Rapport ZW 1950 - 015

Voordracht door J.Hemelrijk in de serie
Actualiteiten op 28 October 1950.

Een symmetrietoets.

1. Inleiding. Wij geven een parametervrije toets voor de hypothese $\mathcal{H}_0'$, dat n waarden $z_1, \ldots, z_n$ waarnemingen zijn van n onafhankelijk verdeelde stochastische grootheden $\underline{z}_1, \ldots, \underline{z}_n$ 1), die alle dezelfde continue waarschijnlijkheidsverdeling bezitten, die symmetrisch ten opzichte van 0 is. Over de vorm van deze waarschijnlijkheidsverdeling maken wij verder geen onderstelling.

Deze toets bezit vele toepassingsmogelijkheden op diverse toepassingsgebieden; voorbeelden zijn de vergelijking van twee meetmethoden, die beide op een aantal verschillende objecten worden toegepast; toetsing van de werkzaamheid van een geneesmiddel door meting van een quantitatieve grootheid vóór en na de toediening daarvan bij een aantal patiënten, lijdende aan dezelfde kwaal, enz.
Deze toepassingen berusten op de gemakkelijk te bewijzen stelling, dat de grootheid

$$(1) \qquad \underline{z} = \underline{a} - \underline{b}$$

symmetrisch t.o.v. 0 verdeeld is, indien $\underline{a}$ en $\underline{b}$ onafhankelijk verdeeld zijn volgens dezelfde waarschijnlijkheidsverdeling.

Ter verduidelijking van de methode geven wij een getallenvoorbeeld waarvan wij de operaties stap voor stap demonstreren.

Voorbeeld: wij onderstellen, dat de volgende 20 waarden z_i bij een experiment gevonden zijn (in volgorde van afdalende grootte gerangschikt):
74; 63; 36; 35; 34; 29; 25; 11; -13; -25; -32; -38; -40; -46; -48; -60; -70; -79; -80; -87.

2. Statistische grootheden en hun verdeling.

Als eerste statistische grootheid voeren wij in het aantal $\underline{n}_1$ der $\underline{z}_i$, die positieve waarden aannemen (in het voorbeeld: $n_1 = 8$). Zoals bekend, bezit $\underline{n}_1$, onder hypothese $\mathcal{H}_0$, een binomiale verdeling van de vorm:

$$(2) \qquad P[\underline{n}_1 = n_1 / \mathcal{H}_0] = 2^{-n}\binom{n}{n_1}$$

---

1) Het stochastische karakter van een stochastische grootheid geven wij aan door onderstreping. Niet onderstreepte letters geven getallen of niet-stochastische variabelen aan.

waarin het achter de verticale streep geplaatste symbool $\mathcal{H}_0$ de hypothese aangeeft, waaronder (2) geldt.

Vervolgens geven wij de $\underline{n}_1$ positieven onder de $\underline{z}_i$, in volgorde van waarneming aan met $\underline{x}_1, \ldots, \underline{x}_{n_1}$, en de $\underline{n}_2 = n - \underline{n}_1$ negatieve, eveneens in volgorde van waarneming, met $-\underline{y}_1, \ldots, -\underline{y}_{n_2}$, zodat $\underline{x}_1, \ldots, \underline{y}_{n_2}$ stochastische grootheden zijn, die, onder de voorwaarde $\underline{n}_1 = n_1$, en hypothese $\mathcal{H}_0$, onderling onafhankelijk verdeeld zijn volgens dezelfde voorwaardelijke waarschijnlijkheidsverdeling en die slechts positieve waarden kunnen aannemen.

Deze grootheden rangschikken wij volgens afdalende volgorde van grootte en wij geven ze in deze volgorde aan met

(3) $\underline{w}_1, \ldots\ldots, \underline{w}_n$ 2)

Zij nu

(4) $r = \left[\frac{n+1}{2}\right]$

dan definiëren wij $\underline{u}$ resp. $\underline{v}$ als het aantal der $\underline{x}_i$, die onder $\underline{w}_1, \ldots\ldots \underline{w}_r$ resp. $\underline{w}_{r+1}, \ldots\ldots, \underline{w}_n$ voorkomen (dus $\underline{u} + \underline{v} = \underline{n}_1$).

In ons voorbeeld is dus $r = 10$, $u = 2$, $v = 8$,

Nu geldt:

LEMMA 1: De voorwaardelijke waarschijnlijkheidsverdeling van $\underline{u}$, onder de voorwaarde $\underline{n}_1 = n_1$ en onder hypothese $\mathcal{H}_0$, is een hypergeometrische verdeling, die gegeven wordt door

$$(5) \qquad P[\underline{u} = u \mid n_1 ; \mathcal{H}_0] = \frac{\binom{r}{u}\binom{n-r}{v}}{\binom{n}{n_1}}$$

waarin $v = n_1 - u$ is en r door (4) gegeven wordt.

