HOOFSTUK IV.

DIE PROBLEEM VAN m-RANGSKIKKINGS.

4.1. ALGEMENE INLEIDING.

Die probleem van m-rangskikkings is voorgestel en bespreek deur FRIEDMAN(1937) en is breedvoerig behandel deur KENDALL(1948).

KENDALL(1948) beskou m „waarnemers". Elke waarnemer rangskik k gegewe „items" volgens 'n bepaalde eienskap en ken dan rangnommers toe waarby r_{ij} $(i=1,2,\ldots,m;$ $j=1,2,\ldots,k)$ die rangnommer is van item x_{ij}. Omdat elke waarnemer opnuut, onafhanklik van die ander waarnemers, rangnommers toeken, het ons:

$$\sum_{j=1}^{k} r_{ij} = \tfrac{1}{2}k(k+1) \text{ vir alle } i=1,2,\ldots,m.$$

Skematies word bogenoemde soos volg voorgestel:

TABEL 4.1.

Skematiese voorstelling van m-rangskikkings.

		Items			
		1	2	. . .	k
Waarnemers:	1	r_{11}	r_{12}	. . .	r_{1k}
	2	r_{21}	r_{22}	. . .	r_{2k}
	.	. . .	. . .	. . .	. . .
	.				
	m	r_{m1}	r_{m2}	. . .	r_{mk}
Totaal		u_1^*	u_2^*		u_k^*

Die kolomtotale $u_j^* = \sum_{i=1}^{m} r_{ij}$; $j=1,2,\ldots,k$, word met mekaar vergelyk en gee 'n aanduiding van die ooreenstemming wat daar tussen die verskillende waarnemers bestaan. Indien alle rangskikkings identies is (volkome ooreenstemming tussen waarnemers), sal die totale u_j^* die waardes m, 2m,...,km aanneem (nie noodwendig in dié volgorde nie). Indien daar egter min verskil tussen die items bestaan, sal die waarnemers beswaarlik tussen hulle kan onderskei en mag die totale u_j^* betreklik min van mekaar verskil. ('n Item wat deur een waarnemer vroeg in die rangskikking geplaas word, kan deur 'n ander waarnemer baie later geplaas word. Sien byvoorbeeld KENDALL(1948) pp. 80-81).

Die nulhipotese H_0 wat ondersoek word, veronderstel dat elke rangskikking, onafhanklik van die ander, ewekansig gekies is uit die versameling van alle moontlike permutasies van die getalle 1,2,...,k.

FRIEDMAN(1937) het as toetsingsgrootheid voorgestel

(4.1) $\chi^2 = 12S^*/[mk(k+1)]$ waar

(4.2) $S^* = \sum_{j=1}^{k} [u_j^* - \frac{1}{2}m(k+1)]^2$

en het bewys dat χ^2 onder H_0 asimptoties 'n χ^2-verdeling met (k-1) g.v.v. besit as $m \to \infty$.

In die geval waar gelyke waarnemings voorkom, het BENARD en VAN ELTEREN(1953) 'n wysiging aan χ^2 aangebring, naamlik

(4.3) $\chi^2 = 12S^*/[mk(k+1) - T^*]$ met T^* gedefinieer in (4.73). (Sien §4.4.3.2 vir die voorwaardes waaronder χ^2 se asimptotiese verdeling afgelei is).

Bogenoemde probleem is deur BENARD en VAN ELTEREN(1953) uitgebrei na die geval waar die aantal waarnemings x_{ijh} van die item x_{ij} (d.w.s. die aantal waarnemings in die „sel" (i,j)), enige nie-negatiewe heelgetal k_{ij} is. Die i[de] rangskikking bevat dan

(4.4) $k_i = \sum_{j=1}^{k} k_{ij}$ [1]) waarnemings wat weer volgens

---

1). Tensy anders vermeld, deurloop i,i' in hierdie hoofstuk die waardes 1,2,...,m; j en j' die waardes 1,2,...,k; p_i die waardes $1,2,\ldots,k_i$; h die waardes $1,2,\ldots,k_{ij}$ en ζ,ζ' die waardes 1,2,...,(k-1).

'n bepaalde eienskap gerangskik word. Daar is dus altesame

(4.5) $N = \sum_i k_i = \sum_i \sum_j k_{ij}$ waarnemings waarby alle k_i en k_{ij} eindig is, tensy anders gespesifiseer.

BENARD en VAN ELTEREN(1953) [2)] definieer

(4.6) $$\tilde{u}_{ij} = \begin{cases} \sum_h [r_{ijh} - \frac{1}{2}(k_i+1)] & \text{indien } k_{ij}>0 \\ 0 & \text{indien } k_{ij}=0 \end{cases}$$

waarby r_{ijh} die rangnommer is van die waarneming x_{ijh} in die i^{de} rangskikking. (Rangnommers word opnuut in elke ry toegeken, onafhanklik van die ander rye). B+v.E(1953) neem dan 'n funksie van die kolomtotale $\tilde{u}_j$ as toetsingsgrootheid, waar $\tilde{u}_j = \sum_i \tilde{u}_{ij}$ (sien §4.4.3.1).

In hierdie hoofstuk word 'n veralgemeende toetsingsgrootheid vir die probleem van m-rangskikkings onder H_0 voorgestel en 'n uitbreiding na 'n meer algemene geval behandel.

4.2. 'n MEER ALGEMENE BENADERING.

4.2.1. Inleiding.

Laat $\Psi_{k_i}(\delta)$, vir alle waardes van k_i, 'n willekeurige funksie wees wat gedefinieer en eindig is vir iedere δ op die ope interval (0,1) by alle waardes van k_i en monotoon stygend (of monotoon dalend) in sy argument δ by vaste k_i. (Sien ook hoofstuk II §2.2.1).

Beskou 'n funksie van die rangnommers, naamlik $\Psi_{k_i}[r_{ijh}/(k_i+1)] = \Psi_{k_i}(\delta_{ijh})$ waar

(4.7) $\delta_{ijh} = r_{ijh}/(k_i+1)$.

Let op dat $1 \leq r_{ijh} \leq k_i$ vir alle $j=1,2,\ldots,k$ en $h=1,2,\ldots,k_{ij}$.

---

2). Voortaan word na hierdie artikel verwys as B+v.E(1953).

Opmerking.

Voortaan skryf ons kortweg $\Psi_i(\delta)$ in plaas van $\Psi_{k_i}(\delta)$, behalwe in gevalle waar $\Psi_{k_i}(\delta)$ se afhanklikheid van k_i beklemtoon moet word.

4.2.2. Gelyke Waarnemings (Knope).

In die geval van gelyke waarnemings binne 'n bepaalde rangskikking (ry), word dieselfde prosedure gevolg as in hoofstuk II. Ten opsigte van notasie is die volgende aanpassing nodig (alleen die geval van gemiddelde range word bespreek):

Dui, vir die i^{de} ry, die gerangskikte stel x_{ijh}-waardes aan deur $z_{i1}, z_{i2}, \ldots, z_{ik_i}$ waarby $z_{i1} \leq z_{i2} \leq \ldots \leq z_{ik_i}$. Gestel daar kom 'n knoop van lengte g_{α_i} voor. Aan elkeen van die gelyke waarnemings

z_{i,ν_i+1} , z_{i,ν_i+2} , ... , $z_{i,\nu_i+g_{\alpha_i}}$ word dieselfde rang toegeken, naamlik:

(4.8) $$g_{\alpha_i}^{-1} \sum_{\mu_i} (\nu_i + \mu_i) = \nu_i + \tfrac{1}{2}(g_{\alpha_i} + 1)$$ [3].

Op ooreenkomstige wyse word die funksie $\Psi_i(\delta_{ip_i})$, waar

(4.9) $\delta_{ip_i} = r_{ip_i}/(k_i+1)$ en r_{ip_i} die rangnommer is van z_{ip_i} in die i^{de} rangskikking, gedefinieer deur

(4.10) $$\Psi_i(\delta_{ip_i}) = g_{\alpha_i}^{-1} \sum_{\mu_i} \Psi_i[(\nu_i+\mu_i)/(k_i+1)]$$

vir alle p_i waarvoor $z_{ip_i} = z_{i,\nu_i+1}$.

Gestel nou daar kom onder die z_{ip_i} 's λ_i knope van lengtes $g_{i1}, g_{i2}, \ldots, g_{i\lambda_i}$ voor. Stel

(4.11) $G_{\alpha_i} = g_{i1} + g_{i2} + \ldots + g_{i\alpha_i}$ waar $\alpha_i \in \{1, 2, \ldots \lambda_i\}$ en $G_0 = 0$.

---

3). Tensy anders gespesifiseer, deurloop α_i die waardes $1, 2, \ldots, \lambda_i$ en μ_i die waardes $1, 2, \ldots, g_{\alpha_i}$.

Definieer

$$(4.12)\quad a_{ip_i} = g_{\alpha_i}^{-1} \sum_{\mu_i} \Psi_i[(G_{\alpha_i-1} + \mu_i)/(k_i+1)] \quad \text{wanneer}$$

$$G_{\alpha_i-1} + 1 \leq p_i \leq G_{\alpha_i} .$$

Stel

$$(4.13)\quad \sigma_{ai}^2 = k_i^{-1} \sum_{p_i} (a_{ip_i} - \bar{\Psi}_i)^2 \quad \text{waar}$$

$$(4.14)\quad \bar{\Psi}_i = k_i^{-1} \sum_{p_i} a_{ip_i}$$

$$= k_i^{-1} \sum_j \sum_h \Psi_i(\delta_{ijh}).$$

4.2.3. Definisies.

Definieer

$$(4.15)\quad a_{ijh} = a_{ip_i} \text{ indien } x_{ijh} = z_{ip_i} .$$

Stel

$$(4.16)\quad u_{ij} = \begin{cases} \sum_h (a_{ijh} - \bar{\Psi}_i) & \text{indien } k_{ij} > 0 \\ 0 & \text{indien } k_{ij} = 0 \end{cases} \quad \text{en}$$

$$(4.17)\quad u_j = m^{-\frac{1}{2}} \sum_i u_{ij}$$

$$= m^{-\frac{1}{2}} \sum_i \sum_h (a_{ijh} - \bar{\Psi}_i).$$

Stel verder

$$(4.18)\quad k_{.j} = \sum_i k_{ij} .$$

Opmerking.

In hierdie behandeling van m-rangskikkings word alle rye weggelaat waarin alle rangnommers dieselfde is (een enkele knoop in die ry) of waarin $k_{ij} = 0$ vir enige $(k-1)$ waardes van j uit $1,2,\ldots,k$. Sodanige rye lewer geen bydrae tot die studie van bogenoemde probleem nie omdat dan geld $u_{ij} = 0$ vir alle j.

4.2.4. 'n Aantal Lemmas.

Lemma 4.1. Die verwagtingswaarde van u_j onder H_o word, vir vaste j, gegee deur

(4.19) $E(u_j | H_o) = 0$ [4].

Bewys: Op analoë wyse as in lemma 2.1 volg:

$$(4.20)\quad E(u_{ij}) = E[\sum_h (a_{ijh} - \bar{\Psi}_i)]$$
$$= E(\sum_h a_{ijh}) - E(\sum_h \bar{\Psi}_i)$$
$$= k_{ij}k_i^{-1} \sum_{p_i} a_{ip_i} - k_{ij}\bar{\Psi}_i$$
$$= 0 \text{ sodat}$$
$$E(u_j) = E(m^{-\frac{1}{2}} \sum_i u_{ij})$$
$$= m^{-\frac{1}{2}} \sum_i E(u_{ij})$$
$$= 0.$$

Lemma 4.2. Die variansie van u_j onder H_o word gegee deur

$$(4.21)\quad var(u_j) = m^{-1} \sum_i k_{ij}(k_i - k_{ij})K_i$$

waar

$$(4.22)\quad K_i = k_i^{-1}(k_i-1)^{-1} \sum_{p_i}(a_{ip_i} - \bar{\Psi}_i)^2$$
$$= (k_i-1)^{-1} \sigma_{ai}^2$$

waar σ_{ai}^2 in (4.13) gedefinieer is oor alle knope.

Bewys: $\sigma_{ij}^2 = var(u_{ij})$
$$= E(u_{ij}^2) \quad \text{omdat } E(u_{ij}) = 0$$
$$= E[\sum_h (a_{ijh} - \bar{\Psi}_i)]^2$$
$$= \binom{k_i}{k_{ij}}^{-1} \sum{}' [\sum_h (a_{ijh} - \bar{\Psi}_i)]^2$$

waar $\sum'$ die sommasie oor alle moontlike kombinasies van k_{ij} groothede uit k_i aandui, wat almal ewekansig is onder H_o, sodat

4). In hierdie hoofstuk word, tensy anders vermeld, onder die hipotese H_o gewerk. $E(u_j)$ byvoorbeeld dui dus $E(u_j | H_o)$ aan. In die gevalle waar knope voorkom, word voorwaardelik onder die gegewe stel waarnemings gewerk.

(4.23) $\sigma_{ij}^2 = k_{ij}(k_i - k_{ij})k_i^{-1}(k_i-1)^{-1}\sum_{p_i}(a_{ip_i} - \bar{\Psi}_i)^2$

netsoos in lemma 2.2

$= k_{ij}(k_i-k_{ij})K_i$ sodat

(4.24) $\sigma_j^2 = \mathrm{var}(u_j)$

$= E(u_j^2)$ omdat $E(u_j) = 0$ (lemma 4.1)

$= E(m^{-\frac{1}{2}} \sum_i u_{ij})^2$

$= m^{-1} \sum_i E(u_{ij}^2) + m^{-1} \sum_{i \neq i'}\sum E(u_{ij}.u_{i'j})$

$= m^{-1} \sum_i \sigma_{ij}^2 + 0$ omdat alle rangskikkings onderling onafhanklik is

$= m^{-1} \sum_i k_{ij}(k_i-k_{ij})K_i.$

Opmerking.

Uit $K_i = k_i^{-1}(k_i-1)^{-1}\sum_{p_i}(a_{ip_i} - \bar{\Psi}_i)^2$ volg dat

(4.25) $K_i \geq D > 0$ vir alle i omdat die rangnommers r_{ip_i} (vir vaste i) nie almal dieselfde is nie (opmerking §4.2.3); omdat $\Psi_i(\delta)$ monotoon stygend is in δ en omdat alle k_i eindig bly as $m \to \infty$.

Ook is

(4.26) $(k_i-k_{ij}) \geq 1$ vir enige i en j, want:

$k_i-k_{ij} = 0$ alleenlik indien

i) alle $k_{ij} = 0$

of ii) $k_{ij} > 0$, $k_{ij'} = 0$ vir alle $j' \neq j$.

Hierdie twee gevalle word deur die opmerking in §4.2.3 uitgesluit.

Dus

(4.27) $\sigma_{ij}^2 \geq k_{ij}D$ en uit (4.24):

(4.28) $\sigma_j^2 \geq Dm^{-1} \sum_i k_{ij}$

$= Dk_{.j}/m > 0.$

Lemma 4.3. Die kovariansie tussen u_j en $u_{j'}$ $(j' \neq j)$ onder H_0 word gegee deur

(4.29) $\operatorname{kov}(u_j, u_{j'}) = -\, m^{-1} \sum_i k_{ij} k_{ij'} K_i \,.$

Bewys: Netsoos in lemma 2.3 volg dat

$$(4.30)\quad \begin{aligned} \operatorname{kov}(u_{ij}, u_{ij'}) &= E(u_{ij}.u_{ij'}) \text{ omdat } E(u_{ij})=0 \text{ uit } (4.20) \\ &= -\, k_{ij}k_{ij'}\, k_i^{-1}(k_i-1)^{-1} \sum_{p_i} (a_{ip_i} - \bar{\Psi}_i)^2 \\ &= -\, k_{ij}k_{ij'}\, K_i \quad \text{sodat} \end{aligned}$$

$$\begin{aligned} \sigma_{jj'} &= \operatorname{kov}(u_j, u_{j'}) \\ &= E(u_j.u_{j'}) \text{ omdat } E(u_j) = 0 \text{ vir alle } j \\ &= E(m^{-\frac{1}{2}} \sum_i u_{ij}.m^{-\frac{1}{2}} \sum_{i'} u_{i'j'}) \\ &= E(m^{-1} \sum_i u_{ij}u_{ij'} + m^{-1} \sum_{i \neq i'} \sum u_{ij}u_{i'j'}) \\ &= m^{-1} \sum_i E(u_{ij}u_{ij'}) + 0 \\ &= -\, m^{-1} \sum_i k_{ij}k_{ij'}\, K_i. \end{aligned}$$

Opmerking.

Die uitdrukkings vir $\operatorname{var}(u_j)$ en $\operatorname{kov}(u_j, u_{j'})$ is analoog aan dié wat afgelei is deur B+v.E(1953) vergelyking (2.3.1).

4.3. DIE TOETSINGSGROOTHEID T.

4.3.1. 'n Limietverdeling onder H_0.

Stel

(4.31) $v_i = \sum_j c_j u_{ij}$ waar die c_j 's willekeurige eindige reële getalle is, dan is

$$(4.32)\quad \begin{aligned} E(v_i) &= E(\sum_j c_j u_{ij}) \\ &= \sum_j c_j E(u_{ij}) \\ &= 0 \text{ uit } (4.20) \quad \text{en} \end{aligned}$$

$$(4.33)\quad \begin{aligned} \sigma_{i.}^2 &= \operatorname{var}(v_i) \\ &= E(v_i^2) \\ &= E(\sum_j c_j u_{ij})^2 \end{aligned}$$

$$= E(\sum_j c_j^2 u_{ij}^2 + \sum_{j \neq j'}\sum c_j c_{j'} u_{ij} u_{ij'})$$

$$= \sum_j c_j^2 E(u_{ij}^2) + \sum_{j \neq j'}\sum c_j c_{j'} E(u_{ij} u_{ij'})$$

$$= \sum_j c_j^2 k_{ij}(k_i - k_{ij})K_i - \sum_{j \neq j'}\sum c_j c_{j'} k_{ij} k_{ij'} K_i$$

uit (4.23) en (4.30)

$$= K_i \sum_j c_j k_{ij}[c_j(k_i - k_{ij}) - \sum_{j'(\neq j)} c_{j'} k_{ij'}]$$

$$= K_i \sum_j c_j k_{ij} \sum_{j'} k_{ij'}(c_j - c_{j'}).$$

STELLING 4.1.

Indien

(4.34) $\sigma_{i.}^2 > 0$ vir $i=1,2,\ldots,m$,

dan besit $\sum_j c_j u_j$ asimptoties, vir $m \to \infty$, 'n normaalverdeling onder H_o.

Bewys: Omdat alle k_i eindig is, geld vir alle waardes van i, j en h dat

(4.35) $0 < \delta_{ijh} < 1$.

Gevolglik is $\Psi_i(\delta_{ijh})$ eindig vir enige waardes van i, j en h (uit die definisie van $\Psi_i(\delta)$).

Omdat alle k_{ij} eindig is, volg dus dat

(4.16) $u_{ij} = \sum_h (a_{ijh} - \overline{\Psi}_i)$ eindig is, sodat

(4.36) $E|u_{ij}|^{2+\delta} < \infty$ vir alle i en j en vir $\delta > 0$.

Vir eindige c_j geld

(4.37) $E|c_j u_{ij}|^{2+\delta} < \infty$.

Uit lemma 3.11 volg nou dat

(4.38) $E|\sum_j c_j u_{ij}|^{2+\delta} < \infty$ omdat k eindig is.

Dus

(4.39) $E|v_i|^{2+\delta} < \infty$ vir $i=1,2,\ldots,m$.

Die groothede v_i en $v_{i'}$ is onderling onafhanklik indien $i' \neq i$. Volgens die sentrale limietstelling (Liapounoff-voorwaarde) besit $m^{-\frac{1}{2}} \sum_{i=1}^{m} v_i$ onder voorwaarde (4.34) asimptoties, vir $m \to \infty$, 'n normaalverdeling.

Maar $m^{-\frac{1}{2}}\sum_i v_i = m^{-\frac{1}{2}}\sum_i\sum_j c_j u_{ij}$

$= \sum_j c_j u_j$ m.b.v. (4.17).

Gevolglik besit $\sum_j c_j u_j$ asimptoties, vir $m\to\infty$, 'n normaalverdeling.

Opmerking.

'n Nodige en voldoende voorwaarde dat

(4.34) $\sigma^2_{i.} > 0$ vir $i=1,2,\ldots,m$,

is dat die c_j 's nie almal gelyk is nie en vir tenminste twee verskillende c_j 's die ooreenkomstige k_{ij} 's > 0.

Bewys: Uit (4.25) volg $K_i \geq D > 0$ vir alle i.

Gevolglik is $\sigma^2_{i.} > 0$ as en slegs as

(4.40) $\sum_j c_j k_{ij} \sum_{j'} k_{ij'}(c_j - c_{j'}) > 0$ (sien (4.33)).

