{ \begin{array}{l} b = \sum_{\nu=1}^{n} b_\nu . \\ K_\mu \text{ is gedefinieerd door (10).} \end{array} \right.$$

(Neem $c_\nu = b_\nu \sqrt{\sum_{\mu=1}^{m} a_\mu^2 K_\mu}$ ).

3.2 De toetsingsgrootheid $\underline{\chi}^2$ is dan en slechts dan symmetrisch in alle $\underline{\tilde{u}}_\nu$ indien alle covarianties $\sigma_{\nu\nu'}$ gelijk zijn.

Er geldt dan, indien $\sigma = \sigma_\nu$ en $\underline{S} = \sum_{\nu=1}^{n} \underline{\tilde{u}}_\nu^2$ ,

(18) $$\underline{\chi}^2 = \frac{n-1}{n} \frac{\underline{S}}{\sigma^2} .$$

Voorbeelden:

a) In het algemene schema met n=2 is uiteraard aan bovengenoemde voorwaarde voldaan, want er is slechts één covariantie. Er geldt $\tilde{u}_1 = -\tilde{u}_2 = \tilde{u}$ dus:

(19) $$\underline{\chi}^2 = \frac{\underline{\tilde{u}}^2}{\sigma^2} .$$

Deze toetsingsgrootheid is onder $H_o$ asymptotisch verdeeld volgens een $\chi^2$-verdeling met 1 vrijheidsgraad, dus is

(20) $$\underline{\chi} = \frac{\underline{\tilde{u}}}{\sigma}$$

asymptotisch normaal (0,1) verdeeld.

Bijzondere gevallen hiervan: tekentoets (alle $k_\mu = 2$) en toets van Wilcoxon (m=1).

b) Indien we in het algemene schema $k_{\mu\nu} = 1$ kiezen ($\mu = 1,\ldots,m$; $\nu = 1,\ldots,n$) krijgen we het schema van Friedman terug. Volgens (18) wordt hier de toetsingsgrootheid bij aanwezigheid van gelijke rangnummers:

(21) $$\underline{\chi}^2 = \frac{12\underline{S}}{mn(n+1) - T} ,$$

waarin: $$T = \frac{1}{n-1} \sum_{\mu=1}^{m} \sum_{\gamma=1}^{g_\mu} (t_{\mu\gamma}^3 - t_{\mu\gamma}).$$

Deze formule komt overeen met de formule vermeld door Kendall (zie [5], page 96).

c) Schema van Durbin:

Alle $k_{\mu\nu}$ zijn 0 of 1, doch zó dat:

1. $$k_{\mu} = \sum_{\nu=1}^{n} k_{\mu\nu} = k \qquad (\mu=1,\ldots,m) \qquad \text{en}$$

2. $$\sum_{\mu=1}^{m} k_{\mu\nu}k_{\mu\nu'} = \lambda \quad (\nu = 1,\ldots,n;\ \ \nu' = 1,\ldots,n;\ \ \nu' \neq \nu).$$

D.w.z. het aantal keren dat zowel in kolom $\nu$ als in kolom $\nu'$ in dezelfde rij een waarneming voorkomt, is voor alle $\nu$ en $\nu'$ gelijk aan $\lambda$ .

Indien er geen gelijke rangnummers voorkomen, geldt nu:

$$\sigma_{\nu\nu'} = \frac{1}{12}\lambda(k+1) \quad \text{en:}$$

(22) $$\underline{\chi}^2 = \frac{12\underline{S}}{n\lambda(k+1)} .$$

Literatuur:

[1] Cramèr, H., Random variables and probability distributions, Cambridge tracts in mathematics No 36, Cambridge 1937.

[2] Durbin, J., Incomplete blocks in ranking experiments, British Journal of Psychology 4 (1951), 85-90.

[3] Friedman, M., The use of ranks to avoid the assumption of normality implicit in the analysis of variance, Journal of the American Statistical Association 32 (1935), 675-699.

[4] Hemelrijk, J., Note on Wilcoxon's two sample test when ties are present, Annals of Mathematical Statistics 23 (1952), 133-135.

[5] Kendall, M.G., Rank Correlation methods, London (1948).

[6] Krishna Iyer, P.V., A non parametric method of testing k samples, Nature 167 (1951), 33.

[7] Kruskal, W.H., A non parametric analogue based upon ranks of one way analysis of variance, Abstract: Annals of Mathematical Statistics 23 (1952), 140.

[8] Mann, H.B. and Whitney, D.R., On a test whether one of two random variables is stochastically larger than the other, Annals of Mathematical Statistics 18 (1947), 50-60.

[9] Rijkoort, P.J., A generalisation of Wilcoxon's test, Indag. Math. 14 (1952), 394-403.

[10] Terpstra, T.J., A non parametric k-sample test and its connection with the H-test, Math. Centrum, Stat. Afdeling. Rapport S 92 (VP 2)

[11] Wilcoxon, F., Individual comparisons by ranking methods, Biometrics Bulletin 1 (1945), 80-83.