BEWIJS: Zoals boven reeds is opgemerkt, zijn onder de voorwaarde $\underline{n}_1 = n_1$ en onder hypothese $\mathcal{H}_0$, de grootheden $\underline{x}_1, \ldots, \underline{y}_{n_2}$ onderling onafhankelijk verdeeld met dezelfde voorwaardelijke verdeling.

De rij grootheden $\underline{w}_1, \ldots, \underline{w}_n$ (zie (3)) is een stochastische permutatie van de rij $\underline{x}_1, \ldots, \underline{y}_{n_2}$, waarbij nu dus alle $n!$ permutaties gelijke waarschijnlijkheden ($\frac{1}{n!}$) bezitten. Onder deze permutaties zijn er echter

$$(6) \qquad \nu(u) = n_1!\, n_2!\binom{r}{u}\binom{n-r}{v}$$

---

2) Wegens de onderstelling van continuïteit is de kans op het optreden van gelijken onder deze grootheden gelijk aan 0.

waarbij $\underline{u}$ de waarde u aanneemt. Hieruit volgt het lemma.

Uit (2) en lemma 1 volgt echter direct:

STELLING 1: Onder hypothese $\mathcal{H}_0$ zijn u en v onafhankelijk verdeeld volgens binomiale verdelingen.

(7) $$P[\underline{u} = u, \underline{v} = v \,/\, \mathcal{H}_0] = 2^{-n}\binom{r}{u}\binom{n-r}{v}$$

Op deze stelling baseren wij de toets voor $\mathcal{H}_0$. Wij geven verder het paar van waarden (u, v) aan als een punt $\mathcal{U} \equiv (u, v)$ in een vlak met coördinaten u en v. Het linker-lid van (7) schrijven wij ter afkorting als $P[\mathcal{U}/\mathcal{H}_0]$.

3. Kritieke zône.

De kritieke zône $Z(\alpha)$, behorende bij gegeven onbetrouwbaarheidsdrempel α, definiëren wij nu als die verzameling van punten $\mathcal{U} \equiv (u, v)$, die aan de volgende eisen voldoen:

(8) $$\sum_{\mathcal{U} \in Z} P[\mathcal{U}/\mathcal{H}_0] \leq \alpha$$

(9) Is $\mathcal{U}_1 \in Z$ en $\mathcal{U}_2 \in Z$, dan is $P[\mathcal{U}_1/\mathcal{H}_0] < P[\mathcal{U}_2/\mathcal{H}_0]$

(10) Z is de grootste verzameling die aan (8) en (9) voldoet.

In tabel I is de begrenzing van deze kritieke zône voor $\alpha = 0{,}05$ resp. $0{,}10$ aangegeven door een getrokken resp. een gestippelde lijn. Voor de duidelijkheid zijn alle mogelijk uitkomsten $\mathcal{U}$ in de tabel opgenomen, hoewel men, wegens de symmetrie, kan volstaan met ongeveer 1/4 daarvan.

De toets $\mathcal{H}_0$ bestaat daaruit, dat $\mathcal{H}_0$ verworpen wordt met onbetrouwbaarheidsdrempel α, als de gevonden uitkomst $\mathcal{U}$ tot $Z(\alpha)$ behoort. Uit stelling 1 volgt, dat dan inderdaad de kans op (ten onrechte) verwerpen van $\mathcal{H}_0$, indien deze juist is, $\leq \alpha$ is.

Tabel I

Simultane waarschijnlijkheidsverdeling van u en v, onder hypothese $\mathcal{H}_0$.

v \ u	0	1	2	3	4	5	6	7	8	9	10
10	1	10	45	120	210	252	210	120	45	10	1
9	10	100	450	1200	2100	2520	2100	1200	450	100	10
8	45	450	2025	5400	9450	11340	9450	5400	2025	450	45
7	120	1200	5400	14400	25200	30240	25200	14400	5400	1200	120
6	210	2100	9450	25200	44100	52920	44100	25200	9450	2100	210
5	252	2520	11340	30240	52920	63504	52920	30240	11340	2520	252
4	210	2100	9450	25200	44100	52920	44100	25200	9450	2100	210
3	120	1200	5400	14400	25200	30240	25200	14400	5400	1200	120
2	45	450	2025	5400	9450	11340	9450	5400	2025	450	45
1	10	100	450	1200	2100	2520	2100	1200	450	100	10
0	1	10	45	120	210	252	210	120	45	10	1

(Alle waarschijnlijkheden in de tabel zijn met $2^{20} = 1049.10^3$ vermenigvuldigd.)