Maar $\sum_j c_j k_{ij} \sum_{j'} k_{ij'}(c_j - c_{j'})$

$$= \sum_j c_j k_{ij} \sum_{j'(<j)} k_{ij'}(c_j - c_{j'}) + \sum_j c_j k_{ij} \sum_{j'(>j)} k_{ij'}(c_j - c_{j'})$$

$$= \sum_j \sum_{j'(<j)} k_{ij}k_{ij'}\, c_j(c_j - c_{j'}) + \sum_j \sum_{j'(<j)} k_{ij}k_{ij'}\, c_{j'}(c_{j'} - c_j)$$

$$= \sum_j \sum_{j'(<j)} k_{ij}k_{ij'}\, [c_j(c_j - c_{j'}) + c_{j'}(c_{j'} - c_j)]$$

$$= \sum_j \sum_{j'(<j)} k_{ij}k_{ij'}(c_j - c_{j'})^2$$

wat positief is as en slegs as bogenoemde voorwaardes geld.

Lemma 4.4. Indien

(4.41) $\lim_{m\to\infty} m^{-1}k_{.j} > 0$ vir $j=1,2,\ldots,k$,

dan voldoen die u_j 's onder H_o aan die voorwaardes

(4.42) $\lim_{m\to\infty} \mathrm{var}(u_j) > 0$ vir $j=1,2,\ldots,k$ en

(4.43) $\lim_{m\to\infty} E|u_j|^{2+\delta} < \infty$ vir $j=1,2,\ldots,k$ en $\delta > 0$.

Bewys: $\lim_{m\to\infty} \mathrm{var}(u_j) \geq D \lim_{m\to\infty} m^{-1}k_{.j}$ uit (4.28)

> 0 m.b.v. (4.41), sodat (4.42) bevredig word.

Uit die bewys van stelling 4.1 volg, omdat alle k_i na bo begrens is, dat

(4.36') $E|u_{ij}|^{2+\delta} \leq M < \infty$ vir $i=1,2,\ldots,m$; $j=1,2,\ldots,k$,

vaste $\delta > 0$ en M 'n eindige positiewe getal.

Dit geld in die besonder ook vir $\delta = 2$.

Nou is

$$E(u_j^4) = E[m^{-\frac{1}{2}} \sum_i u_{ij}]^4$$

$$= E[m^{-2} \sum_i \sum_{i'} \sum_{i''} \sum_{i'''} u_{ij} u_{i'j} u_{i''j} u_{i'''j}]$$ waar alle i 's gaan van 1 tot m

$$= m^{-2} \sum_i \sum_{i'} E(u_{ij}^2) E(u_{i'j}^2) + m^{-2} \sum_i E(u_{ij}^4)$$ want u_{ij} en $u_{i'j}$ is onderling onafhanklik vir $i \neq i'$ en $E(u_{ij}) = 0$ uit (4.20)

$< \infty$ onafhanklik van m uit (4.36').

Hieruit volg dat (4.43) geld vir $\delta \leq 2$.

Lemma 4.5. Indien

(4.41) $\lim_{m\to\infty} m^{-1} k_{.j} > 0$ vir $j=1,2,\ldots,k$,

dan is die gesamentlike verdeling van die variante u_j onder H_o asimptoties normaal as $m \to \infty$.

Bewys: Omdat

(4.42) $\lim_{m\to\infty} \operatorname{var}(u_j) > 0$ vir $j=1,2,\ldots,k$,

(4.43) $\lim_{m\to\infty} E|u_j|^{2+\delta} < \infty$ vir $j=1,2,\ldots,k$ en $\delta > 0$

en omdat $\sum_j c_j u_j$ asimptoties normaal verdeel is as $m \to \infty$ (stelling 4.1) [5], volg die resultaat met behulp van lemma 2.5.

---

5). Aan voorwaarde (4.34) word in die algemeen voldoen - vergelyk stelling 4.1 (opmerking).

Opmerkings.

1). Die karakteristieke funksie van die gesamentlike verdeling van die u_j 's het dus die vorm

$e^{-\frac{1}{2}\vec{\theta}'\Sigma\vec{\theta}}$ waar Σ die momentematriks van die u_j 's voorstel, $\vec{\theta}' = \{\theta_1, \theta_2, \ldots, \theta_k\}$ en $\vec{\theta}$ die ooreenkomstige kolomvektor.

2). Neem as voorwaarde

(4.44) Σ is van rang $(k-1)$.

Omdat $\sum_j u_j = 0$ (sien (4.17)), kan enige u_j uitgedruk word in terme van die ander $(k-1)$ u_j 's, naamlik

(4.45) $u_{j'} = - \sum_{j(\neq j')} u_j$,

maar onder voorwaarde (4.44) kan $u_{j'}$ nie in 'n kleiner aantal as $(k-1)$ u_j 's uitgedruk word nie.

Dui die momentematriks van $u_1, u_2, \ldots, u_{k-1}$ aan deur $\Lambda = \{\lambda_{\zeta\varrho}\}$ waar ζ en ϱ nou die waardes $1, 2, \ldots, (k-1)$ deurloop. Onder voorwaarde (4.44) is Λ dus nie-singulier.

Dui die kofaktore van die element $\lambda_{\zeta\varrho}$ van Λ aan deur $|\Lambda_{\zeta\varrho}|$.

STELLING 4.2.

Indien

(4.41) $\lim_{m\to\infty} m^{-1}k_{.j} > 0$ vir $j=1,2,\ldots,k$ en

(4.44) Σ van rang $(k-1)$ is,

dan besit die toetsingsgrootheid

(4.46) $$T = \sum_{\zeta=1}^{k-1} \sum_{\varrho=1}^{k-1} \frac{|\Lambda_{\zeta\varrho}|}{|\Lambda|} u_\zeta u_\varrho$$

asimptoties, vir $m\to\infty$, onder H_0 'n χ^2-verdeling met $(k-1)$ grade van vryheid.

Bewys: Die bewys volg met behulp van lemma 4.5 en §1.4.

Opmerkings.

1). T kan meer kompak geskryf word in die vorm

$$T = \frac{|\Delta_u|}{|\Delta|}$$ waar die simbole soos volg gedefinieer word:

Beskou die matriks Δ_u verkry uit

$$V_u = \begin{Bmatrix} \sigma_1^2, & \sigma_{12}, & \ldots, & \sigma_{1k}, & u_1 \\ \sigma_{21}, & \sigma_2^2, & \ldots, & \sigma_{2k}, & u_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \sigma_{k1}, & \sigma_{k2}, & \ldots, & \sigma_k^2, & u_k \\ u_1, & u_2, & \ldots, & u_k, & 0 \end{Bmatrix}$$

deur weglating van 'n willekeurige ry en kolom behalwe die laaste ry en kolom, en bereken sy determinant $|\Delta_u|$.

Beskou ook die matriks Δ verkry uit $\Sigma = \{\sigma_{jj'}\}$ $(j, j' = 1, 2, \ldots, k)$ deur weglating van dieselfde ry en kolom as in die eerste geval, en bereken sy determinant $|\Delta|$.

2). Die toetsingsgrootheid T is 'n uitbreiding van die toetsingsgrootheid χ_r^2 van B+v.E(1953), naamlik na die geval waar gewerk word met funksies van rangnommers in plaas van slegs met rangnommers.

3). Stelling 4.2 kan uitgebrei word na die geval waar Σ van rang $\nu \leq (k-1)$ is. Hierop gaan ons nie verder in nie, behalwe dat ons hier 'n stelling van B+v.E(1953) vermeld: Indien 'n aantal k_{ij} 's = 0, kan dit gebeur dat ons in Σ twee of meer komplementêre groepe elemente het (vergelyk B+v.E(1953) p. 361), d.w.s. dat in elke ry (en elke kolom) van Σ slegs elemente van één van die groepe voorkom. Hierdie groepe word genoem „non-compared" groepe elemente, en die aantal sulke groepe word aangedui deur s.

Lemma 4.6. (B+v.E(1953) stelling 2): As en slegs as daar s „non-compared" groepe elemente in 'n k x k momente-matriks van die u_j 's voorkom, is die rang van die matriks gelyk aan (k-s).

4.3.2. Die toetsingsgrootheid van BENARD en VAN ELTEREN(1953).

Stel ons $\Psi_i(\delta_{ijh}) = r_{ijh}$, dan herlei u_{ij} van vergelyking (4.16) na $\tilde{u}_{ij}$ van (4.6), d.w.s. tot die tipe grootheid wat deur B+v.E(1953) gebruik word.

Stel voorts: $\tilde{u}_j = m^{-\frac{1}{2}} \sum_i \tilde{u}_{ij}$ [6].

Definieer

$$\tilde{\chi}_r^2 = \frac{|\tilde{\Delta}_u|}{|\tilde{\Delta}|}$$

waar alle simbole soortgelyk is as in die voorgaande paragrawe, behalwe dat nou deurgaans geld $\Psi_i(\delta_{ijh}) = r_{ijh}$.

B+v.E(1953) het bewys dat die toetsingsgrootheid $\tilde{\chi}_r^2$ onder die voorwaardes

(4.47) die aantal kolomme k is eindig

(4.48) die aantal rye m neig na ∞

(4.49) $\lim\limits_{m\to\infty} \sum\limits_i \tilde{\sigma}_j^{-3} E|\tilde{u}_{ij}|^3 = 0$ vir $j=1,2,\ldots,k$ en

(4.50) die rang van die matriks $\tilde{R} = \{\tilde{\rho}_{jj'}\}$ waar

$\tilde{\rho}_{jj'} = \lim\limits_{m\to\infty} \tilde{\sigma}_{jj'}/\tilde{\sigma}_j\tilde{\sigma}_{j'}$, is gelyk aan (k-1),

asimptoties 'n χ^2-verdeling met (k-1) g.v.v. besit.

Hierby word aangeneem dat alle k_{ij} eindig bly as $m\to\infty$.

---

6). In §4.1 is genoem dat B+v.E(1953) met kolomtotale $\tilde{u}_j = \sum_i \tilde{u}_{ij}$ werk. Ons wysig egter die definisie van $\tilde{u}_j$ sodat $0 < \lim\limits_{m\to\infty} \text{var}(\tilde{u}_j) < \infty$. Hierdeur word die toetsingsgrootheid $\tilde{\chi}_r^2$ (sien die definisie hieronder) egter nie beïnvloed nie aangesien die addisionele m 'e in $\tilde{\chi}_r^2$ mekaar uitkanselleer.

B+v.E(1953) het ook bewys dat die voorwaardes

(4.51) $\lim_{m \to \infty} m^{-1} \sum_i k_{ij} > 0$ vir $j=1,2,\ldots,k$ [7] en

(4.52) die matriks

$$\tilde{C} = \begin{Bmatrix} \tilde{\varkappa}_{11}, \tilde{\varkappa}_{12}, \ldots, \tilde{\varkappa}_{1k} \\ \tilde{\varkappa}_{21}, \tilde{\varkappa}_{22}, \ldots, \tilde{\varkappa}_{2k} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \tilde{\varkappa}_{k1}, \tilde{\varkappa}_{k2}, \ldots, \tilde{\varkappa}_{kk} \end{Bmatrix}$$

met $\tilde{\varkappa}_{jj'} = \lim_{m \to \infty} m^{-1} \sum_i k_{ij} k_{ij'}$, is nie 'n matriks van die

tipe $\begin{pmatrix} P & O' \\ O & Q \end{pmatrix}$ nie waar P en Q vierkante matrikse is en O en O' uit nulle bestaan,

voldoende is vir die geldigheid van (4.49) en (4.50).

Omdat $E(u_{ij} . u_{ij'}) = 0$ as en slegs as $k_{ij} k_{ij'} = 0$ (sien vergelykings (4.25) en (4.30)), kan voorwaarde (4.44) deur voorwaarde (4.52) vervang word (want (4.52) impliseer dat (4.44) geld).

4.4. DIE TOETSINGSGROOTHEID T_k.

4.4.1. Inleiding.

Die toetsingsgrootheid T soos gedefinieer in vergelyking (4.46) is baie algemeen, maar dit is in 'n moeilik hanteerbare vorm.

Ons gaan ons vervolgens beperk tot 'n eenvoudiger toetsingsgrootheid T_k waartoe T reduseer indien alle $k_{ij} = b > 0$. Ten spyte van hierdie beperking, is T_k nog 'n uitbreiding op die probleem van m-rangskikkings soos deur FRIEDMAN(1937) en KENDALL(1948) behandel.

---

7). B+v.E(1953) beweer dat die voorwaarde $\lim_{m \to \infty} m^{-1} \sum_i k_i > 0$ voldoende is vir hulle stelling 5. Die voorwaarde moet egter wees $\lim_{m \to \infty} m^{-1} \sum_i k_{ij} > 0$ vir $j=1,2,\ldots,k$. Hierdie voorwaarde is dieselfde as voorwaarde
(4.41) $\lim_{m \to \infty} m^{-1} k_{.j} > 0$ vir $j=1,2,\ldots,k$.

4.4.2. Definisie en Limietverdeling van T_k.

Stel $k_{ij} = b > 0$ vir alle i en j, dan is

$$k_{i} = kb \text{ vir alle } i,$$

$$k_{\cdot j} = mb \text{ vir alle } j \text{ en}$$

$$N = mkb.$$

Laat

$$(4.53) \quad u_j = m^{-\frac{1}{2}} \sum_i \sum_{\eta=1}^{b} (a_{ij\eta} - \bar{\Psi}_i) \text{ soos in } (4.17) \quad [8].$$

Verder is

$$(4.54) \quad \sigma_j^2 = m^{-1} b(kb-b) \sum_i K_i \text{ uit } (4.24)$$

$$= b(k-1)[mk(kb-1)]^{-1} \sum_i \sum_{p=1}^{kb} (a_{ip} - \bar{\Psi}_i)^2 \text{ uit } (4.22)$$

$$= (k-1)K \text{ waar}$$

$$(4.55) \quad K = b[mk(kb-1)]^{-1} \sum_i \sum_p (a_{ip} - \bar{\Psi}_i)^2.$$

Stel

$$(4.56) \quad \sigma_j^2 = \sigma^2 \text{ vir alle } j.$$

Verder is

$$(4.57) \quad \sigma_{jj'} = - b^2 m^{-1} \sum_i K_i \quad \text{uit } (4.29)$$

$$= - K.$$

$\Lambda = \{\lambda_{\varsigma\varsigma'}\}$ is nou die (k-1)x(k-1) matriks met determinant

$$|\Lambda| = \begin{vmatrix} (k-1)K, & -K & ,\ldots\ldots, & -K \\ -K & ,(k-1)K, & \ldots\ldots, & -K \\ \cdot & \cdot & \cdot & \cdot \\ -K & , -K & ,\ldots\ldots, & (k-1)K \end{vmatrix}$$

$$= K^{k-1} \begin{vmatrix} (k-1), & -1 & ,\ldots\ldots\ldots, & -1 \\ -1 & ,(k-1), & \ldots\ldots\ldots, & -1 \\ \cdot & \cdot & \cdot & \cdot \\ -1 & , -1 & ,\ldots\ldots\ldots, & (k-1) \end{vmatrix}$$

$$= K^{k-1} k^{k-2}, \quad \text{want:}$$

Neem $A = \begin{Bmatrix} a & , -1 \\ -1 & , a \end{Bmatrix}$, 'n 2x2 matriks.

Dan is $|A| = (a+1)(a-1)$.

---

8). η deurloop die waardes 1,2,...,b en p die waardes 1,2,...,kb.

Neem $A' = \begin{Bmatrix} a, & -1, & -1 \\ -1, & a, & -1 \\ -1, & -1, & a \end{Bmatrix}$, 'n 3x3 matriks, dan is

$|A'| = (a+1)^2(a-2)$.

Neem $A'' = \begin{Bmatrix} a, & -1, & -1, & -1 \\ -1, & a, & -1, & -1 \\ -1, & -1, & a, & -1 \\ -1, & -1, & -1, & a \end{Bmatrix}$, 'n 4x4 matriks, dan is

$|A''| = (a+1)^3(a-3)$, ens.

Vir $A^{*} = \begin{Bmatrix} a, & -1, & \ldots, & -1 \\ -1, & a, & \ldots, & -1 \\ \cdot & \cdot & \cdot & \cdot \\ -1, & -1, & \ldots, & a \end{Bmatrix}$, 'n (k-1)x(k-1) matriks, is

(4.58) $|A^{*}| = (a+1)^{k-2}(a-k+2)$.

Stel a = (k-1) in (4.58), dan is die determinant van die (k-1)x(k-1) matriks

$\begin{Bmatrix} (k-1), & -1, & \ldots, & -1 \\ -1, & (k-1), & \ldots, & -1 \\ \cdot & \cdot & \cdot & \cdot \\ -1, & -1, & \ldots, & (k-1) \end{Bmatrix}$ gelyk aan k^{k-2}.

As kofaktore van Λ vind ons:

$$|\Lambda_{\varsigma\varsigma}| = \begin{vmatrix} (k-1)K, & -K, & \ldots, & -K \\ -K, & (k-1)K, & \ldots, & -K \\ \cdot & \cdot & \cdot & \cdot \\ -K, & -K, & \ldots, & (k-1)K \end{vmatrix}, \text{ 'n } (k-2)x(k-2) \text{ determinant}$$

$$= K^{k-2} \begin{vmatrix} (k-1), & -1, & \ldots, & -1 \\ -1, & (k-1), & \ldots, & -1 \\ \cdot & \cdot & \cdot & \cdot \\ -1, & -1, & \ldots, & (k-1) \end{vmatrix}$$

$= 2K^{k-2}k^{k-3}$ m.b.v. (4.58) met a = (k-1).

Soortgelyk is

$|\Lambda_{\varsigma\varsigma'}| = K^{k-2}k^{k-3}$ vir $\varsigma' \neq \varsigma$.

Vervolgens is

(4.59) $\dfrac{|\Lambda_{\varsigma\varsigma}|}{|\Lambda|} = \dfrac{2}{Kk}$ en

(4.60) $\dfrac{|\Lambda_{\varsigma\varsigma'}|}{|\Lambda|} = \dfrac{1}{Kk}$ vir $\varsigma' \neq \varsigma$ sodat

$$(4.61)\quad T_k = \sum_{\zeta=1}^{k-1}\sum_{\zeta'=1}^{k-1} \frac{|\Lambda_{\zeta\zeta'}|}{|\Lambda|} u_\zeta u_{\zeta'} \qquad \text{- sien (4.46)}$$

$$= \frac{2}{Kk}\sum_{\zeta} u_\zeta^2 + \frac{1}{Kk}\sum_{\zeta\neq\zeta'}\sum u_\zeta u_{\zeta'}$$

$$= \frac{2}{Kk}\Big[\sum_{\zeta} u_\zeta^2 + \sum_{\zeta<\zeta'}\sum u_\zeta u_{\zeta'}\Big].$$

Omdat $\sum_{j=1}^{k} u_j = 0$ uit die definisie van u_j, is

$$u_k = -\sum_{\zeta=1}^{k-1} u_\zeta,$$

$$u_k^2 = (-\sum_{\zeta} u_\zeta)^2 = \sum_{\zeta} u_\zeta^2 + 2\sum_{\zeta<\zeta'}\sum u_\zeta u_{\zeta'} \quad \text{sodat}$$

$$(4.62)\quad 2\sum_{\zeta} u_\zeta^2 + 2\sum_{\zeta<\zeta'}\sum u_\zeta u_{\zeta'} = \sum_{\zeta} u_\zeta^2 + u_k^2$$

$$= \sum_{j=1}^{k} u_j^2 .$$

Uit (4.61) en (4.62) volg dan

$$(4.63)\quad T_k = \frac{1}{Kk}\sum_{j=1}^{k} u_j^2$$

$$= \frac{(k-1)}{k\sigma^2}\sum_{j} u_j^2 \quad \text{uit (4.54) en (4.56).}$$

STELLING 4.3.

Die toetsingsgrootheid

$$T_k = \frac{(k-1)}{k\sigma^2}\sum_{j} u_j^2$$

besit onder H_0 asimptoties, vir $m\to\infty$, 'n χ^2-verdeling met $(k-1)$ grade van vryheid.

Bewys: T_k is 'n spesiale geval van T wat onder voorwaardes (4.41) en (4.44) asimptoties χ^2 verdeel is as $m\to\infty$ (stelling 4.2).

Nou is $\lim_{m\to\infty} m^{-1}k_{.j} = \lim_{m\to\infty} m^{-1}mb = b > 0$ sodat (4.41) bevredig word. Aan voorwaarde (4.44) word ook voldoen want $|\Lambda| = K^{k-1}k^{k-2} > 0$.

Die stelling volg dus direk.

Opmerkings.

1). FRIEDMAN(1937) se bewys deur karakteristieke funksies kan uitgebrei word na hierdie geval.

2). $$E(T_k) = \frac{(k-1)}{k\sigma^2} E(\sum_j u_j^2) = \frac{(k-1)}{k\sigma^2} \sum_j E(u_j^2) = \frac{(k-1)}{k\sigma^2} \cdot k\sigma^2 = (k-1) = \text{aantal g.v.v. van } T_k.$$

4.4.3. Spesiale gevalle van T_k.

4.4.3.1. BENARD en VAN ELTEREN(1953).

Stel $\Psi_i(\delta_{ij\eta}) = r_{ij\eta}$, dan herlei

$K_i = [kb(kb-1)]^{-1} \sum_p (a_{ip} - \bar{\Psi}_i)^2$ – sien (4.22) – na

(4.64) $K_i = [k^3b^3 - \sum_{\alpha_i} g_{\alpha_i}^3]/[12bk(kb-1)]$ uit (2.61), sodat

(4.65) $$\sigma^2 = (k-1)K \text{ uit (4.54) en (4.56)} = m^{-1}(k-1)b^2 \sum_i K_i = b(k-1)[12mk(kb-1)]^{-1} \sum_i (k^3b^3 - \sum_{\alpha_i} g_{\alpha_i}^3).$$

Verder is

(4.66) $u_j = m^{-\frac{1}{2}} \sum_i \sum_\eta [r_{ij\eta} - \frac{1}{2}(kb+1)]$ uit (4.53)

sodat T_k herlei na

(4.67) $$T_B = \frac{12(kb-1)}{b \sum_i (k^3b^3 - \sum_{\alpha_i} g_{\alpha_i}^3)} \sum_j [\sum_i \sum_\eta \left\{r_{ij\eta} - \tfrac{1}{2}(kb+1)\right\}]^2.$$

T_B is die toetsingsgrootheid wat deur BENARD en VAN ELTEREN(1953) in §3.3 bespreek word waarby aangeneem word dat alle kovariansies $\sigma_{jj'}$ (sien (4.30) waarby $\Psi_i(\delta_{ij\eta}) = r_{ij\eta}$) gelyk is. Nodig en voldoende hiervoor is dat alle k_{ij} gelyk is, sê $k_{ij} = b > 0$ vir alle i en j, en dit is presies die geval wat hierbo bespreek is.

Opmerking.

Indien $\Psi_i(\delta_{ij\eta}) = \delta_{ij\eta}$ geneem word, herlei T_k ook na T_B.

4.4.3.2. FRIEDMAN(1937).

Stel $k_{ij} = 1$ vir alle i en j, d.w.s. b = 1 in §4.4.3.1. Dan herlei T_k na

(4.68) $T_C = \dfrac{(k-1)}{k\sigma^2} \sum_j [m^{-\frac{1}{2}} \sum_i (a_{ij} - \bar{\Psi}_i)]^2$ uit (4.53) en (4.63)

waarby a_{ij} gedefinieer is in (4.12) en

(4.69) $\sigma^2 = (mk)^{-1} \sum_i \sum_j (a_{ij} - \bar{\Psi}_i)^2$ uit (4.54) en (4.55) met b = 1

$= m^{-1} \sum_i \sigma^2_{\Psi_i}$ waar

(4.70) $\sigma^2_{\Psi_i} = k^{-1} \sum_j (a_{ij} - \bar{\Psi}_i)^2$, sodat dan volg

(4.71) $T_C = \dfrac{(k-1)}{k \sum_i \sigma^2_{\Psi_i}} \sum_j [\sum_i (a_{ij} - \bar{\Psi}_i)]^2.$

Definieer

(4.72) $S^* = \sum_j [\sum_i \{r_{ij} - \frac{1}{2}(k+1)\}]^2$ (vergelyk (4.2)) en

(4.73) $T^* = (k-1)^{-1} \sum_i \sum_{\alpha_i} (g^3_{\alpha_i} - g_{\alpha_i})$ (vergelyk BENARD en VAN ELTEREN(1953) vergelyking (3.4.2)).

Stel weer $\Psi_i(\delta_{ij}) = r_{ij}$, dan is

$\sigma^2 = [12mk]^{-1} \sum_i (k^3 - \sum_{\alpha_i} g^3_{\alpha_i})$ uit (4.65) met b=1 sodat

(4.74) $\sigma^2_{\Psi_i} = [12k]^{-1}(k^3 - \sum_{\alpha_i} g^3_{\alpha_i})$, en herlei T_C na:

(4.75) $T_D = 12(k-1)S^* / \sum_i (k^3 - \sum_{\alpha_i} g^3_{\alpha_i})$

$= 12(k-1)S^*/[mk(k^2-1) - \sum_i \{\sum_{\alpha_i} g^3_{\alpha_i} - \sum_{\alpha_i} g_{\alpha_i}\}]$

$= 12S^*/[mk(k+1) - (k-1)^{-1} \sum_i \sum_{\alpha_i} (g^3_{\alpha_i} - g_{\alpha_i})]$

$= 12S^*/[mk(k+1) - T^*]$

$= \chi^2$ van FRIEDMAN(1937) soos aangegee deur B+v.E(1953) vergelykings (3.4.1) en (3.4.2).

Indien daar geen knope voorkom nie, d.i. as alle $g_{\alpha_i} = 1$, is $T^* = 0$ en herlei T_D na:

(4.76) $T_F = 12S^*/mk(k+1)$
$= 12[mk(k+1)]^{-1} \sum_j (\sum_i \tilde{r}_{ij})^2$ waar

(4.77) $\tilde{r}_{ij} = r_{ij} - \frac{1}{2}(k+1)$.

Vergelyk FRIEDMAN(1937) vergelyking (18) en FRASER(1957) p. 262 probleem 22.

FRIEDMAN(1937) het bewys dat T_F asimptoties 'n χ^2-verdeling met (k-1) g.v.v. besit as $m \to \infty$.

Uit stelling 4.3 volg dat die meer algemene toetsingsgrootheid T_C, wat na $T_D = \chi^2$ herlei indien $\Psi_i(\delta_{ij}) = r_{ij}$ en voorts na T_F indien daar geen knope voorkom nie, asimptoties χ^2 verdeel is onder H_0 met (k-1) g.v.v. as $m \to \infty$.

Opmerkings.

1). FRIEDMAN(1937) se toetsingsgrootheid is 'n spesiale geval van BENARD en VAN ELTEREN(1953) se toetsingsgrootheid T_B.

2). KENDALL(1948) het die grootheid T_F geskryf in die vorm

(4.78) $T_F = m(k-1)W$ waar

(4.79) $W = 12S^*/[m^2k(k^2-1)]$.

Kendall noem W die „coefficient of concordance".

3). Indien daar geen gelyke waarnemings voorkom nie, alle $\Psi_i(\delta) = \Psi(\delta)$ en alle $k_{ij} = 1$, dan herlei $\sigma^2_{\Psi_i}$ uit (4.70) na

(4.80) $\sigma^2_{\Psi_i} = k^{-1} \sum_j [\Psi(\delta_{ij}) - \bar{\Psi}]^2$ waar $\bar{\Psi} = k^{-1} \sum_j \Psi[j/(k+1)]$
$= \sigma^2_{\Psi}$ (sê).

Dus $\sigma^2_j = m^{-1} \sum_i \sigma^2_{ij}$ uit (4.24)
$= m^{-1} \sum_i \sigma^2_{\Psi}$
$= \sigma^2_{\Psi}$.

Laat S_m die maksimum-waarde wees wat $S = \sum_j u_j^2$ onder bogenoemde voorwaardes kan aanneem. Dit word verkry as alle rangskikkings identies is. Die kolomtotale is dan (nie noodwendig in dié volgorde nie):

$$\sqrt{m}\ \Psi(\tfrac{1}{k+1}),\ \sqrt{m}\ \Psi(\tfrac{2}{k+1}),\ \ldots\ ,\ \sqrt{m}\ \Psi(\tfrac{k}{k+1}) \quad \text{sodat}$$

$$\begin{aligned} S_m &= \sum_j [\sqrt{m}\ \Psi(\tfrac{j}{k+1}) - \sqrt{m}\ \bar{\Psi}]^2 \\ &= m \sum_j [\Psi(\tfrac{j}{k+1}) - \bar{\Psi}]^2 \\ &= mk\sigma_\Psi^2 \text{ uit (4.80).} \end{aligned}$$

Gevolglik herlei T_C uit (4.68) na

$$\begin{aligned} T_C' &= \frac{(k-1)}{k\sigma_\Psi^2} \sum_j u_j^2 \quad \text{m.b.v. (4.69) en (4.80)} \\ &= m(k-1)S/S_m\ . \end{aligned}$$

4). TERPSTRA(1962) wys daarop dat KENDALL(1948) se bewys dat $m(k-1)W$ asimptoties χ^2 verdeel is as $m\to\infty$, foutief is (stelling VI).

5). In EHRENBERG, A.S.C.(1952):Biometrika 39, pp. 82-87, is 'n toetsingsgrootheid voorgestel vir die probleem van m-rangskikkings. TERPSTRA(1955,1956) het die asimptotiese verdeling daarvan bepaal vir m eindig en $k\to\infty$, terwyl VAN ELTEREN(1957) die asimptotiese verdeling bepaal het vir k eindig en $m\to\infty$ en aangetoon het dat die toetsingsgrootheid in 'n triviale geval ekwivalent is aan die toetsingsgrootheid van FRIEDMAN(1937).

4.4.3.3. $\Psi_i(\delta_{ij\eta}) \equiv \Xi_{r_{ij\eta}} \equiv E(\xi_{r_{ij\eta}})$

waar ξ_{ip_i} $(\xi_{i1} < \xi_{i2} < \ldots < \xi_{ik_i})$ die waardes in rang van 'n ewekansige steekproef van grootte $k_i = kb$ uit 'n normaal (0,1)-populasie is (vergelyk TERRY(1952) en §2.5.2.2).

In hierdie geval is: [9)]

$$\bar{\Xi}_i = \bar{\Psi}_i$$
$$= (kb)^{-1} \sum_p \Xi_{ip}$$
$$= 0 \text{ weens simmetrie.}$$

Verder is

(4.81) $u_j = m^{-\frac{1}{2}} \sum_i \sum_\eta a_{ij\eta}$ en

(4.82) $\sigma^2 = (k-1)K$ uit (4.54)

$$= (k-1)b[mk(kb-1)]^{-1} \sum_i \sum_p a_{ip}^2 \quad \text{waar}$$

(4.83) $a_{ip} = g_{\alpha_i}^{-1} \sum_{\mu_i} \Psi_i[(G_{\alpha_i-1}+\mu_i)/(k_i+1)]$

wanneer $G_{\alpha_i-1}+1 \leq p \leq G_{\alpha_i}$

met $\Psi_i(\delta_{ip}) = \Xi_{ip}$ in hierdie geval.

Uit stelling 4.3 volg dat

$$T_k = \frac{(k-1)}{k\sigma^2} \sum_j u_j^2$$ onder H_0 asimptoties, vir $m \to \infty$,

'n χ^2-verdeling met (k-1) g.v.v. besit.

Opmerking.

Hierdie is 'n veralgemening van 'n probleem voorgestel deur FRASER(1957) (p. 263 probleem 23). Stel ons b = 1, dan herlei bostaande resultate na genoemde probleem van FRASER(1957) en kry ons:

$$u_j = m^{-\frac{1}{2}} \sum_i a_{ij} \quad \text{en}$$

$$\sigma^2 = (mk)^{-1} \sum_i \sum_j a_{ij}^2 \quad \text{sodat}$$

(4.84) $T_k = \dfrac{m(k-1)}{\sum_i \sum_j a_{ij}^2} \sum_j u_j^2 .$

---

9). Ons kan in hierdie geval ook met 'n toetsingsgrootheid van die vorm van (4.46) werk indien die k_{ij} 's nie almal gelyk is nie.

Indien geen knope voorkom nie, is $a_{ij} = \Xi_{ij}$, sodat

(4.85) $$T_k = \frac{(k-1)}{\sum\limits_i \sum\limits_j \Xi_{ij}^2} \sum_j [\sum_i \Xi_{ij}]^2 .$$

T_k is onder H_o asimptoties χ^2 verdeel met (k-1) g.v.v. as $m \to \infty$.

4.4.3.4. $\Psi_i(\delta_{ij\eta}) \equiv Q(\delta_{ij\eta})$

waar Q(q) die q^{de} kwantiel van 'n normaal (0,1)-verdeling is (sien VAN DER WAERDEN(1957) en §2.5.2.3).

In hierdie geval is [10)]

$$\bar{Q}_i = \bar{\Psi}_i$$
$$= (kb)^{-1} \sum_p Q(\delta_{ip})$$
$$= 0 \text{ weens simmetrie.}$$

Verder is

(4.86) $u_j = m^{-\frac{1}{2}} \sum\limits_i \sum\limits_\eta a_{ij\eta}$ en

(4.87) $\sigma^2 = (k-1)b[mk(kb-1)]^{-1} \sum\limits_i \sum\limits_p a_{ip}^2$ uit (4.82)

met a_{ip} soos in (4.83) waarby $\Psi_i(\delta_{ip}) = Q(\delta_{ip})$.

Soos in die geval van §4.4.3.3 volg dat

$T_k = \frac{(k-1)}{k\sigma^2} \sum\limits_j u_j^2$ onder H_o asimptoties, vir $m \to \infty$, 'n χ^2-verdeling met (k-1) g.v.v. besit.

Opmerking.

Stel b = 1, dan is

$$u_j = m^{-\frac{1}{2}} \sum_i a_{ij} \quad \text{en}$$

$$\sigma^2 = (mk)^{-1} \sum_i \sum_j a_{ij}^2 \quad \text{sodat}$$

(4.88) $$T_k = \frac{m(k-1)}{\sum\limits_i \sum\limits_j a_{ij}^2} \sum_j u_j^2 .$$

Indien daar geen knope voorkom nie, is $a_{ip} = Q(\delta_{ip})$ sodat

(4.89) $$T_k = \frac{(k-1)}{\sum\limits_i \sum\limits_j Q^2(\delta_{ij})} \sum_j [\sum_i Q(\delta_{ij})]^2 .$$

10). Sien voetnoot 9.

T_k is onder H_o asimptoties χ^2 verdeel met $(k-1)$ g.v.v. as $m \to \infty$.

4.4.3.5. Neem vervolgens die meer algemene funksies $\Psi_i(\delta_{ijh}, x_{ijh})$, $i=1,2,\ldots,m$, wat strengstygend (of strengdalend) is in beide argumente en eindig is vir alle eindige x_{ijh} en vir $0 < \delta_{ijh} < 1$, met

$$\left|\frac{\partial \Psi_i}{\partial x}\right| \geq \varepsilon > 0 \text{ vir alle } i \text{ en } x.$$

Laat (sien byvoorbeeld vergelyking (4.12)):

$$(4.90)\quad a_{ip_i} = g_{\alpha_i}^{-1} \sum_{\mu_i} \Psi_i\left[z_{i,G_{\alpha_i-1}+\mu_i}, (G_{\alpha_i-1}+\mu_i)/(k_i+1)\right]$$

vir $G_{\alpha_i-1}+1 \leq p_i \leq G_{\alpha_i}$.

Definieer

$$(4.91)\quad a_{ijh} = a_{ip_i} \quad \text{indien } x_{ijh} = z_{ip_i},$$

$$(4.92)\quad \bar{\Psi}_i = k_i^{-1} \sum_{p_i} \Psi_i(x_{ip_i}, \delta_{ip_i}) = k_i^{-1} \sum_{p_i} a_{ip_i}.$$

Laat weer

$$(4.93)\quad u_{ij} = \sum_h (a_{ijh} - \bar{\Psi}_i) \quad \text{indien } k_{ij} > 0,$$

$$(4.94)\quad u_j = m^{-\frac{1}{2}} \sum_i u_{ij}.$$

Die resultate van §e 4.2 tot 4.4.2 kan vir hierdie geval aangepas word. Ons gee slegs 'n kort aanduiding.

Dui die x_{ijh}-waardes aan deur $z_1, z_2, \ldots, z_N$ en laat $Z = (z_1, z_2, \ldots, z_N)$.

Soos in lemmas 4.1 tot 4.3 kan bewys word dat

$$(4.95)\quad E(u_j | H_o; Z) = 0,$$

$$(4.96)\quad \sigma_j^2 = \mathrm{var}(u_j | H_o; Z) = m^{-1} \sum_i k_{ij}(k_i - k_{ij}) k_i^{-1}(k_i-1)^{-1} \sum_{p_i}(a_{ip_i} - \bar{\Psi}_i)^2 \quad \text{en}$$

$$(4.97)\quad \sigma_{jj'} = \mathrm{kov}(u_j, u_{j'} | H_o; Z) = -\, m^{-1} \sum_i k_{ij} k_{ij'} k_i^{-1}(k_i-1)^{-1} \sum_{p_i}(a_{ip_i} - \bar{\Psi}_i)^2.$$

Ons kies 'n getal $\omega > 0$ en vereis dat in elke ry tenminste twee x 'e in verskillende „selle" met minstens ω verskil. Alle rye wat nie aan hierdie voorwaarde voldoen nie, word weggelaat.

Omdat alle x_{ijh} eindig is, volg soos in die opmerking na lemma 4.2 dat $\sigma_j^2 > 0$ vir m eindig, en $\lim_{m\to\infty} \sigma_j^2 \geq \omega' > 0$.

Neem ons weer $v_i = \sum_j c_j u_{ij}$ waar die c_j 's eindige reële getalle is wat nie almal gelyk is nie, en $\sigma_{i.}^2 = \mathrm{var}(v_i)$, dan kan ons op analoë wyse as tevore die volgende stellings bewys:

STELLING 4.4.

Indien

(4.98) $\sigma_{i.}^2 > 0$ vir $i=1,2,\ldots,m$,

besit $\sum_j c_j u_j$ asimptoties, vir $m\to\infty$, onder H_0 'n normaalverdeling.

Bewys: Soos stelling 4.1.

Soortgelyke resultate word ook gevind as in lemmas 4.4 en 4.5, sodat ons kry:

STELLING 4.5.

Indien

(4.41) $\lim_{m\to\infty} m^{-1}k_{.j} > 0$ vir $j=1,2,\ldots,k$ en

(4.99) $\Sigma = \{\sigma_{jj'}\}$ van rang $(k-1)$ is vir elke gegewe Z

dan besit

$$T = \sum_{\zeta}\sum_{\zeta} \frac{|\Lambda_{\zeta\zeta'}|}{|\Lambda|} u_\zeta u_{\zeta'}$$

asimptoties, vir $m\to\infty$, onder H_0 'n χ^2-verdeling met $(k-1)$ g.v.v.

Bewys: Soos stelling 4.2.

Stel $k_{ij} = b > 0$ vir alle i en j, dan is (sien §4.4.2)

(4.100) $\sigma_j^2 = (k-1)b[mk(kb-1)]^{-1} \sum_i \sum_p (a_{ip} - \bar{\Psi}_i)^2$ vir alle j

$= \sigma^2$ (sê), en

$$u_j = m^{-\frac{1}{2}} \sum_i \sum_{\eta=1}^{b} (a_{ij\eta} - \bar{\Psi}_i).$$

Definieer weer

$$T_k = \frac{(k-1)}{k\sigma^2} \sum_j u_j^2$$

dan besit T_k asimptoties, vir $m \to \infty$, onder H_o 'n χ^2-verdeling met $(k-1)$ g.v.v. (sien stelling 4.3).

Stel nou $\Psi_i(x_{ij\eta}, \delta_{ij\eta}) \equiv x_{ij\eta}$ (vergelyk PITMAN(1937) en §2.6.2).

Laat $\bar{x}_i = \bar{\Psi}_i$

$$= (kb)^{-1} \sum_p x_{ip}.$$

Dan is

(4.101) $u_j = m^{-\frac{1}{2}} \sum_i \sum_\eta (x_{ij\eta} - \bar{x}_i)$ en

(4.102) $\sigma^2 = (k-1)b[mk(kb-1)]^{-1} \sum_i \sum_p (x_{ip} - \bar{x}_i)^2.$

$T_k = \frac{(k-1)}{k\sigma^2} \sum_j u_j^2$ besit asimptoties 'n χ^2-verdeling onder H_o met $(k-1)$ g.v.v. as $m \to \infty$.

Opmerking.

Stel $b = 1$, dan reduseer bostaande geval na dié voorgestel deur FRASER(1957) (p.262 probleem 21). In dié geval is:

$$\bar{x}_i = k^{-1} \sum_j x_{ij} \;, \quad \bar{x}_{.j} = m^{-1} \sum_i x_{ij} \;, \quad \bar{x} = (mk)^{-1} \sum_i \sum_j x_{ij} \;,$$

$$u_j = m^{-\frac{1}{2}} \sum_i (x_{ij} - \bar{x}_i),$$

$$\sigma^2 = (mk)^{-1} \sum_i \sum_j (x_{ij} - \bar{x}_i)^2$$

$$= m^{-1} \sum_i S_i^2 \quad \text{waar}$$

(4.103) $S_i^2 = k^{-1} \sum_j (x_{ij} - \bar{x}_i)^2$ sodat

(4.104) $T_k = \frac{(k-1)}{k \sum_i S_i^2} \sum_j [\sum_i (x_{ij} - \bar{x}_i)]^2$

$$= \frac{m^2(k-1)}{k \sum_i S_i^2} \sum_j (\bar{x}_{.j} - \bar{x})^2.$$

T_k besit asimptoties, vir $m \to \infty$, onder H_o 'n χ^2-verdeling met (k-1) grade van vryheid.

4.5. ANDER UITBREIDINGS.

4.5.1. Inleiding.

In die probleem van m-rangskikkings ($\S^e$ 4.1 tot 4.4) is die asimptotiese verdelings van voorgestelde toetsingsgroothede ondersoek indien die aantal rangskikkings $m \to \infty$. Die vraag ontstaan wat sal gebeur indien alle $k_{ij} \to \infty$, of indien $k \to \infty$.

Alleen eersgenoemde geval lewer 'n sinvolle bydrae tot die studie van ons probleem. (Sien ook §4.4.3.2 opmerking 5 waar 'n geval aangehaal word waar laasgenoemde geval ($k \to \infty$) bestudeer word).