4. Bruikbaarheid van de toets.

Definitie: Een toets heet bruikbaar, of asymptotisch onderscheidend[3]) met betrekking tot de klasse B van hypothesen omtrent de simultane waarschijnlijkheidsverdeling der z_i indien voor iedere $\alpha > 0$ in iedere hypothese $\mathcal{H}$ van B geldt:

(11) $$\lim_{n \to \infty} P[\underline{u} \in Z(\alpha) / \mathcal{H}] = 1$$

waarin n het aantal grootheden z_i is.

Voor de formulering van de bruikbaarheidsstelling, die wij zullen bewijzen, zijn nog enige voorafgaande definities nodig:

Zij z een stochastische grootheid, die positieve en negatieve waarden aan kan nemen, dan definiëren wij x_z resp. $-y_z$ als stochastische grootheden, die als waarschijnlijkheidsverdeling de voorwaardelijke waarschijnlijkheidsverdelingen van z bezitten onder de voorwaarden $z > 0$ resp. $z < 0$. De verdelingsfuncties van x_z en y_z met F en G aangevende, definiëren wij M_x en M_y als die verzamelingen van waarden x resp. y waarvoor geldt:

---

3) Engels: "consistent"; de vertaling van deze term is overgenomen uit D. van Dantzig, kadercursus Mathematische Statistiek, Math. Centrum, Amsterdam, 1947 - 1950.

(12) $F(x) = P[\underline{x}_z \leqq x] \geqq \frac{1}{2}$ <u>en</u> $\lim_{\varepsilon \searrow 0} F(x-\varepsilon) \leqq \frac{1}{2}$

resp.

(13) $G(y) = P[\underline{y}_z \leqq y] \geqq \frac{1}{2}$ <u>en</u> $\lim_{\varepsilon \searrow 0} G(y-\varepsilon) \leqq \frac{1}{2}$

M_x en M_y zijn dan, wegens de monotonie van F en G (eventueel ontaarde) afgesloten intervallen, die wij de <u>voorwaardelijke mediaanintervallen</u> van $\underline{z}$, onder de voorwaarden $\underline{z} > 0$ resp. $\underline{z} < 0$, noemen.

Wij definiëren nu m_x en m_y in geval $M_x \cap M_y$ leeg is, als die punten $m_x \in M_x$ en $m_y \in M_y$, die de kleinste afstand bezitten, en in geval $M_x \cap M_y$ niet leeg is, als het midden $m = m_x = m_y$ van deze doorsnede.

Wij bewijzen vervolgens

<u>STELLING 2: De beschreven toets voor H_0 is asymptotisch onderscheidend voor $n \to \infty$ met betrekking tot de klasse B van hypothesen, inhoudende dat</u>

<u>1e. alle $\underline{z}_i$ ($i = 1, \ldots\ldots, n$) onderling onafhankelijk verdeeld zijn volgens dezelfde continue waarschijnlijkheidsverdeling als een stochastische grootheid $\underline{z}$</u> ;

<u>2e. voor deze verdeling geldt:</u>

(15) $P[\underline{z}_i > 0] \neq \frac{1}{2}$

<u>of</u>

(16) $m_x \neq m_y$

<u>of beide</u>

<u>BEWIJS:</u> Zij

(17) $P[\underline{z} > 0] = \theta$

dan bezit $\underline{n}_1$ een binomiale verdeling met

(18) $\mathcal{E}\underline{n}_1 = \theta n \qquad \sigma^2_{\underline{n}_1} = n\theta(1-\theta)$

Wij beschouwen nu de <u>voorwaardelijke verdeling van $\underline{u}$</u>, onder de voorwaarde $\underline{n}_1 = n_1$ en <u>in de onderstelling dat één der hypothesen H van B juist is.</u> Geven wij de verdelingsfuncties van $\underline{x}_z$ en $\underline{y}_z$ weer aan met F resp. G dan volgt uit de continuïteit der verdeling van $\underline{z}$, dat

(19) $F(m_x) = G(m_y) = \frac{1}{2}$

is. Verder zij

(20) $\frac{n_1}{n_2} = c$

en w zij het midden van het interval der waarden w', die voldoen aan

(21) $\quad cF(w') + G(w') = \frac{1}{2}(1+c)$

Het bestaan van w volgt daaruit, dat $cF+G$ monotoon met w' stijgt van 0 tot $1+c$.

Indien $m_x \neq m_y$ is, ligt w in het open interval, dat door m_x en m_y wordt begrensd. Dit ziet men b.v. als volgt: is $m_x < m_y$ (voor $m_x > m_y$ verloopt het bewijs analoog), dan is

$$\delta = P[\underline{x}_z > m_y] < P[\underline{x}_z > m_x] = \tfrac{1}{2}$$

Was nu $w \geqq m_y$, dan was het linker lid van (21)

$$\geqq c(1-\delta) + \tfrac{1}{2} > \tfrac{1}{2}(1+c)$$

Derhalve is $w < m_y$. Op analoge wijze bewijst men dat $w > m_x$ is.