4.5.2. Geval $k_{ij} \to \infty$ vir $j=1,2,\ldots,k$; $i=1,2,\ldots,m$, waarby k en m eindig is.

Hierdie is 'n twee-faktor variansie-analise waar rangnommers afsonderlik binne elke ry toegeken word. Elke ry kan beskou word as bestaande uit k willekeurige groot steekproewe (die m rye dui m replikate aan). Die teorie van hoofstuk II is dus toepasbaar op elke ry. In notasie sluit ons aan by dié van hoofstuk II.

Die waarnemings in die sel (i,j) kan beskou word as 'n steekproef afkomstig uit 'n populasie met verdelingsfunksie F_{ij}. Onder H_o' nl. dat alle verdelingsfunksies F_{ij} identies is vir gegewe i, sê $= F_i$, is alle permutasies van waarnemings in elke ry ewekansig (H_o).

Neem

$$(4.105) \quad y_{ij} = \begin{cases} k_{ij}^{-\frac{1}{2}} \sum\limits_h (a_{ijh} - \bar{\Psi}_i) & \text{indien } k_{ij} > 0 \\ 0 & \text{indien } k_{ij} = 0 \end{cases}$$

en $y_j = \sum\limits_i y_{ij}$.

Hier word veronderstel dat die Ψ_i funksies is van rangnommers alleen, d.w.s. van die vorm $\Psi_i(\delta_{ijh})$.

Lemma 4.7. Indien

(4.106) $\text{maks}_{p_i}[\Psi_i(\delta_{ip_i})-\bar{\Psi}_i]^2 = o(k_i)$ vir $k_i \to \infty$ en $i=1,2,..,m$

(4.107) (Iedere diskontinuïteit van $F_i(x)) \leq \theta < 1$ vir alle i

(4.108) $\lim_{k_i \to \infty} k_{ij}/k_i \geq \varepsilon > 0$ vir alle i en j,

dan is

(4.109) $w_i = \sum_j c_{ij} y_{ij}$

waar die c_{ij}'s willekeurige eindige reële getalle is, asimptoties normaal verdeel as $k_i \to \infty$, $i=1,2,...,m$.

Bewys: Die bewys is, vir vaste i, analoog aan dié van stelling 2.3.

Opmerking.

Voorwaarde (4.108) is nie nodig nie solank $k_{ij} \to \infty$ as $k_i \to \infty$ (vergelyk stelling 2.3 opmerking 2).

STELLING 4.6.

Onder voorwaardes (4.106), (4.107) en (4.108) en onder H_0' is $\sum_j e_j \hat{y}_j$, waarby die e_j's willekeurige eindige reële getalle is en

(4.110) $\hat{y}_j = (N-k_{.j})^{\frac{1}{2}} N^{-\frac{1}{2}} y_j \sigma_j^{-1}$

met $\sigma_j^2 = \text{var}(y_j)$ en $k_{.j} = \sum_i k_{ij}$,

asimptoties normaal verdeel as $N \to \infty$.

Bewys: w_i en $w_{i'}$ is onderling onafhanklik indien $i \neq i'$.

Met behulp van lemma 4.7 volg dus dat

(4.111) $W = \sum_i^{l} w_i$ asimptoties normaal verdeel is as $N \to \infty$

(as $N \to \infty$, geld dat alle $k_{ij} \to \infty$ omdat k en m eindig is en omdat ons aanneem dat alle k_{ij} van dieselfde orde grootte is - voorwaarde (4.108)).

Omdat die c_{ij} 's willekeurig is (behalwe in soverre dat almal nie gelyk mag wees nie) kan ons sonder verlies aan algemeenheid stel

(4.112) $c_{ij} = (N-k_{.j})^{\frac{1}{2}} N^{-\frac{1}{2}} e_j \sigma_j^{-1}$ vir alle i.

Dan is

$$(4.113)\quad W = \sum_i \sum_j c_{ij} y_{ij}$$

$$= \sum_i \sum_j (N-k_{.j})^{\frac{1}{2}} N^{-\frac{1}{2}} e_j \sigma_j^{-1} y_{ij}$$

$$= \sum_j (N-k_{.j})^{\frac{1}{2}} N^{-\frac{1}{2}} e_j \sigma_j^{-1} \sum_i y_{ij}$$

$$= \sum_j (N-k_{.j})^{\frac{1}{2}} N^{-\frac{1}{2}} e_j \sigma_j^{-1} y_j$$

$$= \sum_j e_j \hat{y}_j \quad \text{m.b.v. } (4.110).$$

$\sum_j e_j \hat{y}_j$ is dus asimptoties normaal verdeel as $N \to \infty$.

Lemma 4.8. Onder die voorwaardes (4.106), (4.107) en (4.108) en onder H_0 geld die voorwaardes

$$(4.114)\quad \lim_{N\to\infty} \operatorname{var}(\hat{y}_j) > 0 \quad \text{en}$$

$$(4.115)\quad \lim_{N\to\infty} E|\hat{y}_j|^{2+\delta} < \infty \text{ vir } j=1,2,\ldots,k \text{ en } \delta > 0.$$

Bewys: Soos in die opmerking na stelling 2.2 volg m.b.v. STOKER(1955) stelling 2.4 dat

$$\liminf_{N\to\infty} \operatorname{var}(y_{ij}) \geq \varepsilon > 0.$$

Maar $\operatorname{var}(y_j) = \sum_i \operatorname{var}(y_{ij})$ – sien byvoorbeeld (4.24).

Dus $\liminf_{N\to\infty} \operatorname{var}(y_j) \geq \liminf_{N\to\infty} \operatorname{var}(y_{ij}) > 0$ vir enige i.

$$\operatorname{var}(\hat{y}_j) = (N-k_{.j})N^{-1}\operatorname{var}(y_j)\ \sigma_j^{-2}$$

$$= 1 - k_{.j}N^{-1} \quad \text{sodat}$$

$$\lim_{N\to\infty} \operatorname{var}(\hat{y}_j) = 1 - \lim_{N\to\infty} k_{.j}N^{-1} > 0 \quad \text{waarmee (4.114)}$$

bevredig word.

Indien, wat verder deurgaans veronderstel word,

$$k_i^{-1} \sum_{p_i} (a_{ip_i} - \bar{\Psi}_i)^4 = O(k_i) \text{ vir } i = 1,2,\ldots,m$$

dan volg op analoë wyse as in bylaag A dat

$$E(y_{ij})^4 < \infty \text{ vir alle i en j.}$$

Aangesien y_{ij} en $y_{i'j}$ onafhanklik is indien $i \neq i'$, volg maklik dat voorwaarde (4.115) bevredig word.

Lemma 4.9. Die variante $\hat{y}_j$ is onder voorwaardes (4.106), (4.107) en (4.108) gesamentlik asimptoties normaal verdeel as $N \to \infty$.

Bewys: Die bewys volg direk uit stelling 4.6 en lemma 4.8 met behulp van lemma 2.5.

Opmerking.

Die karakteristieke funksie van die $\hat{y}_j$'s is dus van die vorm

$e^{-\frac{1}{2}\vec{\theta}'\Sigma\vec{\theta}}$ waar Σ die momentematriks van die $\hat{y}_j$'s voorstel, $\vec{\theta}' = \{\theta_1,\theta_2,\ldots,\theta_k\}$ en $\vec{\theta}$ die ooreenkomstige kolomvektor.

Lemma 4.10. Die momentematriks Σ van die variante $\hat{y}_j$, $j=1,2,\ldots,k$ besit onder H_o 'n rang van (k-1).

Bewys: Met behulp van (4.23) en (4.30) volg:

$$(4.116)\quad \mathrm{var}(y_{ij}) = (k_i-k_{ij})k_i^{-1}(k_i-1)^{-1}\sum_{p_i}(a_{ip_i}-\bar{\Psi}_i)^2 = (k_i-k_{ij})K_i \text{ en}$$

$$(4.117)\quad \mathrm{kov}(y_{ij},y_{ij'}) = -(k_{ij}k_{ij'})^{\frac{1}{2}}k_i^{-1}(k_i-1)^{-1}\sum_{p_i}(a_{ip_i}-\bar{\Psi}_i)^2 = -(k_{ij}k_{ij'})^{\frac{1}{2}}K_i \text{ vir } j \neq j' \text{ waar}$$

$$(4.22)\quad K_i = k_i^{-1}(k_i-1)^{-1}\sum_{p_i}(a_{ip_i}-\bar{\Psi}_i)^2.$$

Die momentematriks van die y_{ij} 's (vir vaste i) is van rang (k-1), want tussen die elemente van elke ry bestaan 'n lineêre verband, naamlik

$$(4.118)\quad \sum_j k_{ij}\,\mathrm{var}(y_{ij}) + \sum_{j\neq j'}\sum (k_{ij}k_{ij'})^{\frac{1}{2}}\,\mathrm{kov}(y_{ij},y_{ij'})$$
$$= \sum_j k_{ij}(k_i-k_{ij})K_i - \sum_{j\neq j'}\sum k_{ij}k_{ij'}K_i$$
$$= \sum_j k_{ij}K_i[k_i-k_{ij}-\sum_{j'(\neq j)}k_{ij'}]$$
$$= 0$$

en geen ander lineêre onafhanklike verband bestaan nie.

Die transformasie van y_{ij} na $\hat{y}_j$ is lineêr en nie-singulier. Gevolglik is die rang van Σ, die momentematriks van die $\hat{y}_j$'s, ook (k-1).

Opmerking.

Dui, soos in die opmerking na lemma 4.5, die momentematriks van enige (k-1) $\hat{y}_j$'s aan deur $\Lambda = \{\lambda_{\zeta\varsigma}\}$ met kofaktore $|\Lambda_{\zeta\varsigma}|$.

STELLING 4.7.

Indien

(4.106) $\underset{p_i}{\text{maks}}[\Psi_i(\delta_{ip_i}) - \bar{\Psi}_i]^2 = o(k_i)$ vir $k_i \to \infty$ en $i=1,2,..,n$

(4.107) (Iedere diskontinuïteit van $F_i(x)) \leq \theta < 1$ vir alle i

(4.108) $\lim\limits_{k_i \to \infty} k_{ij}/k_i \geq \varepsilon > 0$ vir alle i en j,

dan is

$$T = \sum_{\zeta=1}^{k-1} \sum_{\varsigma=1}^{k-1} \frac{|\Lambda_{\zeta\varsigma}|}{|\Lambda|} \hat{y}_\zeta \hat{y}_\varsigma \quad \text{onder } H_0' \text{ asimptoties}$$

χ^2 verdeel met (k-1) grade van vryheid as $N \to \infty$.

Bewys: Uit lemma 4.9 volg dat die momentematriks van die $\hat{y}_\zeta$'s van die vorm

$$e^{-\frac{1}{2}\vec{t}'\Lambda\vec{t}} \quad \text{is,}$$

en m.b.v. lemma 4.10 dat Λ van rang (k-1) is.

Die bewys volg nou m.b.v. §1.4.

Opmerking.

T kan tot meer hanteerbare vorms vereenvoudig word deur sekere beperkings te stel, byvoorbeeld alle $k_{ij} = b$ (sien §4.4). Aangesien die prosedure dieselfde is as in hoofstuk II, gaan ons nie verder daarop in nie.

4.6. DIE ALTERNATIEWE HIPOTESE.

4.6.1. Inleiding.

Tot dusver in hierdie hoofstuk het ons die geval onder H_0 beskou waarby veronderstel word dat alle moontlike rangskikkings van rangnommers binne elke ry gelykkansig is. Hier word egter die meer algemene hipotese H beskou waarby aangeneem word dat al die km items uit dieselfde of verskillende populasies afkomstig is, en wel dat die waarnemings x_{ijh}, $h=1,2,\ldots,k_{ij}$ van die item x_{ij}, afkomstig is uit 'n populasie met kontinue verdelingsfunksie F_{ij}, $i=1,2,\ldots,m$; $j=1,2,\ldots,k$.

Die limietverdeling van die voorgestelde toetsingsgrootheid word onder H bepaal indien $m \to \infty$.

4.6.2. Definisies.

Neem

$$(4.119) \quad t_{ij} = k_{ij}^{-1} \sum_{h=1}^{k_{ij}} \Psi_i(\delta_{ijh})$$

waar $\Psi_i(\delta)$ eindig is vir $0 < \delta < 1$ en $\delta_{ijh} = r_{ijh}/(k_i+1)$ soos in §4.1. Stel

$$(4.120) \quad t'_{ij} = t_{ij} - E(t_{ij}|H) \;,$$

$$(4.121) \quad t_j = m^{-\frac{1}{2}} \sum_{i=1}^{m} t'_{ij} \;,$$

$$(4.122) \quad \sigma_{jj'}^{(i)} = \text{kov}(t'_{ij}, t'_{ij'}|H) = E(t'_{ij}, t'_{ij'}|H) \text{ vir } j \neq j' \text{ want}$$

$$(4.123) \quad E(t'_{ij}|H) = 0;$$

$$(4.124) \quad \sigma_{jj}^{(i)} = \text{var}(t'_{ij}|H),$$

$$(4.125) \quad \sigma'_{jj'} = \text{kov}(t_j, t_{j'}|H) = E(t_j . t_{j'}|H) \text{ vir } j \neq j' \quad \text{want}$$

$$(4.126) \quad E(t_j|H) = 0 \quad \text{en}$$

$$(4.127) \quad \sigma'_{jj} = \sigma_j'^2 = \text{var}(t_j|H).$$

Hieruit volg maklik dat

$$(4.128) \quad \sigma'_{jj'} = m^{-1} \sum_i \sigma_{jj'}^{(i)} \quad \text{met}$$

$$(4.129) \quad \sigma_j'^2 = m^{-1} \sum_i \sigma_{jj}^{(i)}.$$

4.6.3. Die Toetsingsgrootheid T_A. Limietverdeling onder H.

Stel

$$(4.130) \quad V_i = \sum_j c_j t'_{ij}$$ waar die c_j 's willekeurige eindige reële getalle is wat nie almal gelyk is nie.

STELLING 4.8.

Indien

(4.131) $\operatorname{var}(V_i|H) > 0$ vir $i=1,2,\ldots,m$,

dan besit $\sum_j c_j t_j$ onder H asimptoties, vir $m\to\infty$, 'n normaalverdeling.

Bewys: Uit die definisie van Ψ_i volg dat t_{ij}, en dus ook t'_{ij}, eindig is vir alle i en j ($k_i < \infty$ vir alle i).

Dus

(4.132) $E|t'_{ij}|^{2+\delta} < \infty$ vir alle i en j en vir $\delta > 0$.

Soos in stelling 4.1 volg eerstens dat

(4.133) $E|V_i|^{2+\delta} < \infty$ vir alle $i=1,2,\ldots,m$

en vervolgens dat $\sum_j c_j t_j = m^{-\frac{1}{2}} \sum_i V_i$ asimptoties, vir $m\to\infty$, 'n normaalverdeling besit.

Opmerking.

$\operatorname{var}(V_i|H)$ kan ook geskryf word:

(4.134)
$$\begin{aligned}\operatorname{var}(V_i) &= E[V_i - E(V_i)]^2\\ &= E[\sum_j c_j t'_{ij} - E(\sum_j c_j t'_{ij})]^2\\ &= E[\sum_j c_j t'_{ij}]^2 \qquad \text{want } E(t'_{ij}) = 0\\ &= \sum_j \sum_{j'} c_j c_{j'} E(t'_{ij}.t'_{ij'})\\ &= \sum_j \sum_{j'} c_j c_{j'} \sigma_{jj'}^{(i)} .\end{aligned}$$

Lemma 4.11. Indien

(4.131) $\operatorname{var}(V_i|H) > 0$ vir $i=1,2,\ldots,m$ en

(4.135) $\sigma_{jj}^{(i)} > 0$ vir alle i en j,

dan voldoen die t_j 's aan die voorwaardes

(4.136) $\lim_{m\to\infty} \operatorname{var}(t_j|H) > 0$ vir $j=1,2,\ldots,k$ en

(4.137) $\lim_{m\to\infty} E|t_j|^{2+\delta} < \infty$ vir $j=1,2,\ldots,k$ en $\delta > 0$.

Bewys: Uit (4.129) volg

$var(t_j|H) = m^{-1}\sum_i \sigma_{jj}^{(i)}$ sodat m.b.v. (4.135) volg

(4.136) $\lim_{m\to\infty} var(t_j|H) > 0.$

In stelling 4.8 is bewys dat

(4.132) $E|t'_{ij}|^{2+\delta} < \infty$ vir alle i en j en $\delta > 0$.

Soos in lemma 4.4 volg nou dat

(4.137) $\lim_{m\to\infty} E|t_j|^{2+\delta} < \infty$ vir $j=1,2,\ldots,k$ en $\delta > 0$.

Lemma 4.12. Indien voorwaardes (4.131) en (4.135) geld, is die gesamentlike verdeling van die variante t_j asimptoties normaal as $m\to\infty$.

Bewys: Omdat

(4.136) $\lim_{m\to\infty} var(t_j|H) > 0$ vir $j=1,2,\ldots,k$,

(4.137) $\lim_{m\to\infty} E|t_j|^{2+\delta} < \infty$ vir $j=1,2,\ldots,k$ en $\delta > 0$ (lemma 4.11) en omdat $\sum_j c_j t_j$ asimptoties normaal verdeel is as $m\to\infty$ (stelling 4.8), volg die resultaat m.b.v. lemma 2.5.

Opmerkings.

1). Die karakteristieke funksie van die gesamentlike verdeling van die t_j 's is dus

$e^{-\frac{1}{2}\vec{\theta}'\Sigma\vec{\theta}}$ waar $\Sigma = \{\sigma'_{jj}\}$ die momentematriks van die t_j 's voorstel.

2). Neem as voorwaarde

(4.138) Σ is van rang (k-1).

Dui, soos in opmerking 2 na lemma 4.5, die momentematriks van enige (k-1) t_j 's aan deur $\Lambda' = \{\lambda'_{\zeta\varsigma}\}$ met kofaktore $|\Lambda'_{\zeta\varsigma}|$ waar $\zeta,\varsigma = 1,2,\ldots,(k-1)$.

STELLING 4.9.

Indien

(4.131) $\operatorname{var}(V_i|H) > 0$ vir $i=1,2,\ldots,m$,

(4.135) $\sigma_{jj}^{(i)} > 0$ vir alle i en j en

(4.138) Σ van rang $(k-1)$ is,

dan besit die toetsingsgrootheid

(4.139) $$T_A = \sum_{\zeta}\sum_{\zeta'} \frac{|\Lambda'_{\zeta\zeta'}|}{|\Lambda'|} t_\zeta t_{\zeta'}$$

onder H asimptoties, vir $m\to\infty$, 'n χ^2-verdeling met $(k-1)$ grade van vryheid.

Bewys: Die stelling volg direk uit lemma 4.12 en §1.4.

Opmerkings.

1). Die berekening van $\sigma_{jj'}^{(i)}$ skyn baie ingewikkeld te wees, behalwe onder H_o in welke geval ons dit in lemmas 4.2 en 4.3 aangegee het.

2). Deur allerlei vereenvoudigings kan spesiale gevalle uit T_A afgelei word.

3). Die stelling kan ook bewys word deur van karakteristieke funksies gebruik te maak soos FRIEDMAN(1937), maar die voorwaardes is baie meer beperkend as in ons geval (byvoorbeeld: Vir elke vaste j en j' moet $\sigma_{jj'}^{(i)}$ gelyk wees vir alle waardes van $i=1,2,\ldots,m$).

4.7. ASIMPTOTIESE RELATIEWE DOELTREFFENDHEID.

Beskou die alternatiewe hipotese

(4.140) H_a: $F_{ij}(x) = F_i(x+d_j m^{-\frac{1}{2}})$

waar die d_j 's willekeurige eindige reële getalle is waarvoor sonder verlies aan algemeenheid veronderstel kan word dat

(4.141) $\sum_j d_j = 0$.

Ons neem nou deurgaans as voorwaardes

(4.142) $\lim_{m\to\infty} E(u_j|H_a) = B_j < \infty$ vir alle j en

(4.143) $\lim\limits_{m\to\infty} \dfrac{\text{kov}(u_j, u_{j'}|H_a)}{\text{kov}(u_j, u_{j'}|H_o)} = 1$ vir alle j en j' .

Aan laasgenoemde voorwaarde word waarskynlik in die algemeen voldoen, d.w.s. onder algemene reëlmatigheidsvoorwaardes en kontinuïteit van die funksies Ψ_i en f_i, waar $f_i \equiv F_i'$.

In die res van hierdie hoofstuk word veronderstel dat alle Ψ_i en f_i kontinu is, dat daar geen gelyke waarnemings voorkom nie (alle F_i is ook kontinu) en dat alle $k_{ij} = b > 0$.

Nou volg uit (4.17), (4.119), (4.120) en (4.121) met H_a in plaas van H:

(4.144) $t_j = b^{-1} u_j - b^{-1} E(u_j|H_a)$

met $b^{-1} E(u_j|H_a) = m^{-\frac{1}{2}} \sum\limits_i [b^{-1} \sum\limits_\eta E\{\Psi_i(\delta_{ij\eta})|H_a\} - \bar{\Psi}_i]$.

Uit stelling 4.2 volg dat

(4.46) $T = \sum\limits_\zeta \sum\limits_{\zeta'} \dfrac{|\Lambda_{\zeta\zeta'}|}{|\Lambda|} u_\zeta u_{\zeta'}$

onder H_o asimptoties 'n χ^2-verdeling met (k-1) g.v.v. besit as $m\to\infty$. (Aan die voorwaardes van stelling 4.2 word voldoen - sien stelling 4.3 - want vir alle $k_{ij} = b$ herlei T na T_k van stelling 4.3).

Omdat $\lim\limits_{m\to\infty} \text{var}(t_j|H_a)$ eindig en positief is (lemma 4.11), volg dat

(4.145) $0 < \lim\limits_{m\to\infty} \text{var}(u_j|H_a) < \infty$ vir alle j.

Uit stelling 4.8 volg dat $\sum\limits_j c_j t_j$ asimptoties normaal verdeel is onder H_a en voorwaarde (4.131) as $m\to\infty$, waar die c_j 's willekeurige eindige reële getalle is.

Maar

(4.146) $\sum\limits_j c_j t_j = \sum\limits_j c_j [b^{-1} u_j - b^{-1} E(u_j|H_a)]$ uit (4.144)

$= b^{-1} \sum\limits_j c_j u_j - b^{-1} \sum\limits_j c_j E(u_j|H_a)$, sodat ook

$\sum\limits_j c_j u_j$ asimptoties normaal verdeel is onder bogenoemde voorwaardes.

Gevolglik besit die u_j 's onder H_a asimptoties, vir $m\to\infty$, gesamentlik 'n normaalverdeling indien voorwaarde (4.142) ook nog bevredig word.

Uit voorwaarde (4.143) volg dat die u_j 's onder H_a asimptoties, vir $m\to\infty$, dieselfde momentematriks besit as onder H_0, naamlik Λ. Nou volg uit §1.4 dat T onder H_a asimptoties, vir $m\to\infty$, 'n nie-sentrale χ^2-verdeling besit met (k-1) g.v.v. en nie-sentraliteitsparameter

$$(4.147)\quad \lambda^2 = \sum_{\zeta}\sum_{\zeta'} \lim_{m\to\infty} \frac{|\Lambda_{\zeta\zeta'}|}{|\Lambda|} E(u_\zeta|H_a).E(u_{\zeta'}|H_a) \; .$$

Uit (4.142) en (4.145) volg dat λ^2 eindig is.

Aangesien alle $k_{ij} = b > 0$ en daar geen knope voorkom nie, herlei T na:

$$(4.148)\quad T_k = \frac{(k-1)}{k\sigma^2} \sum_j u_j^2 \quad \text{uit (4.63)}$$

$$= \frac{(k-1)}{k\sigma^2} \sum_j [m^{-\frac{1}{2}} \sum_i \sum_{\eta=1}^{b} \{\Psi_i(\delta_{ij\eta}) - \bar{\Psi}_i\}]^2 \text{ uit (4.17)}$$

met σ^2 gegee deur (4.54), (4.55) en (4.56).

Soortgelyk aan T herlei λ^2 m.b.v. (4.59) en (4.60) na:

$$(4.149)\quad \lambda_k^2 = \lim_{m\to\infty} \left\{\frac{2}{Kk} \sum_\zeta [E(u_\zeta|H_a)]^2 + \frac{1}{Kk} \sum_{\zeta\neq\zeta'}\sum E(u_\zeta|H_a)E(u_{\zeta'}|H_a)\right\}$$

$$= \lim_{m\to\infty} \frac{(k-1)}{k\sigma^2}\left\{2\sum_\zeta [E(u_\zeta|H_a)]^2 + \sum_{\zeta\neq\zeta'}\sum E(u_\zeta|H_a)E(u_{\zeta'}|H_a)\right\} .$$

Hierbo is

$$(4.150)\quad E(u_j|H_a) = E[\{m^{-\frac{1}{2}} \sum_i \sum_\eta [\Psi_i(\delta_{ij\eta}) - \bar{\Psi}_i]\} |H_a]$$

$$= m^{-\frac{1}{2}} \sum_i \sum_\eta E[\Psi_i(\delta_{ij\eta})|H_a] - bm^{-\frac{1}{2}} \sum_i \bar{\Psi}_i \; .$$

Nou is, vir vaste i,

$$(4.151)\quad E[\Psi_i(\delta_{ij\eta})|H_a) = E[\Psi_i(\frac{r_{ij\eta}}{kb+1})|H_a]$$

$$= \sum_{p=1}^{kb} \Psi_i(\frac{p}{kb+1}) P[r_{ij\eta} = p|H_a] \; .$$

Dui vervolgens die verdelingsfunksie van die verdeling van $x_{ij\eta}$ aan deur $F''_{ij\eta}$ waarby $i=1,2,\ldots,m$; $j=1,2,\ldots,k$ en $\eta=1,2,\ldots,b$. Dan geld volgens §4.6.1 dat

(4.152) $F''_{ij1} \equiv F''_{ij2} \equiv \ldots \equiv F''_{ijb} \equiv F_{ij}$.

Vir vaste i is daar kb verdelingsfunksies $F''_{ij\eta}$; $j=1,2,\ldots,k$; $\eta=1,2,\ldots,b$. Dui nou hierdie verdelingsfunksies aan deur F'_{ig} waar $g=1,2,\ldots,kb$ sodanig dat $F'_{i1} \equiv F''_{i11}$, $F'_{i2} \equiv F''_{i12}$, ... , $F'_{ib} \equiv F''_{i1b}$, $F'_{i,b+1} \equiv F''_{i21}$, $F'_{i,b+2} \equiv F''_{i22}$, ens.

Uit (4.151) volg dus nou verder

$$(4.153)\quad E[\Psi_i(\delta_{ij\eta})|H_a] = \sum_p \Psi_i\left(\frac{p}{kb+1}\right) P[r_{ij\eta} = p|H_a]$$

$$= \Psi_i\left(\frac{1}{kb+1}\right) \int_{-\infty}^{\infty} \prod_{\substack{p'=1\\(\neq g)}}^{kb} [1 - F'_{ip'}]\, dF'_{ig} +$$

$$+ \Psi_i\left(\frac{2}{kb+1}\right) \sum_{\substack{g'=1\\(\neq g)}}^{kb} \int_{-\infty}^{\infty} \prod_{\substack{p'=1\\(\neq g,g')}}^{kb} [1 - F'_{ip'}]F'_{ig'}\, dF'_{ig} +$$

$$+ \Psi_i\left(\frac{3}{kb+1}\right) \sum_{\substack{g'<g''\\(\neq g)}}^{kb}\sum \int_{-\infty}^{\infty} \prod_{\substack{p'=1\\(\neq g,g',g'')}}^{kb} [1 - F'_{ip'}]F'_{ig'}\, F'_{ig''}\, dF'_{ig} + \ldots\ldots +$$

$$+ \Psi_i\left(\frac{kb}{kb+1}\right) \int_{-\infty}^{\infty} \prod_{\substack{p'=1\\(\neq g)}}^{kb} F'_{ip'}\, dF'_{ig}$$

waar $F'_{ig} \equiv F''_{ij\eta}$, d.w.s $g = (j-1)b + \eta$ [11)]

$$= \sum_{p=1}^{kb} \Psi_i\left(\frac{p}{kb+1}\right) \cdot$$

$$\cdot \sum_{\substack{g'<g''<\ldots<g^{(p-1)}\\(\text{almal } \neq g)}}^{kb}\sum \ldots \sum \int_{-\infty}^{\infty} \prod_{\substack{p'=1\\(\neq \{g,g',\ldots,g^{(p-1)}\})}}^{kb} [1-F'_{ip'}]F'_{ig'}F'_{ig''}\cdots F'_{ig^{(p-1)}}\, dF'_{ig} \,.$$

λ_k^2 word nou gevind deur terugsubstitusie van (4.153) en (4.150) in (4.149). Bostaande uitdrukking is in die algemeen onhanteerbaar. Ons beskou slegs die volgende spesiale geval.

---

11). $\prod_{p'=1}^{kb} F'_{ip'}$ dui die produk $F'_{i1}F'_{i2}\cdot \ldots \cdot F'_{i,kb}$ aan.

Spesiale geval: $\Psi_i(\delta_{ij\eta}) \equiv r_{ij\eta}$.

Nou is $\bar{\Psi}_i = (kb)^{-1} \sum_j \sum_\eta r_{ij\eta}$ uit (4.14)

$= \frac{1}{2}(kb+1)$,

(4.154) $u_j = m^{-\frac{1}{2}} \sum_i \sum_\eta [r_{ij\eta} - \frac{1}{2}(kb+1)]$ en

(4.155) $\sigma^2 = b^2(k-1)(kb+1)/12$ uit (4.65).

T_k herlei dus tot

(4.156) $T_k' = 12[mb^2k(kb+1)]^{-1} \sum_j [\sum_i \sum_\eta \{r_{ij\eta} - \frac{1}{2}(kb+1)\}]^2$.

Uit (4.153) volg:

$$E(r_{ij\eta}|H_a) = \int_{-\infty}^{\infty} \prod_{\substack{p'=1\\(\neq g)}}^{kb} [1 - F'_{ip'}]\, dF'_{ig} +$$

$$+ 2 \sum_{\substack{g'=1\\(\neq g)}}^{kb} \int_{-\infty}^{\infty} \prod_{\substack{p'=1\\(\neq g,g')}}^{kb} [1 - F'_{ip'}]F'_{ig'}\, dF'_{ig} +$$

$$+ 3 \sum_{\substack{g'<g''\\(\neq g)}}^{kb} \int_{-\infty}^{\infty} \prod_{\substack{p'=1\\(\neq g,g',g'')}}^{kb} [1-F'_{ip'}]F'_{ig'} F'_{ig''} dF'_{ig} + \ldots. +$$

$$+ kb \int_{-\infty}^{\infty} \prod_{\substack{p'=1\\(\neq g)}}^{kb} F'_{ip'}\, dF'_{ig} .$$

Deur die terme uit te skryf en te hergroepeer, volg dan:

(4.157) $E(r_{ij\eta}|H_a) = 1 + \sum_{g'(\neq g)}^{kb} \int_{-\infty}^{\infty} F'_{ig'}\, dF'_{ig}$ waarby

$F'_{ig} \equiv F''_{ij\eta} \equiv F_{ij}$, d.w.s. $g = (j-1)b + \eta$.

Hierby is

(4.158) $$\sum_{g'(\neq g)}^{kb} \int_{-\infty}^{\infty} F'_{ig'}\, dF'_{ig} = \sum_{j'=1}^{k} \sum_{\eta'=1}^{b} \int_{-\infty}^{\infty} F''_{ij'\eta'}\, dF''_{ij\eta} - \int_{-\infty}^{\infty} F'_{ig}\, dF'_{ig}$$

$$= b \sum_{j'} \int_{-\infty}^{\infty} F_{ij'}\, dF_{ij} - \tfrac{1}{2}$$

$$= b \sum_{j'} \int_{-\infty}^{\infty} F_i(x+d_{j'}m^{-\frac{1}{2}})\, dF_i(x+d_jm^{-\frac{1}{2}}) - \tfrac{1}{2}$$ met behulp van (4.140),

$$= b \sum_{j'} \int_{-\infty}^{\infty} F_i[x + (d_{j'} - d_j)m^{-\frac{1}{2}}]\, dF_i(x) - \tfrac{1}{2}$$

$$= b \sum_{j'} \int_{-\infty}^{\infty} [F_i(x) + (d_{j'} - d_j)m^{-\frac{1}{2}} f_i\{x + \theta(d_{j'} - d_j)m^{-\frac{1}{2}}\}]dF_i(x) - \tfrac{1}{2}$$

m.b.v. Taylor se middelwaardestelling, waar $0<\theta<1$, aangesien f_i orals kontinu veronderstel word

$$= b \sum_{j'}[\tfrac{1}{2} + (d_{j'} - d_j)m^{-\frac{1}{2}} \int_{-\infty}^{\infty} f_i\{x + \theta(d_{j'} - d_j)m^{-\frac{1}{2}}\}dF_i(x)] - \tfrac{1}{2}$$

$$= \tfrac{1}{2}(kb-1) + bm^{-\frac{1}{2}} \sum_{j'} (d_{j'} - d_j) \int_{-\infty}^{\infty} f_i(x')dF_i(x) \text{ waar}$$

$$x' = x + \theta(d_{j'} - d_j)m^{-\frac{1}{2}}.$$

Nou is vir die geval $\Psi_i(\delta_{ij\eta}) \equiv r_{ij\eta}$,

(4.159) $$E(u_j|H_a) = m^{-\frac{1}{2}} \sum_i \sum_\eta E[r_{ij\eta}|H_a] - \tfrac{1}{2}m^{+\frac{1}{2}}b(kb+1) \text{ met behulp van (4.154)}$$

$$= m^{-\frac{1}{2}} \sum_i \sum_\eta [1 + \tfrac{1}{2}(kb-1) + bm^{-\frac{1}{2}} \sum_{j'} (d_{j'} - d_j) \int_{-\infty}^{\infty} f_i(x')dF_i(x)] - $$
$$- \tfrac{1}{2}m^{+\frac{1}{2}}b(kb+1) \text{ uit (4.157) en (4.158)}$$

$$= b^2m^{-1} \sum_{j'} (d_{j'} - d_j) \sum_i \int_{-\infty}^{\infty} f_i(x')dF_i(x).$$

Indien vir m voldoende groot geld

(4.160) $$\int_{-\infty}^{\infty} f_i(x + cm^{-\frac{1}{2}})dF_i(x) < \infty$$ waar c 'n willekeurige konstante is, dan geld, omdat f_i kontinu veronderstel word,

$$\int_{-\infty}^{\infty} f_i[x + \theta(d_{j'} - d_j)m^{-\frac{1}{2}}]dF_i(x) = \int_{-\infty}^{\infty} f_i(x)dF_i(x) + o(1),$$
waarby $0<\theta<1$.

Aan voorwaarde (4.160) word byvoorbeeld voldoen indien f_i gelykmatig begrens is.

Nou is

(4.161) $$E(u_j|H_a) = b^2m^{-1} \sum_{j'} (d_{j'} - d_j) \sum_i \int_{-\infty}^{\infty} f_i(x)dF_i(x) + o(1)$$

$$= -b^2km^{-1}d_j \sum_i \int_{-\infty}^{\infty} f_i(x)dF_i(x) + o(1) \text{ met behulp van (4.141).}$$

Daar word dus in hierdie geval voldoen aan voorwaarde (4.142).

Uit (4.149), (4.155) en (4.161) volg:

$$(4.162)\quad \lambda_k^2 = \frac{12}{b^2k(kb+1)} \lim_{m\to\infty} [2\sum_\zeta \{E(u_\zeta|H_a)\}^2 + \sum_{\zeta\neq\zeta'}\sum E(u_\zeta|H_a)E(u_{\zeta'}|H_a)]$$

$$= \frac{12}{b^2k(kb+1)} \lim_{m\to\infty} [2\sum_\zeta\Big\{\frac{b^2k}{m}d_\zeta \sum_i \int_{-\infty}^{\infty} f_i(x)dF_i(x)\Big\}^2 + \sum_{\zeta\neq\zeta'}\sum\Big\{\frac{b^2k}{m}d_\zeta \sum_i \int_{-\infty}^{\infty} f_i(x)dF_i(x)\Big\}\Big\{\frac{b^2k}{m}d_{\zeta'} \sum_i \int_{-\infty}^{\infty} f_i(x)dF_i(x)\Big\}]$$

$$= \frac{12b^2k}{(kb+1)} \lim_{m\to\infty} m^{-2}[\sum_i \int_{-\infty}^{\infty} f_i(x)dF_i(x)]^2[2\sum_\zeta d_\zeta^2 + \sum_{\zeta\neq\zeta'}\sum d_\zeta d_{\zeta'}]$$

$$= \frac{12b^2k}{(kb+1)} \lim_{m\to\infty} m^{-2}[\sum_i \int_{-\infty}^{\infty} f_i(x)dF_i(x)]^2[\sum_\zeta d_\zeta^2 + (\sum_\zeta d_\zeta)^2]$$

$$= \frac{12b^2k}{(kb+1)} \lim_{m\to\infty} m^{-2}[\sum_i \int_{-\infty}^{\infty} f_i(x)dF_i(x)]^2 \cdot \sum_{j=1}^{k} d_j^2 \quad \text{met behulp}$$

van (4.141) aangesien $\zeta=1,2,\ldots,(k-1)$.

Vir $b = 1$ herlei hierdie uitdrukking tot:

$$(4.163)\quad \lambda_k^2 = \frac{12k}{(k+1)} \lim_{m\to\infty} m^{-2}[\sum_i \int_{-\infty}^{\infty} f_i^2(x)\, dx]^2 \sum_j d_j^2$$

wat dieselfde is as dié in VAN ELTEREN en NOETHER(1959) vergelyking (9).

Beskou ons vervolgens die meer eenvoudige hipotese

$$(4.164)\quad H_a':\ F_{ij}(x) = F[x+(\delta_i+d_j)m^{-\frac{1}{2}}]$$

waar die δ_i 's en d_j 's willekeurige eindige reële getalle is en die d_j 's nog voldoen aan die voorwaarde

$$(4.141)\quad \sum_j d_j = 0,$$

dan herlei λ_k^2 uit (4.162) na:

$$(4.165)\quad \lambda_k^2 = \frac{12b^2k}{(kb+1)} \lim_{m\to\infty} m^{-2}[m \int_{-\infty}^{\infty} f^2(x)\, dx]^2 \sum_j d_j^2$$

$$= \frac{12b^2k}{(kb+1)}\Big[\int_{-\infty}^{\infty} f^2(x)\, dx\Big]^2 \sum_j d_j^2 .$$

Die nie-sentraliteitsparameter van die F-toets vir hierdie geval kan gevind word deur die voorbeeld van TANG(1938) op bladsye 137-138 na hierdie geval te veralgemeen, waardeur ons vind

(4.166) $\lambda_F^2 = b\sigma_F^{-2} \sum_j d_j^2$. (Vergelyk ook VAN ELTEREN en NOETHER(1959) vir die geval b = 1).

Die asimptotiese relatiewe doeltreffendheid van T_k' met betrekking tot die F-toets is dus

(4.167) $$E_{T_k', F} = \lambda_k^2 / \lambda_F^2$$

$$= \frac{12kb}{(kb+1)} \sigma_F^2 \left[\int_{-\infty}^{\infty} f^2(x)\, dx \right]^2 .$$

Opmerkings.

1). Stel b = 1, dan herlei T_k' in (4.156) na

(4.168) $$T_k'' = \frac{12}{mk(k+1)} \sum_j [\sum_i \{ r_{ij} - \tfrac{1}{2}(k+1) \}]^2$$

$$= \frac{12}{mk(k+1)} \sum_j [\sum_i r_{ij} - \tfrac{1}{2}m(k+1)]^2 .$$

2). Vir b = 1 is

(4.169) $$E_{T_k'', F} = \frac{12k}{(k+1)} [\sigma_F \int_{-\infty}^{\infty} f^2(x)\, dx]^2$$

wat dieselfde is as die resultaat gevind deur VAN ELTEREN en NOETHER(1959) vergelyking (7).

3). Vir die normaalverdeling is uit (4.167)

(4.170) $$E_{T_k'', F} = \frac{kb}{(kb+1)} . \frac{3}{\pi} .$$

Namate kb toeneem, benader $E_{T_k'', F}$ die waarde $3/\pi = 0.955$ (sien VAN ELTEREN en NOETHER(1959) vergelyking (7N) en die daaropvolgende tabel).

4). Vir die homogene verdeling is uit (4.167)

(4.171) $$E_{T_k'', F} = \frac{kb}{(kb+1)}$$

wat die waarde 1 benader as kb toeneem.

Opmerking.

Dit blyk dat T_k' met $\Psi_i(\delta_{ij\eta}) \equiv r_{ij\eta}$ asimptoties onafhanklik is van verskuiwingseffekte tussen die verskillende rye - sien (4.162). Die toets gebaseer op T_k' kan dus gebruik word om vir verskuiwingseffekte tussen die kolomme te toets onafhanklik van eventuele verskuiwingseffekte tussen rye. Soortgelyk kan rangnommers opnuut binne die kolomme toegeken word en die toets aangepas word om te toets vir verskuiwings tussen die rye (onafhanklik van die voorkoms, al dan nie, van kolom-verskuiwings). Die toetse vir ry- en kolom-effekte moet egter op afsonderlike onafhanklike stelle waarnemings toegepas word.

Slotopmerking.

In bostaande spesiale geval is dit redelik maklik om in te sien dat onder sekere reëlmatigheidsvoorwaardes, waaronder die kontinuïteit van alle F_i, aan (4.143) voldoen word.

HOOFSTUK V.

DIE ANALISE VAN VARIANSIE.

5.1. INLEIDING.

Word in die geval van §4.5.2 rangnommers aan al die N waarnemings gesamentlik toegeken, kry ons 'n twee-faktor variansie-analise. Ons gaan dié geval in hierdie hoofstuk behandel.

5.2. KORT OORSIG OOR REEDS BEKENDE WERK.

MOOD(1950) het 'n verdelingsvrye prosedure behandel waarvolgens vir ry-effekte, kolomeffekte, interaksie en kombinasies daarvan getoets kan word in die tweefaktor rangskikking met b waarnemings per sel. Die toetsingsgroothede is gebaseer op die afwykings van die waarnemings vanaf hulle mediane. Die verskillende toetsingsgroothede word kortliks bespreek (sien MOOD(1950) en TATE en CLELLAND(1957)).

Gestel daar is m rye, k kolomme en b waarnemings per sel. Laat $A = \frac{1}{2}mb$ of $\frac{1}{2}(mb-1)$, naamlik dié een waarvoor A 'n heelgetal is. Laat $\hat{x}_{.j}$ die mediaan van die mb waarnemings in die j^{de} kolom wees en $\hat{x}_{i.}$ die mediaan van die kb waarnemings in die i^{de} ry.

5.2.1. Mediaantoets vir Interaksie.

MOOD(1950) het as toetsingsgrootheid vir interaksie voorgestel

$$\chi_I^2 = P^2 b \sum_i \sum_j \frac{(P_{ij} - P_{i.}P_{.j}/P)^2}{P_{i.}P_{.j}(P_{i.}P_{.j} - b)} \qquad 1)$$

waar die simbole die volgende betekenis het: Word ry- en/of kolommediane van die waarnemings in die tabel afgetrek, kry ons negatiewe waardes (aangedui deur

---

1). In hierdie hoofstuk deurloop i,i' die waardes 1,2,...,m en j,j',j" die waardes 1,2,...,k, tensy anders gespesifiseer.

minustekens), nulle, en positiewe waardes (aangedui deur plustekens). Nou is

P = aantal plustekens plus helfte van die nulle in die tabel,

P_{ij} = aantal plustekens plus helfte van die nulle in die sel (i,j),

$P_{i.}, P_{.j}$ = aantal plustekens plus helfte van die nulle in die i^{de} ry respektiewelik j^{de} kolom.

Hierdie toets vir interaksie word toegepas nadat ry- en kolomeffekte (indien enige) verwyder is deur herhaalde aftrekking. Trek naamlik die kolommediane af van die waarnemings in die onderskeie kolomme. (Daar kan netsowel met rye begin word). Vir die nuwe stelsel word ry-mediane bepaal en van die waarnemings in die onderskeie rye afgetrek. Hierdie prosedure word herhaal totdat die ry- en kolommediane almal prakties nul is.

Volgens MOOD(1950) is χ^2_I asimptoties χ^2 verdeel met $(m-1)(k-1)$ g.v.v. as $b \to \infty$. (Die benadering is bevredigend as $P_{i.}P_{.j}/P \geq 2$ en $(m-1)(k-1) \geq 2$).

5.2.2. Mediaantoets vir ry- en interaksie-effekte gesamentlik.

Reduseer die waarnemings binne elke kolom om hul kolommediaan sodat die toets onafhanklik is van moontlike kolomeffekte.

Die toetsingsgrootheid

$$\chi^2_{mI} = \frac{m(mb-1)}{A(mb-A)} \sum_i \sum_j (P_{ij} - \frac{A}{m})^2$$

besit asimptoties 'n χ^2-verdeling met $k(m-1)$ g.v.v. as $b \to \infty$. (Die benadering is goed vir $b \geq 5$ of $mkb \geq 20$).

Soortgelyk vir kolom- en interaksie-effekte gesamentlik.

5.2.3. Mediaantoets vir ry-effekte as interaksie nie betekenisvol is nie.

Reduseer die waarnemings binne elke kolom om hul kolommediaan (om moontlike kolomeffekte uit te skakel). Die toetsingsgrootheid

$$\chi_m^2 = \frac{m(mb-1)}{kA(mb-A)} \sum_i (P_{i.} - \frac{kA}{m})^2$$

is asimptoties χ^2 verdeel met $(m-1)$ g.v.v. as $b \to \infty$.

Soortgelyk vir kolomeffekte.

5.3. 'n NUWE UITBREIDING.

5.3.1. Inleiding.

In hierdie hoofstuk word twee tipes toetse bespreek, naamlik toetse vir homogeniteit van gemiddeldes en homogeniteit van variansies.

In die behandeling van MOOD(1950) hierbo word gewerk met die afwykings van die gegewe waarnemings vanaf hulle mediane.

By toetse vir die homogeniteit van variansies kan dieselfde gedagte verder uitgebrei word deur die waarnemings te reduseer om 'n bepaalde lokaliteitsparameter (rekenkundige gemiddeld, mediaan, modus, ens.), aan die gereduseerde waardes rangnommers toe te ken en dan 'n toetsingsgrootheid te konstrueer as 'n funksie van die rangnommers. Die resultate van hoofstuk II kan vermoedelik na hierdie geval uitgebrei word onderhewig aan bepaalde voorwaardes, bv. simmetrie van die verdeling.

SUKHATME(1958) het in die twee-steekproef geval vir verspreidingsalternatiewe soos volg te werk gegaan: Hy beskou onafhanklike steekproewe $X_1, X_2, \ldots, X_m$ en $Y_1, Y_2, \ldots, Y_n$ uit populasies met verdelingsfunksies $F(x-\xi)$ en $G(x-\eta)$ respektiewelik waar ξ en η onbekende populasieparameters is. SUKHATME(1958) maak skattings

$\hat{\xi}(X_1, X_2, \ldots, X_m)$ en $\hat{\eta}(Y_1, Y_2, \ldots, Y_n)$ van ξ en η respektiewelik. Dan werk hy met 'n toetsingsgrootheid gebaseer op die U-statistiese groothede van HOEFFDING(1948), toegepas op die waarnemings $X_i - \hat{\xi}$ en $Y_j - \hat{\eta}$. SUKHATME toon onder andere aan dat sy toetsingsgrootheid asimptoties normaal verdeel en asimptoties verdelingsvry is.

Dié geval is deur CROUSE(1960) uitgebrei na m steekproewe.

In hierdie hoofstuk gebruik ons egter 'n geheel ander benadering wat, soos aangetoon sal word, 'n groot ooreenkoms toon met die gewone analise van variansie.

5.3.2. Definisies.

Gestel ons het 'n stel waarnemings x_{ijh} verdeel in m rye en k kolomme met b waarnemings per sel waar m en k eindig is. Laat N=mkb. Ons werk onder die nulhipotese H_o dat alle waarnemings x_{ijh} afkomstig is uit dieselfde populasie met verdelingsfunksie F.

Rangskik die waarnemings x_{ijh} gesamentlik volgens grootte in stygende volgorde. Ken aan al die N waardes x_{ijh} rangnommers toe waarby r_{ijh} die rangnommer is van x_{ijh}. Dui die gerangskikte x_{ijh}-waardes aan deur

$$z_1 \leq z_2 \leq \ldots \leq z_N.$$

Ons beskou nou 'n funksie van die rangnommers, naamlik $\Psi_N[r_{ijh}/(N+1)]$.

Indien daar tussen die waardes x_{ijh} knope van lengtes $g_1, g_2, \ldots, g_\lambda$ voorkom, definieer ons (sien hoofstuk II):

$$(5.1)\quad G_\alpha = g_1 + g_2 + \ldots + g_\alpha \quad \text{met } \alpha \in \{1, 2, \ldots, \lambda\} \quad \text{en}$$
$$G_o = 0,$$

waarby aangeneem word dat daar λ knope voorkom. Ongelyke waardes word beskou as knope van lengte een.

Laat

(5.2) $a_p = g_\alpha^{-1} \sum_{\mu=1}^{g_\alpha} \Psi_N[(G_{\alpha-1}+\mu)/(N+1)]$ wanneer

$G_{\alpha-1}+1 \leq p \leq G_\alpha$.

Definieer

(5.3) $a_{ijh} = a_p$ indien $x_{ijh} = z_p$.

Indien alle knope van lengte een is, is

$a_{ijh} = \Psi_N[r_{ijh}/(N+1)]$ vir alle i, j en h.

Laat

(5.4) $a_{ij.} = b^{-1} \sum_h a_{ijh}$ [2)]

(5.5) $a_{i..} = (kb)^{-1} \sum_j \sum_h a_{ijh}$

(5.6) $a_{.j.} = (mb)^{-1} \sum_i \sum_h a_{ijh}$

(5.7) $a_{...} = (mkb)^{-1} \sum_i \sum_j \sum_h a_{ijh} = N^{-1} \sum_{p=1}^{N} a_p$

(5.8) $\sigma_a^2 = N^{-1} \sum_i \sum_j \sum_h (a_{ijh} - a_{...})^2 = N^{-1} \sum_p (a_p - a_{...})^2$.

5.3.3. Die Toetsingsgroothede T_s, T_m, T_k en T_I .

Ons definieer die volgende toetsingsgroothede:

(5.9) $T_s = (N-1)bN^{-1}\sigma_a^{-2} \sum_i \sum_j (a_{ij.} - a_{...})^2$,

(5.10) $T_m = (N-1)kbN^{-1}\sigma_a^{-2} \sum_i (a_{i..} - a_{...})^2$,

(5.11) $T_k = (N-1)mbN^{-1}\sigma_a^{-2} \sum_j (a_{.j.} - a_{...})^2$ en

(5.12) $T_I = (N-1)bN^{-1}\sigma_a^{-2} \sum_i \sum_j (a_{ij.} - a_{i..} - a_{.j.} + a_{...})^2$.

Lemma 5.1. $T_s = T_m + T_k + T_I$.

Bewys: $\sum_i \sum_j (a_{ij.} - a_{...})^2 =$

$= \sum_i \sum_j [a_{ij.} - a_{i..} - a_{.j.} + a_{...} + (a_{i..} - a_{...}) + (a_{.j.} - a_{...})]^2$

$= \sum_i \sum_j (a_{ij.} - a_{i..} - a_{.j.} + a_{...})^2 + k\sum_i (a_{i..} - a_{...})^2 + m\sum_j (a_{.j.} - a_{...})^2$.

2). In hierdie hoofstuk deurloop h,h' die waardes 1,2,..,b en p die waardes 1,2,...,N, tensy anders gespesifiseer.

Nou is

$$T_s = (N-1)bN^{-1}\sigma_a^{-2}\sum_i\sum_j(a_{ij.}-a_{...})^2$$

$$= \frac{(N-1)b}{N\sigma_a^2}\sum_i\sum_j(a_{ij.}-a_{i..}-a_{.j.}+a_{...})^2 + \frac{(N-1)bk}{N\sigma_a^2}\sum_i(a_{i..}-a_{...})^2 + $$

$$+ \frac{(N-1)mb}{N\sigma_a^2}\sum_j(a_{.j.}-a_{...})^2$$

$$= T_I + T_m + T_k \text{ .}$$

5.3.4. Limietverdelings onder H_o van T_s, T_m, T_k en T_I as $b\to\infty$.

STELLING 5.1.

Indien

(5.13) $\underset{p}{\text{maks}}(a_p-a_{...})^2 = o(N)$ en

(5.14) (Iedere diskontinuïteit van $F(x)) \leq \theta < 1$,

dan is T_s onder H_o asimptoties χ^2 verdeel met $(mk-1)$ grade van vryheid as $N\to\infty$ 3).

Bewys: Definieer

(5.15) $t'_{ij} = \sum_h(a_{ijh} - a_{...})$

$= b(a_{ij.} - a_{...})$.

Omdat t'_{ij} analoog is aan t_i van hoofstuk II, volg netsoos in lemmas 2.1, 2.2 en 2.3 dat

(5.16) $E(t'_{ij}) = bN^{-1}\sum_i\sum_j\sum_h(a_{ijh} - a_{...}) = 0$,

(5.17) $\text{var}(t'_{ij}) = b(N-b)(N-1)^{-1}\sigma_a^2$ en

(5.18) $\text{kov}(t'_{ij}, t'_{i'j'}) = -b^2(N-1)^{-1}\sigma_a^2$ mits $i'\neq i$ en/of $j'\neq j$.

Soos in stelling 2.4 volg dat die variansie-kovariansiematriks van die variante

$$\hat{t}'_{ij} = (N-b)^{\frac{1}{2}}\, t'_{ij}[N\,\text{var}(t'_{ij})]^{-\frac{1}{2}}$$

van rang $(mk-1)$ is.

---

3). Die bewerings $N\to\infty$ en $b\to\infty$ kan as ekwivalent beskou word omdat m en k vas gehou word as $b\to\infty$.

Kies nou willekeurige eindige konstantes c'_{ij} wat nie almal gelyk is nie.

Onder voorwaardes (5.13) en (5.14) geld, soos in stelling 2.3, dat $\sum_i \sum_j c'_{ij} t'_{ij}$ asimptoties normaal verdeel is as $N \to \infty$. Ten slotte volg, soos in stelling 2.5, dat $\sum_i \sum_j (N-b)N^{-1}(t'_{ij})^2[\operatorname{var}(t'_{ij})]^{-1}$ asimptoties, vir $N \to \infty$, 'n χ^2-verdeling met (mk-1) g.v.v. besit.

Hierby is:

$$\sum_i \sum_j (N-b)N^{-1}(t'_{ij})^2[\operatorname{var}(t'_{ij})]^{-1}$$

$$= (N-b)N^{-1} \sum_i \sum_j [b(a_{ij.}-a_{...})]^2[b(N-b)(N-1)^{-1}\sigma_a^2]^{-1}$$

$$= (N-1)bN^{-1}\sigma_a^{-2} \sum_i \sum_j (a_{ij.} - a_{...})^2$$

$$= T_s \ .$$

Hiermee is die stelling bewys.

Opmerking.

$$(5.19)\quad E(T_s) = E[(N-1)bN^{-1}\sigma_a^{-2} \sum_i \sum_j (a_{ij.} - a_{...})^2]$$

$$= (N-1)bN^{-1}\sigma_a^{-2} \sum_i \sum_j E(a_{ij.} - a_{...})^2$$

$$= (N-1)bN^{-1}\sigma_a^{-2} \sum_i \sum_j E[b^{-1}\sum_h (a_{ijh}-a_{...})b^{-1}\sum_{h'}(a_{ijh'} - a_{...})]$$

$$= (N-1)N^{-1}b^{-1}\sigma_a^{-2} \sum_i \sum_j E[\sum_h (a_{ijh}-a_{...})\sum_{h'}(a_{ijh'} - a_{...})]$$

$$= (N-1)N^{-1}b^{-1}\sigma_a^{-2} \sum_i \sum_j \operatorname{var}(t'_{ij}) \qquad \text{omdat } E(t'_{ij}) = 0$$

$$= (N-1)N^{-1}b^{-1}\sigma_a^{-2} \sum_i \sum_j b(N-b)(N-1)^{-1}\sigma_a^2 \qquad \text{uit } (5.17)$$

$$= mk(mkb-b)/N$$

$= (mk-1) =$ aantal g.v.v. van T_s, soos wat die geval behoort te wees.

Definieer vervolgens

$$(5.20)\quad t_{ij} = \sum_h (a_{ijh} - a_{i..} - a_{.j.} + a_{...})$$

$$(5.21)\quad t_{i.} = \sum_j \sum_h (a_{ijh} - a_{...}) \quad \text{en}$$

$$(5.22)\quad t_{.j} = \sum_i \sum_h (a_{ijh} - a_{...}) \ .$$

Lemma 5.2. Die variante t_{ij}, $t_{i.}$ en $t_{.j}$ is, vir enige vaste i en j, onderling ongekorreleerd.

Bewys: $\operatorname{kov}(t_{i.}, t_{.j}) = E[\sum_{j'}\sum_{h}(a_{ij'h}-a_{...})\sum_{i'}\sum_{h'}(a_{i'jh'}-a_{...})]$

$$= \sum_{i'}\sum_{j'} E[\sum_{h}(a_{ij'h}-a_{...})\sum_{h'}(a_{i'jh'}-a_{...})]$$

$$= E[\sum_{h}(a_{ijh}-a_{...})\sum_{h'}(a_{ijh'}-a_{...})] +$$

$$+ \sum\sum_{(i,j')\neq(i',j)} \sum\sum E[\sum_{h}(a_{ij'h}-a_{...})\sum_{h'}(a_{i'jh'}-a_{...})] \text{.} \quad 4)$$

Neem die terme afsonderlik.

(5.23) $E[\sum_{h}(a_{ijh}-a_{...})\sum_{h'}(a_{ijh'}-a_{...})] = E(t_{ij}'^2)$

$$= b(N-b)(N-1)^{-1}\sigma_a^2$$

uit (5.17),

(5.24) $E[\sum_{h}(a_{ij'h}-a_{...})\sum_{h'}(a_{i'jh'}-a_{...})] = -\,b^2(N-1)^{-1}\sigma_a^2$ uit (5.18) waarby $i'\neq i$ en/of $j'\neq j$ in laasgenoemde geval.

Dus $\operatorname{kov}(t_{i.}, t_{.j}) = b(N-b)(N-1)^{-1}\sigma_a^2 - (mk-1)b^2(N-1)^{-1}\sigma_a^2$

$= 0$ want $N = mkb$.

$t_{i.}$ en $t_{.j}$ is dus ongekorreleerd. Dit geld vir enige moontlike waardes van i en j.

Verder is:

$\operatorname{kov}(t_{i.}, t_{ij}) = E[\sum_{j'}\sum_{h}(a_{ij'h}-a_{...})\sum_{h'}(a_{ijh'}-a_{i..}-a_{.j.}+a_{...})]$

$$= E[\sum_{j'}\sum_{h}(a_{ij'h}-a_{...})\sum_{h'}(a_{ijh'}-a_{...})-\sum_{j'}\sum_{h}(a_{ij'h}-a_{...})\sum_{h'}(a_{i..}-a_{...})$$

$$-\sum_{j'}\sum_{h}(a_{ij'h}-a_{...})\sum_{h'}(a_{.j.}-a_{...})] \text{.}$$

Neem die terme afsonderlik.

$E[\sum_{j'}\sum_{h}(a_{ij'h}-a_{...})\sum_{h'}(a_{ijh'}-a_{...})] =$

$$= E[\sum_{h}(a_{ijh}-a_{...})\sum_{h'}(a_{ijh'}-a_{...})] +$$

$$+ \sum_{j'(\neq j)} E[\sum_{h}(a_{ij'h}-a_{...})\sum_{h'}(a_{ijh'}-a_{...})]$$

$$= b(N-b)(N-1)^{-1}\sigma_a^2 - (k-1)b^2(N-1)^{-1}\sigma_a^2$$

$$= b^2k(m-1)(N-1)^{-1}\sigma_a^2 \text{.}$$

---

4). Hierdie sommasie beteken $i'\neq i$ en/of $j'\neq j$.

$$E[\sum_{j'}\sum_{h}(a_{ij'h} - a_{...})\sum_{h'}(a_{i..} - a_{...})]$$

$$= b\sum_{j'} E[\sum_{h}(a_{ij'h} - a_{...})(a_{i..} - a_{...})]$$

$$= b\sum_{j'} E[\sum_{h}(a_{ij'h} - a_{...})k^{-1}b^{-1}\sum_{j''}\sum_{h'}(a_{ij''h'} - a_{...})]$$

$$= k^{-1}\sum_{j'}\sum_{j''} E[\sum_{h}(a_{ij'h} - a_{...})\sum_{h'}(a_{ij''h'} - a_{...})]$$

$$= k^{-1}\sum_{j''} E[\sum_{h}(a_{ij''h} - a_{...})\sum_{h'}(a_{ij''h'} - a_{...})] +$$

$$+ k^{-1}\sum_{j'\neq j''}\sum E[\sum_{h}(a_{ij'h} - a_{...})\sum_{h'}(a_{ij''h'} - a_{...})]$$

$$= k^{-1}kb(N-b)(N-1)^{-1}\sigma_a^2 - k^{-1}k(k-1)b^2(N-1)^{-1}\sigma_a^2$$

$$= b^2k(m-1)(N-1)^{-1}\sigma_a^2 .$$

$$E[\sum_{j'}\sum_{h}(a_{ij'h} - a_{...})\sum_{h'}(a_{.j.} - a_{...})]$$

$$= b\sum_{j'} E[\sum_{h}(a_{ij'h} - a_{...})m^{-1}b^{-1}\sum_{i'}\sum_{h'}(a_{i'jh'} - a_{...})]$$

$$= m^{-1}\sum_{j'}\sum_{i'} E[\sum_{h}(a_{ij'h} - a_{...})\sum_{h'}(a_{i'jh'} - a_{...})]$$

$$= 0 \text{ soos in } \mathrm{kov}(t_{i.}, t_{.j}) .$$

Ons kry dus:

$$\mathrm{kov}(t_{i.}, t_{ij}) = b^2k(m-1)(N-1)^{-1}\sigma_a^2 - b^2k(m-1)(N-1)^{-1}\sigma_a^2$$

$$= 0.$$

$t_{i.}$ en t_{ij} is dus ongekorreleerd.

Op analoë wyse as in die vorige geval, is $\mathrm{kov}(t_{.j}, t_{ij}) = 0$, d.w.s. $t_{.j}$ en t_{ij} is ongekorreleerd.

Hiermee is die lemma bewys.

Lemma 5.3. Die $(mk+m+k)$ variante t''_{ij}, $t'_{i.}$ en $t'_{.j}$ vir [5] $i=1,2,\ldots,m$; $j=1,2,\ldots,k$ is gesamentlik asimptoties normaal verdeel onder die voorwaardes (5.13) en (5.14) indien $N\to\infty$ (d.i. $b\to\infty$).

Bewys: Ons maak gebruik van stelling 2.1, naamlik dat $S_N = \sum_p b_{Np}d_p$ asimptoties normaal verdeel is as $N\to\infty$ indien die voorwaarde

5). Hier is $t''_{ij} = b^{-\frac{1}{2}}t_{ij}$, $t'_{i.} = b^{-\frac{1}{2}}t_{i.}$ en $t'_{.j} = b^{-\frac{1}{2}}t_{.j}$.

$$\text{(2.29)}\quad \lim_{N\to\infty} \frac{N \operatorname*{maks}_p (a_{Np}-\bar{a}_N)^2 \operatorname*{maks}_p (b_{Np}-\bar{b}_N)^2}{\sum_p (a_{Np}-\bar{a}_N)^2 \sum_p (b_{Np}-\bar{b}_N)^2} = 0$$

bevredig word (sien stelling 2.1).

Stel $a_{Np} = a_p$ vir $p=1,2,\ldots,N$ en

$$b^{\frac{1}{2}} b_{Np} = (mk)^{-1}[mkc_{i'j'} - m\sum_j c_{i'j} - k\sum_i c_{ij'} + \sum_i\sum_j c_{ij}] +$$

$$+ km^{-1}(mc_{i'.} - \sum_i c_{i.}) + mk^{-1}(kc_{.j'} - \sum_j c_{.j}) \quad \text{vir alle}$$

p waarvoor z_p in die sel (i',j') lê. Hier word ook veronderstel dat die c_{ij} 's, $c_{i.}$'s en $c_{.j}$'s $(mk+m+k)$ willekeurige eindige konstantes is. (Die $c_{i.}$'s en $c_{.j}$'s is willekeurige konstantes onafhanklik van die c_{ij} 's).

Nou is:

$$\text{(5.25)}\quad b^{\frac{1}{2}} S_N = b^{\frac{1}{2}} \sum_p b_{Np} d_p$$

$$= \sum_{i'}\sum_{j'}\sum_h [(mk)^{-1}(mkc_{i'j'} - m\sum_j c_{i'j} - k\sum_i c_{ij'} + \sum_i\sum_j c_{ij})$$

$$+ km^{-1}(mc_{i'.} - \sum_i c_{i.}) + mk^{-1}(kc_{.j'} - \sum_j c_{.j})]\, a_{i'j'h}$$

$$= \sum_{i'}\sum_{j'}\sum_h c_{i'j'} a_{i'j'h} - k^{-1}\sum_{i'}\sum_{j'}\sum_h\sum_j c_{i'j} a_{i'j'h} - m^{-1}\sum_{i'}\sum_{j'}\sum_h\sum_i c_{ij'} a_{i'j'h} +$$

$$+ (mk)^{-1}\sum_{i'}\sum_{j'}\sum_h\sum_i\sum_j c_{ij} a_{i'j'h} + km^{-1}\sum_{i'}\sum_{j'}\sum_h (mc_{i'.} - \sum_i c_{i.}) a_{i'j'h}$$

$$+ mk^{-1}\sum_{i'}\sum_{j'}\sum_h (kc_{.j'} - \sum_j c_{.j}) a_{i'j'h}$$

$$= \sum_i\sum_j\sum_h c_{ij} a_{ijh} - k^{-1}\sum_i\sum_j\sum_h\sum_{j'} c_{ij} a_{ij'h} - m^{-1}\sum_i\sum_j\sum_h\sum_{i'} c_{ij} a_{i'jh} +$$

$$+ (mk)^{-1}\sum_i\sum_j\sum_h\sum_{i'}\sum_{j'} c_{ij} a_{i'j'h} + k\sum_{i'} c_{i'.} \sum_{j'}\sum_h a_{i'j'h} -$$

$$- km^{-1}\sum_i c_{i.} \sum_{i'}\sum_{j'}\sum_h a_{i'j'h} + m\sum_{i'}\sum_{j'}\sum_h c_{.j'} a_{i'j'h} -$$

$$- mk^{-1}\sum_{i'}\sum_{j'}\sum_h\sum_j c_{.j} a_{i'j'h}$$

$$= \sum_i\sum_j c_{ij}\sum_h a_{ijh} - b\sum_i\sum_j c_{ij} a_{i..} - b\sum_i\sum_j c_{ij} a_{.j.} + b\sum_i\sum_j c_{ij} a_{...} +$$

$$+ k\sum_i c_{i.} \sum_{j'}\sum_h a_{ij'h} - bk^2\sum_i c_{i.} a_{...} + m\sum_j c_{.j}\sum_{i'}\sum_h a_{i'jh} -$$

$$- bm^2\sum_j c_{.j} a_{...}$$

$$= \sum_i \sum_j c_{ij} [\sum_h (a_{ijh} - a_{i..} - a_{.j.} + a_{...})] + k \sum_i c_{i.} \sum_{j'} \sum_h (a_{ij'h} - a_{...})$$
$$+ m \sum_j c_{.j} \sum_{i'} \sum_h (a_{i'jh} - a_{...})$$

$$= \sum_i \sum_j c_{ij} t_{ij} + k \sum_i c_{i.} t_{i.} + m \sum_j c_{.j} t_{.j} \quad \text{sodat}$$

$$S_N = \sum_i \sum_j [c_{ij} t''_{ij} + c_{i.} t'_{i.} + c_{.j} t'_{.j}].$$

Die variante t''_{ij}, $t'_{i.}$ en $t'_{.j}$ vir $i=1,2,\ldots,m$; $j=1,2,\ldots,k$ is dus gesamentlik asimptoties normaal verdeel onder voorwaarde (2.29) as $N \to \infty$.[6] Ons toon vervolgens aan dat (2.29) bevredig word.

Sonder verlies aan algemeenheid kan veronderstel word dat die c_{ij} 's nie almal gelyk is nie, en netso vir die $c_{i.}$'s en die $c_{.j}$'s.

Omdat bogenoemde konstantes almal eindig is, is

$$b \cdot \max_p (b_{Np} - \bar{b}_N)^2 \leq K < \infty.$$

Verder is $\lim_{N \to \infty} b N^{-1} \sum_p (b_{Np} - \bar{b}_N)^2 \geq \varepsilon > 0$ omdat die getalle $b^{\frac{1}{2}} b_{Np}$ vir die verskillende selle nie almal gelyk kan wees as alle c's groepsgewyse nie almal gelyk is nie.

Onder voorwaarde (5.14) kry ons dan as voldoende voorwaarde vir die geldigheid van (2.29) die voorwaarde

(5.13) $\max_p (a_p - a_{...})^2 = o(N)$.

Lemma 5.4. Onder voorwaardes (5.13) en (5.14) besit elk van die drie groepe variante t''_{ij}, $t'_{i.}$ en $t'_{.j}$ ($i=1,2,..,m$; $j=1,2,\ldots,k$) asimptoties, vir $N \to \infty$, gesamentlike normaalverdelings wat onderling onafhanklik is.

Bewys: Stel $\vec{M} = \{m_1, m_2, \ldots, m_n\}$ en $\vec{t} = \{t_1, t_2, \ldots, t_n\}$ met $n = mk+m+k$, $m_g = E(t_g)$, $g=1,2,\ldots,n$, en

(5.26) $$t_g = \begin{cases} t''_{ij}, \ 1 \leq i \leq m, \ 1 \leq j \leq k & \text{vir } 1 \leq g \leq mk \\ t'_{i.}, \ 1 \leq i \leq m & \text{vir } mk+1 \leq g \leq mk+m \\ t'_{.j}, \ 1 \leq j \leq k & \text{vir } mk+m+1 \leq g \leq n \end{cases}$$

sodanig dat $t_1 = t''_{11}$, $t_2 = t''_{12}$,.. $t_k = t''_{1k}$, $t_{k+1} = t''_{21}$, $t_{k+2} = t''_{22}$, ens.

6). Die bewering geld eintlik alleen as aan 'n ekstra voorwaarde voldoen word. Sien bylaag C.

Omdat die variante t''_{ij}, $t'_{i.}$ en $t'_{.j}$ vir alle i en j gesamentlik asimptoties 'n normaalverdeling besit (lemma 5.3), word die karakteristieke funksie van hierdie gesamentlike verdeling gegee deur (sien CRAMéR(1946) p. 310):

$$(5.27)\quad \phi(\vec{u}) = \text{eksp}\left(i\sum_{g=1}^{n} m_g u_g - \tfrac{1}{2}\sum_{f=1}^{n}\sum_{g=1}^{n}\lambda_{fg}u_f u_g\right)$$

waar λ_{fg} die kovariansie tussen t_f en t_g aandui en $\vec{u} = \{u_1, u_2, \ldots, u_n\}$.

Maar

$$\sum_{f,g=1}^{n}\sum\lambda_{fg}u_f u_g = \sum_{f=1}^{n}\sum_{g=1}^{mk}\lambda_{fg}u_f u_g + \sum_{f=1}^{n}\sum_{g=mk+1}^{mk+m}\lambda_{fg}u_f u_g + $$

$$+ \sum_{f=1}^{n}\sum_{g=mk+m+1}^{n}\lambda_{fg}u_f u_g$$

$$= \sum_{f,g=1}^{mk}\sum\lambda_{fg}u_f u_g + \sum_{f,g=mk+1}^{mk+m}\sum\lambda_{fg}u_f u_g + \sum_{f,g=mk+m+1}^{n}\sum\lambda_{fg}u_f u_g$$

omdat die variante t''_{ij}, $t'_{i.}$ en $t'_{.j}$ onderling ongekorreleerd is (lemma 5.2) en dus

$\lambda_{fg} = 0$ is vir:

i) $1 \leq f \leq mk$ en $mk+1 \leq g \leq n$

ii) $mk+1 \leq f \leq mk+m$ en $1 \leq g \leq mk$ of $mk+m+1 \leq g \leq n$

iii) $mk+m+1 \leq f \leq n$ en $1 \leq g \leq mk+m$.

Ons kry dus

$$(5.28)\quad \phi(\vec{u}) = \text{eksp}\left(i\sum_{g=1}^{n} m_g u_g - \tfrac{1}{2}\sum_{f,g=1}^{mk}\sum\lambda_{fg}u_f u_g - \tfrac{1}{2}\sum_{f,g=mk+1}^{mk+m}\sum\lambda_{fg}u_f u_g\right.$$

$$\left. - \tfrac{1}{2}\sum_{f,g=mk+m+1}^{n}\sum\lambda_{fg}u_f u_g\right)$$

$$= \text{eksp}\left(i\sum_{g=1}^{mk} m_g u_g - \tfrac{1}{2}\sum_{f,g=1}^{mk}\sum\lambda_{fg}u_f u_g\right)\cdot$$

$$\cdot\ \text{eksp}\left(i\sum_{g=mk+1}^{mk+m} m_g u_g - \tfrac{1}{2}\sum_{f,g=mk+1}^{mk+m}\sum\lambda_{fg}u_f u_g\right)\cdot$$

$$\cdot\ \text{eksp}\left(i\sum_{g=mk+m+1}^{n} m_g u_g - \tfrac{1}{2}\sum_{f,g=mk+m+1}^{n}\sum\lambda_{fg}u_f u_g\right)$$

$= \phi_1(\vec{u}_1)\phi_2(\vec{u}_2)\phi_3(\vec{u}_3)$, die produk van drie

karakteristieke funksies wat elkeen die karakteristieke funksie is van 'n normaal-verdeelde groep variante. (Hierby is $\vec{u}_1 = \{u_1, u_2, \ldots, u_{mk}\}$, $\vec{u}_2 = \{u_{mk+1}, \ldots, u_{mk+m}\}$ en $\vec{u}_3 = \{u_{mk+m+1}, \ldots, u_n\}$). Verder is
$\phi_1(\vec{u}_1)$ die karakteristieke funksie van die variante t''_{ij},
$\phi_2(\vec{u}_2)$ die karakteristieke funksie van die variante $t'_{i.}$ en
$\phi_3(\vec{u}_3)$ die karakteristieke funksie van die variante $t'_{.j}$.

Die gesamentlike verdeling van die variante t''_{ij} is asimptoties normaal (CRAMéR(1946) p. 310). Soortgelyk vir die gesamentlike verdelings van die variante $t'_{i.}$ en $t'_{.j}$. Hierdie drie normaalverdelings is verder ook onderling onafhanklik omdat die karakteristieke funksie van die gesamentlike verdeling geskryf kan word as die produk van die afsonderlike karakteristieke funksies en gevolglik is die drie groepe variante onderling onafhanklik.

Opmerking.

Uit die voorgaande bewys volg dat enige funksie van die variante t''_{ij} asimptoties onafhanklik is van enige funksie van die variante $t'_{i.}$ en van enige funksie van die variante $t'_{.j}$ (vergelyk CRAMéR(1962) p. 16 stelling 3).

Definieer nou:

(5.29) $\hat{t}_{ij} = (N-1)^{\frac{1}{2}} t_{ij} N^{-\frac{1}{2}} \sigma_a^{-1}$

(5.30) $\hat{t}_{i.} = (N-kb)^{\frac{1}{2}} t_{i.} [N \text{ var}(t_{i.})]^{-\frac{1}{2}}$ en

(5.31) $\hat{t}_{.j} = (N-mb)^{\frac{1}{2}} t_{.j} [N \text{ var}(t_{.j})]^{-\frac{1}{2}}$.

Vir $t_{i.} = \sum_j \sum_h (a_{ijh} - a_{...})$ kan soos in lemmas 2.2 en 2.3 bewys word dat

(5.32) $\text{var}(t_{i.}) = kb(N-kb)(N-1)^{-1} \sigma_a^2$ en

(5.33) $\text{kov}(t_{i.}, t_{i'.}) = -k^2 b^2 (N-1)^{-1} \sigma_a^2$ vir $i' \neq i$.

Vir $t_{.j} = \sum_i \sum_h (a_{ijh} - a_{...})$ is

(5.34) $\text{var}(t_{.j}) = mb(N-mb)(N-1)^{-1} \sigma_a^2$ en

(5.35) $\text{kov}(t_{.j}, t_{.j'}) = -m^2 b^2 (N-1)^{-1} \sigma_a^2$ vir $j' \neq j$.

STELLING 5.2.

Onder voorwaardes (5.13) en (5.14) besit T_m onder H_o asimptoties, vir $N\to\infty$, 'n χ^2-verdeling met (m-1) grade van vryheid.

Bewys: Omdat die gesamentlike verdeling van die variante $t'_{i.}$ asimptoties, vir $N\to\infty$, normaal is (lemma 5.4), is die gesamentlike verdeling van die variante $\hat{t}_{i.}$ ook asimptoties normaal.

Met behulp van vergelykings (5.30), (5.32) en (5.33) volg:

$$\mathrm{var}(\hat{t}_{i.}) = 1 - m^{-1} \quad \text{en}$$
$$\mathrm{kov}(\hat{t}_{i.}, \hat{t}_{i'.}) = - m^{-1} .$$

Die variante $\hat{t}_{i.}$ besit dus gesamentlik asimptoties 'n normaalverdeling as $N\to\infty$ met momentematriks

$$M_6 = \begin{Bmatrix} 1 - m^{-1}, & - m^{-1}, & \ldots\ldots, & - m^{-1} \\ - m^{-1}, & 1 - m^{-1}, & \ldots\ldots, & - m^{-1} \\ \cdot & \cdot & \cdot & \cdot \\ - m^{-1}, & - m^{-1}, & \ldots\ldots, & 1 - m^{-1} \end{Bmatrix}$$

$$= I - \vec{m}\vec{m}'$$

waar I die eenheidsmatriks is, $\vec{m}'$ die ryvektor $\{m^{-\frac{1}{2}}, m^{-\frac{1}{2}}, \ldots, m^{-\frac{1}{2}}\}$ en $\vec{m}$ die ooreenkomstige kolomvektor.

Uit lemma 1.2 volg dat M_6 van rang (m-1) is omdat $\sum\limits_i m^{-1} = 1$. Uit stelling 1.1 opmerking 2 (vergelyk ook CRAMÉR(1946) p. 419) volg dan dat $\sum\limits_i \hat{t}^2_{i.}$ asimptoties, vir $N\to\infty$, χ^2 verdeel is met (m-1) g.v.v.

Hierby is

$$\sum_i \hat{t}^2_{i.} = (N-kb)N^{-1} \sum_i t^2_{i.} [\mathrm{var}(t_{i.})]^{-1} \qquad \text{uit (5.30)}$$

$$= (N-1)[Nkb\sigma^2_a]^{-1} \sum_i [\sum_j \sum_h (a_{ijh} - a_{...})]^2 \quad \text{uit (5.21) en (5.32)}$$

$$= (N-1)kbN^{-1} \sigma_a^{-2} \sum_i (a_{i..} - a_{...})^2$$

$$= T_m \quad \text{uit (5.10)} .$$

Hiermee is die stelling bewys.

Opmerking.

(5.36) $E(T_m) = E[(N-1)kbN^{-1}\sigma_a^{-2}\sum_i (a_{i..} - a_{...})^2]$

$$= (N-1)kbN^{-1}\sigma_a^{-2}\sum_i E(a_{i..} - a_{...})^2$$

$$= \frac{(N-1)kb}{N\sigma_a^2}\sum_i E\left[\frac{1}{kb}\sum_j\sum_h(a_{ijh}-a_{...})\frac{1}{kb}\sum_{j'}\sum_{h'}(a_{ij'h'}-a_{...})\right]$$

$$= \frac{(N-1)}{kbN\sigma_a^2}\sum_i\left\{\sum_j E[\sum_h (a_{ijh} - a_{...})\sum_{h'}(a_{ijh'} - a_{...})] + \right.$$

$$\left. + \sum_{j\neq j'}\sum E[\sum_h (a_{ijh} - a_{...})\sum_{h'}(a_{ij'h'} - a_{...})]\right\}$$

$$= \frac{(N-1)}{kbN\sigma_a^2}\sum_i [\sum_j b(N-b)(N-1)^{-1}\sigma_a^2 - \sum_{j\neq j'}\sum b^2(N-1)^{-1}\sigma_a^2] \text{ uit (5.23) en (5.24)}$$

$$= mk^{-1}N^{-1}[k(N-b) - k(k-1)b]$$

$$= (m-1) \quad \text{omdat } N = mkb$$

= aantal g.v.v. van T_m, in ooreenstemming met die teorie van die χ^2-verdeling.

STELLING 5.3.

Onder voorwaardes (5.13) en (5.14) besit T_k onder H_0 asimptoties, vir $N\to\infty$, 'n χ^2-verdeling met (k-1) grade van vryheid.

Bewys: Op analoë wyse as in stelling 5.2 kan met behulp van lemmas 1.2 en 5.4 en stelling 1.1 opmerking 2 bewys word dat $\sum_j \hat{t}^2_{.j}$ asimptoties, vir $N\to\infty$, 'n χ^2-verdeling besit met (k-1) g.v.v.

Hier is

$$\sum_j \hat{t}^2_{.j} = (N-mb)N^{-1}\sum_j t^2_{.j}[\mathrm{var}(t_{.j})]^{-1}$$

$$= (N-1)[Nmb\sigma_a^2]^{-1}\sum_j [\sum_i\sum_h(a_{ijh} - a_{...})]^2 \quad \text{uit (5.22) en (5.34)}$$

$$= (N-1)mbN^{-1}\sigma_a^{-2}\sum_j(a_{.j.} - a_{...})^2$$

$$= T_k \quad \text{uit (5.11).}$$

Hiermee is die stelling bewys.

Opmerking.

Soos in die opmerking na stelling 5.2 kan bewys word dat

(5.37) $E(T_k) = (k-1) =$ aantal g.v.v. van T_k .

STELLING 5.4.

Onder voorwaardes (5.13) en (5.14) besit T_I onder H_o asimptoties 'n χ^2-verdeling met $(m-1)(k-1)$ grade van vryheid as $N \to \infty$.

Bewys: T_s, T_m en T_k besit asimptoties, vir $N \to \infty$, χ^2-verdelings met respektiewelik $(mk-1)$, $(m-1)$ en $(k-1)$ grade van vryheid (stellings 5.1, 5.2 en 5.3).

Verder is $T_s = T_m + T_k + T_I$ (lemma 5.1) waarby T_m, T_k en T_I onderling onafhanklik is (lemma 5.4).

Die aantal grade van vryheid van T_s is dus hoogstens gelyk aan die som van die aantal grade van vryheid van T_m, T_k en T_I (SCHEFFé(1959) p. 422).

$(mk-1) \leq (m-1) + (k-1) + a$ waar a die rang van die momentematriks van die variante t_{ij} voorstel.

Dus $a \geq (m-1)(k-1)$.

Beskou

$$(5.20)\quad t_{ij} = \sum_h (a_{ijh} - a_{i..} - a_{.j.} + a_{...})$$

$$= b(a_{ij.} - a_{i..} - a_{.j.} + a_{...}) \quad .$$

Daar is km variante t_{ij} waartussen daar minstens $(k+m-1)$ onafhanklike lineêre verbande bestaan, want:

(5.38) $\sum_i \sum_j t_{ij} = 0$ lê een lineêre verband tussen die t_{ij}'s vas.

Verder is

$$(5.39)\quad \sum_j t_{ij} = 0 \text{ vir } i=1,2,\ldots,m \; .$$

Hierdeur word $(m-1)$ verdere lineêre verbande tussen die t_{ij}'s vasgelê. (Indien $\sum_j t_{ij} = 0$ vir alle waardes van i behalwe één, dan volg die m^{de} betrekking uit (5.38)).

Soortgelyk lê

(5.40) $\sum_i t_{ij} = 0$ vir $j=1,2,\ldots,k$

nog (k-1) onafhanklike lineêre verbande vas. Daar bestaan dus minstens (m+k-1) onafhanklike lineêre betrekkings tussen die t_{ij}'s, sodat die rang a van die momentematriks van die t_{ij}'s hoogstens gelyk is aan

$mk - (m+k-1) = (m-1)(k-1)$. (Sien CRAMéR(1946) pp. 110 en 312).

Maar $a \geq (m-1)(k-1)$. Dus

(5.41) $a = (m-1)(k-1)$.

Omdat die variante t''_{ij} asimptoties, vir $N \to \infty$, gesamentlik 'n normaalverdeling besit met momentematriks van rang $a = (m-1)(k-1)$, volg uit stellings 5.1, 5.2 en 5.3 en lemma 5.4 dat T_I asimptoties 'n χ^2-verdeling met (m-1)(k-1) g.v.v. besit. (Let op dat

(5.42) $T_I = T_s - T_k - T_m$ uit lemma 5.1, sodat

$$E(T_I) = E(T_s) - E(T_k) - E(T_m)$$
$$= (mk-1) - (k-1) - (m-1)$$
$$= (m-1)(k-1)).$$

Hiermee is die stelling bewys.

5.3.5. Die Betekenis van T_s, T_m, T_k en T_I.

T_s kan beskou word as 'n toets vir „sel-effekte" omdat dit meet in watter mate die selgemiddeldes $a_{ij.}$ van die totaalgemiddeld $a_{...}$ verskil.

Die toetsingsgrootheid T_m kan gebruik word as 'n toets in hoeverre die ry-gemiddeldes $a_{i..}$ van die totaalgemiddeld $a_{...}$ verskil. T_m meet dus die sogenaamde „ry-effekte".

Analoog aan T_m meet T_k die „kolom-effekte".

Elke selgemiddeld $a_{ij.}$ is onderhewig aan i) ry-effekte; ii) kolom-effekte; iii) ry-kolom interaksie.

In die toetsingsgrootheid

$$(5.43)\quad T_I = (N-1)bN^{-1}\,\sigma_a^{-2}\sum_i\sum_j(a_{ij.} - a_{i..} - a_{.j.} + a_{...})^2$$

$$= (N-1)bN^{-1}\sigma_a^{-2}\sum_i\sum_j[(a_{ij.}-a_{...})-(a_{i..} - a_{...})-(a_{.j.} - a_{...})]^2$$

word getrag om ry- en kolomeffekte uit te skakel deur aftrekking van $a_{i..}$ en $a_{.j.}$ sodat T_I beskou kan word as 'n maat van die interaksie tussen rye en kolomme. Skematies kry ons:

TABEL 5.1.

Opsomming van Toetsingsgroothede.

Variasiebron	Toetsingsgrootheid	G.v.V.
Ry-effekte	$T_m = \frac{(N-1)kb}{N\sigma_a^2}\sum_i(a_{i..} - a_{...})^2$	$(m-1)$
Kolomeffekte	$T_k = \frac{(N-1)mb}{N\sigma_a^2}\sum_j(a_{.j.} - a_{...})^2$	$(k-1)$
Interaksie	$T_I = \frac{(N-1)b}{N\sigma_a^2}\sum_i\sum_j(a_{ij.}-a_{i..}-a_{.j.}+a_{...})^2$	$(m-1)(k-1)$
Sel-effekte	$T_s = \frac{(N-1)b}{N\sigma_a^2}\sum_i\sum_j(a_{ij.} - a_{...})^2$	$(mk-1)$

5.3.6. Spesiale gevalle van T_s, T_m, T_k en T_I.

Deur $\Psi_N(\delta)$ gelyk te stel aan spesifieke funksies, kan verskillende spesiale gevalle van die toetsingsgroothede verkry word (sien byvoorbeeld §2.5). Ons bespreek slegs twee gevalle.

Eenvoudigheidshalwe neem ons aan dat in geval van knope daar deur die lotingsmeganisme rangnommers toegeken word, sodat ons die voorkoms van knope verder kan ignoreer.

5.3.6.1. $\Psi_N(\delta_{ijh}) \equiv \delta_{ijh}$ (sien §2.5.2.1).

$$(5.44)\quad T_m = (N-1)kbN^{-1}\,\sigma_a^{-2}\sum_i(a_{i..} - a_{...})^2 \qquad \text{uit } (5.10)$$

$$= \frac{12}{Nkb(N+1)}\sum_i[\sum_j\sum_h r_{ijh} - \tfrac{1}{2}kb(N+1)]^2\ .$$

Netso is

(5.45) $T_k = \dfrac{12}{Nmb(N+1)} \sum_j [\sum_i \sum_h r_{ijh} - \tfrac{1}{2}mb(N+1)]^2$ en

(5.46) $T_I = \dfrac{12}{N(N+1)bk^2m^2} \sum_i \sum_j [mk \sum_h r_{ijh} - m \sum_j \sum_h r_{ijh} - k \sum_i \sum_h r_{ijh} +$
$+ \tfrac{1}{2}mkb(N+1)]^2$.

5.3.6.2. $\Psi_N(\delta_{ijh}) \equiv Q(\delta_{ijh})$ (sien §2.5.2.3).

In hierdie geval is:

(5.47) $T_m = (N-1) \sum_i [\sum_j \sum_h Q(\delta_{ijh})]^2 / [kb \sum_p Q^2(\delta_p)]$,

(5.48) $T_k = (N-1) \sum_j [\sum_i \sum_h Q(\delta_{ijh})]^2 / [mb \sum_p Q^2(\delta_p)]$ en

(5.49) $T_I = (N-1)[bk^2m^2 \sum_p Q^2(\delta_p)]^{-1} \sum_i \sum_j [mk \sum_h Q(\delta_{ijh}) -$
$-m \sum_j \sum_h Q(\delta_{ijh}) - k \sum_i \sum_h Q(\delta_{ijh})]^2$.

Soortgelyk kan daar nog ander toetsingsgroothede gekonstrueer word (sien vorige hoofstukke se spesiale gevalle van $\Psi_N(\delta)$).

5.4. VERGELYKING TUSSEN DIE PARAMETRIESE EN NIE-PARAMETRIESE TOETSE.

By die gewone tweefaktor analise van variansie vir homogeniteit van gemiddeldes met b waarnemings per sel, het ons die volgende skema:

TABEL 5.2.

Variansietabel.

Variasiebron	Som van Vierkante	G.v.V
Ry-effekte	$q_1 = kb \sum_i (x_{i..} - x_{...})^2$	$(m-1)$
Kolom-effekte	$q_2 = mb \sum_j (x_{.j.} - x_{...})^2$	$(k-1)$
Interaksie	$q_3 = b \sum_i \sum_j (x_{ij.} - x_{i..} - x_{.j.} + x_{...})^2$	$(m-1)(k-1)$
Subtotaal	$Q_s = b \sum_i \sum_j (x_{ij.} - x_{...})^2$	$(mk-1)$
Residu	$q_4 = \sum_i \sum_j \sum_h (x_{ijh} - x_{ij.})^2$	$mk(b-1)$
Totaal	$Q = \sum_i \sum_j \sum_h (x_{ijh} - x_{...})^2$	$(mkb-1)$

Hierby stem q_1 ooreen met T_m (sien tabel 5.1) in soverre dat albei ry-effekte meet, q_2 met T_k, q_3 met T_I en Q_s met T_s. Let op dat

$$T_Q \stackrel{\text{def}}{=} (N-1)N^{-1}\,\sigma_a^{-2}\sum_i\sum_j\sum_h(a_{ijh}-a_{...})^2$$

$$= (N-1)\,\sigma_a^{-2}\,.\,\sigma_a^2 \qquad \text{uit (5.8)}$$

$$= (N-1)$$

= 'n bepaalde waarde vir vaste m, k en b wat dus geen waarskynlikheidsverdeling besit nie. Die rede hiervoor is dat $\sum_i\sum_j\sum_h a_{ijh} \equiv$ konstant.

Vir ry-effekte toets ons in die geval van tabel 5.2 met

(5.50) $F_1 = mk(b-1)q_1/[(m-1)q_4]$.

F_1 besit 'n F-verdeling met (m-1) en mk(b-1) g.v.v.

As $b\to\infty$, gaan die verdeling van $(m-1)F_1$ oor in 'n χ^2-verdeling met (m-1) g.v.v. (KENNEY en KEEPING(1951) vol. II p. 183) en kry ons presies dieselfde toestand as in die verdelingsvrye geval hierbo waar ons vir ry-effekte toets met T_m wat asimptoties χ^2 verdeel is met (m-1) g.v.v. as $b\to\infty$.

Soortgelyk toets ons vir kolom-effekte met

(5.51) $F_2 = mk(b-1)q_2/[(k-1)q_4]$

waarby $(k-1)F_2$ asimptoties, vir $b\to\infty$, 'n χ^2-verdeling met (k-1) g.v.v. besit (vergelyk T_k).

Vir interaksie toets ons met

(5.52) $F_3 = mk(b-1)q_3/[(m-1)(k-1)q_4]$

waarby $(m-1)(k-1)F_3$ asimptoties, vir $b\to\infty$, 'n χ^2-verdeling besit met (m-1)(k-1) g.v.v. (vergelyk T_I).

Asimptoties, as $b\to\infty$, toon ons toetse en die gewone variansie-analise toetse 'n groot ooreenkoms. Hierdie ooreenkoms blyk verder duidelik indien ons die verskillende toetsingsgroothede se formules met mekaar vergelyk. Byvoorbeeld vir ry-effekte vergelyk ons $(m-1)F_1$ met T_m.

$$(5.53)\quad (m-1)F_1 = mk(b-1)q_1/q_4 \quad \text{uit } (5.50)$$
$$= mk(b-1)kb\sum_i(x_{i..}-x_{...})^2/\sum_i\sum_j\sum_h(x_{ijh}-x_{ij.})^2$$
$$= (1-\tfrac{1}{b})bk\sum_i(x_{i..}-x_{...})^2/N^{-1}\sum_i\sum_j\sum_h(x_{ijh}-x_{ij.})^2 .$$

Uit (5.10) volg:

$$(5.54)\qquad T_m = (1-\tfrac{1}{N})kb\sum_i(a_{i..}-a_{...})^2/\sigma_a^2 .$$

Stel

$$(5.55)\qquad \sigma_{ij}^2 = b^{-1}\sum_h(x_{ijh}-x_{ij.})^2 \quad \text{en}$$

$$(5.56)\qquad \sigma_N^2 = (mk)^{-1}\sum_i\sum_j\sigma_{ij}^2 .$$

Dan is

$$(5.57)\quad (m-1)F_1 = (1-\tfrac{1}{b})\,kb\sum_i(x_{i..}-x_{...})^2/\sigma_N^2 .$$

Neem ons byvoorbeeld $\Psi_N(\delta) = Q(\delta) \equiv \Phi^{-1}(\delta)$, dan sal σ_a^2 asimptoties gelyk wees aan 1 (STOKER(1955) p. 57).

Die toepassing van die F-toets is gebaseer op die veronderstelling dat die x 'e normaalvariante is. Gevolglik is σ_N^2 asimptoties in waarskynlikheid gelyk aan σ^2, die variansie van die normaalvariante. Dit skyn voor die hand liggend te wees dat T_m en $(m-1)F_1$ asimptoties ekwivalent is onder H_o met F die normaal(0,1) v.f.

'n Belangrike vraag is: hoe groot moet b wees alvorens die T-toetse (d.w.s. T_s, T_m, T_k en T_I) toegepas kan word. Hierdie vraag is moeilik te beantwoord. Vir die geval van 2 rye, 2 kolomme en 10 waarnemings per sel met $\Psi_N(\delta)=Q(\delta)$ is vir 'n aantal gevalle die T- en F-toetse se oorskrydingswaarskynlikhede bereken waarby die waarnemings uit WOLD, H(1954): „Tracts for Computers: Random normal deviates" getrek is. Die toepassing van die F-toetse is alleen geregverdig indien ons te doen het met normaal-variante met gelyke variansies.

Die berekende oorskrydingswaarskynlikhede van die T- en F-toetse word in tabelvorm saamgevat waarby die frekwensies vir die toetse vir ry-effekte deur {·}, vir kolomeffekte deur (·) en vir interaksie deur [·] in elke sel aangedui word.

TABEL 5.3.

Oorskrydingswaarskynlikhede van die T- en F-toetse.

		Oorskrydingswaarskynlikhede						
	T / F	0 - 0.025	0.025- 0.05	0.05- 0.10	0.10- 0.20	0.20- 0.40	0.40- 0.60	0.60- 0.80
Oorskrydingswaarskynlikhede	0 - 0.025	{2},[1]						
	0.025- 0.05		{2},[2]		[1]			
	0.05 - 0.10			{1},[2] (1)	(1)			
	0.10 - 0.20				{2},[1] (1)	{1},[1] (1)		
	0.20 - 0.40				(1)	{3},[2] (3)		
	0.40 - 0.60						{1},[2] (2)	{2} (1)
	0.60 - 0.80						(1)	{3} (1)

Dit lyk asof die T- en F-toetse in die geval onder beskouing min of meer ooreenstemmende oorskrydingswaarskynlikhede lewer, maar geen besliste uitspraak kan gelewer word alvorens baie meer berekeninge gedoen is nie.

5.5. ASIMPTOTIESE RELATIEWE DOELTREFFENDHEID.

5.5.1. Inleiding.

Vervolgens wil ons die asimptotiese relatiewe doeltreffendheid bereken van die T-toetse met betrekking tot die F-toets vir alternatiewe van verskuiwing en met betrekking tot Bartlett se toets vir alternatiewe van verspreiding. Vir eers word aangeneem dat daar k_{ij} waarnemings in die sel (i,j) is (ter wille van aansluiting by die notasie van hoofstuk III).

5.5.2. Toetse vir Ry-effekte.

Beskou alle waarnemings binne 'n bepaalde ry as een steekproef, of k steekproewe uit dieselfde populasie. Dan is daar m sulke steekproewe van grootte $k_{i.}$, $i=1,2,\ldots,m$. Dui die verdelingsfunksie van die i^{de} ry aan deur $F_{i.}$, $i=1,2,\ldots,m$. Ons wil nou die nulhipotese H_o: $F_{1.} \equiv F_{2.} \equiv \ldots \equiv F_{m.} \equiv F$ toets onder die aanname dat daar geen kolom- of interaksie-effekte is nie. H_o word getoets teen 'n alternatiewe hipotese H, naamlik dat nie alle $F_{i.}$ identies is nie.

Stel nou

(5.58) $\nu_{Ni.} = k_{i.}/N$,

(5.59) $\nu_{i.} = \lim\limits_{N\to\infty} \nu_{Ni.}$,

(5.60) $F_{i.k_{i.}}(x) = (\text{aantal } x_{ijh} \leq x)/k_{i.}$ vir vaste i en vir $j=1,2,\ldots,k$; $h=1,2,\ldots,k_{ij}$,

(5.61) $H_N(x) = \sum\limits_i \nu_{Ni.} F_{i.k_{i.}}(x)$,

(5.62) $H(x) = \sum\limits_i \nu_{Ni.} F_{i.}(x)$ en

(5.63) $t_{i.} = \int\limits_{-\infty}^{\infty} J_N[H_N(x)]dF_{i.k_{i.}}(x)$ waarby J_N 'n willekeurige funksie is wat voldoen aan soortgelyke voorwaardes as in hoofstuk III (voorwaardes (3.10)-(3.14) en (3.93)).

Soos in lemma 3.14 volg dat

$$(5.64)\quad \lim_{N\to\infty} E(t_{i.}|H) = \int_I J(H)dF_{i.} \ .$$

Dui ons die variansie van $t_{i.}$ onder H aan deur var$(t_{i.}|H)$, kan daar op analoë wyse as in lemma 3.15 'n uitdrukking vir var$(t_{i.}|H)$ afgelei word.

Laat

$$(5.65)\quad \hat{t}_{i.} = (N-k_{i.})^{\frac{1}{2}}[t_{i.} - E(t_{i.}|H)][N\ \mathrm{var}(t_{i.}|H)]^{-\frac{1}{2}}.$$

Onder H_o reduseer $\hat{t}_{i.}$ na die vorm aangedui in (5.30).

Nou definieer ons as toetsingsgrootheid

$$(5.66)\quad T_m = \sum_{i=1}^{m} \hat{t}_{i.}^2 \ .$$

In notasie is daar nou aangesluit by dié van hoofstuk III. Alle resultate van hoofstuk III kan aangepas word vir hierdie geval. Soos in stelling 3.6 volg onder analoë voorwaardes dat T_m onder H asimptoties, vir $N\to\infty$, 'n χ^2-verdeling besit met (m-1) grade van vryheid.

Op soortgelyke wyse as in §§e 3.6.2 en 3.6.3 kan toetse vir verskuiwing en verskil in verspreiding gedefinieer word met a.r.d.-eienskappe wat ooreenstem met dié van hoofstuk III.

Analoog aan die metode van hierdie paragraaf, kan toetse vir kolomeffekte gekonstrueer word onder die aanname dat daar geen ry-effekte en interaksie voorkom nie.

5.5.3. Toetse vir Sel-effekte.

Beskou die waarnemings binne elke sel as één steekproef uit 'n verdeling met verdelingsfunksie F_{ij}. Dan is daar mk steekproewe van grootte k_{ij}, i=1,2,..,m; j=1,2,...,k. Die nulhipotese H_o: $F_{11} \equiv F_{12} \equiv \ldots \equiv F_{mk} \equiv F$ moet getoets word onder die aanname dat daar geen ry-of kolomeffekte of ry-kolom interaksie bestaan nie.

Stel

(5.67) $\nu_{Nij} = k_{ij}/N$,

(5.68) $\nu_{ij} = \lim_{N\to\infty} \nu_{Nij}$,

(5.69) $F_{ijk_{ij}}(x) = (\text{aantal } x_{ijh} \leq x)/k_{ij}$ vir vaste i en j en vir $h=1,2,\ldots,k_{ij}$,

(5.70) $H_N(x) = \sum_i \sum_j \nu_{Nij} F_{ijk_{ij}}(x)$,

(5.71) $H(x) = \sum_i \sum_j \nu_{Nij} F_{ij}(x)$ en

(5.72) $t'_{ij} = \int_{-\infty}^{\infty} J_N[H_N(x)]dF_{ijk_{ij}}(x)$ waar J_N voldoen aan soortgelyke voorwaardes as in hoofstuk III.

Dui die verwagtingswaarde en variansie van t'_{ij} onder H aan deur $E(t'_{ij}|H)$ en $\text{var}(t'_{ij}|H)$.

Laat

(5.73) $\hat{t}'_{ij} = (N-k_{ij})^{\frac{1}{2}}[t'_{ij} - E(t'_{ij}|H)][N \text{ var}(t'_{ij}|H)]^{-\frac{1}{2}}$.

Dan definieer ons as toetsingsgrootheid

(5.74) $T_s = \sum_i \sum_j \hat{t}'^2_{ij}$.

Soos in stelling 3.6 kan aangetoon word dat T_s asimptoties, vir $N\to\infty$, 'n χ^2-verdeling besit met (mk-1) grade van vryheid en kan soos voorheen a.r.d.-uitdrukkings afgelei word vir alternatiewe van verskuiwing en verskil in verspreiding.

5.6. SLOTOPMERKINGS.

1). Die toetse gebaseer op T_m, T_k en T_s in die geval $\Psi_N(\delta) = Q(\delta)$ byvoorbeeld vir verskuiwing, is (onder die genoemde aannames, waaronder dat geen ander effek as dié een waarvoor getoets word, teenwoordig is nie) asimptoties ten minste net so goed soos die F-toets ($E_{T_m,F} \geq 1$).

2). Met behulp van tabelle in VAN DER WAERDEN(1957) is die toepassing van bogenoemde T-toetse (opmerking 1 hierbo) dikwels makliker as dié van die F-toetse.

3). Die vraag oor hoe groot b moet wees alvorens die T-toetse toegepas kan word, is nog nie bevredigend beantwoord nie.

4). Die volgende probleme is onder andere nog onopgelos:

a). Die uitbreiding van die voorgaande teorie na die geval waar die aantal waardes in die onderskeie selle verskillend is.

b). A.r.d.-uitdrukkings vir interaksie-toetse.

c). 'n Volledige verdelingsvrye ekwivalent van die parametriese variansie-analise, naamlik waar byvoorbeeld die toetse vir ry- en kolomeffekte (asimptoties) onafhanklik bly wanneer daar ry- en/of kolomeffekte of interaksie bestaan (m.a.w. H_0 nie geld nie).

5). 'n Ander interessante benadering tot die probleem van 'n verdelingsvrye variansie-analise word gevind in die volgende twee artikels:
HODGES, J.L. en LEHMANN, E. L.(1962) „Rank methods for combination of independent experiments in the analysis of variance", Ann. Math. Stat. 33, pp. 482-497, en LEHMANN, E. L.(1963) „Asymptotically nonparametric inference: An laternative approach to linear models", Ann. Math. Stat. 34, pp. 1494-1506.