Zij nu

(22) $\quad \eta = P[w < \underline{x}_z < m_x]$ en $\varepsilon = P[m_y < \underline{y}_z < w]$

als $m_x > m_y$ is;

(23) $\quad -\eta = P[m_x < \underline{x}_z < w]$ en $-\varepsilon = P[w < \underline{y} < m_y]$

als $m_x < m_y$ is, en

(24) $\eta = \varepsilon = 0$, indien $m_x = m_y$ is, dan volgt door substitutie van (22) en (23) in (21), dat steeds geldt:

(25) $\quad \varepsilon = c\eta$

Noemen wij $\underline{k}_x$ resp. $\underline{k}_y$ de aantallen der $\underline{x}_i$ resp. $\underline{y}_i$, die groter dan w zijn, dan zijn $\underline{k}_x$ en $\underline{k}_y$, onder de voorwaarde $n_1 = n_1$, onderling onafhankelijk verdeeld volgens binomiale verdelingen met

(26) $\quad \mathcal{E}\underline{k}_x = (\frac{1}{2}+\eta)n_1$ en $\sigma^2(\underline{k}_x) = n_1(\frac{1}{4}-\eta^2)$

en

(27) $\quad \mathcal{E}\underline{k}_y = (\frac{1}{2}-c\eta)n_2$ en $\sigma^2(\underline{k}_y) = n_2(\frac{1}{4}-c^2\eta^2)$

Zij verder $\underline{w}_r$ de r^e stochastische grootheid uit de rij der $\underline{w}_i$ (zie (3)) en $\underline{m}$ het aantal der $\underline{x}_i$ en $\underline{y}_i$, die tussen w en $\underline{w}_r$ liggen ($\underline{w}_r$ inbegrepen), terwijl $\underline{m}'$ het aantal der $\underline{x}_i$ hieronder is, waarbij $\underline{m}$ en $\underline{m}'$ positief worden genomen, indien $w > \underline{w}_r$ is en negatief, indien $w < \underline{w}_r$ is, dan is

(28) $\quad \underline{m} = r - \underline{k}_x - \underline{k}_y$

en

(29) $\quad |\underline{m}| - |\underline{m}'| \geqq 0$

Volgens (26) en (27) is

$$(30)\quad \begin{cases} \mathcal{E}\underline{m} = r - \mathcal{E}\underline{k}_x - \mathcal{E}\underline{k}_y = r - \tfrac{1}{2}n = 0 \text{ of } \tfrac{1}{2} \\ \sigma^2_{\underline{m}} = \sigma^2(\underline{k}_x) + \sigma^2(\underline{k}_y) \end{cases}$$

en volgens (29) voor n is even of groot en oneven

$$(31)\quad \begin{cases} |\mathcal{E}\underline{m}'| \leqq |\mathcal{E}\underline{m}| = r - \tfrac{1}{2}n \\ \sigma^2_{\underline{m}'} \leqq \sigma^2_{\underline{m}} \end{cases}$$

Nu is

$$(32)\quad \underline{u} = \underline{k}_x + \underline{m}'$$

dus, onder de voorwaarde $\underline{n}_1 = n_1$, aangegeven door het symbool n_1 achter de verticale streep:

$$(33)\quad \begin{cases} \mathcal{E}(\underline{u}/n_1) = (\tfrac{1}{2} + \eta)n_1 + \mathcal{E}\underline{m}' \\ \sigma^2(\underline{u}/n_1) = \{\sigma(\underline{k}_x) + \sigma_{\underline{m}'}\}^2 \end{cases}$$

Daar $\sigma^2(\underline{k}_x) \leqq \tfrac{1}{4}n_1$; $\sigma^2(\underline{k}_y) \leqq \tfrac{1}{4}n_2$, dus $\sigma^2_{\underline{m}'} \leqq \tfrac{1}{4}n$ is, hebben wij:

$$(34)\quad \sigma^2(\underline{u}/n_1) \leqq \left(\tfrac{1}{2}\sqrt{n_1} + \tfrac{1}{2}\sqrt{n}\right)^2 = \tfrac{1}{4}\left(n + n_1 + \sqrt{n n_1}\right) = d^2 = O(n)$$

Hieruit verkrijgen wij een schatting voor de <u>onvoorwaardelijke</u> spreiding van $\underline{u}$ door toepassing van de volgende formule:

$$(35)\quad \sigma^2_{\underline{u}} = \sigma^2\{\mathcal{E}(\underline{u}/\underline{n}_1)\} + \mathcal{E}\{\sigma^2(\underline{u}/\underline{n}_1)\}$$ [4)]

Wij hebben n.l. volgens (18)

$$(36)\quad \sigma^2\{\mathcal{E}(\underline{u}/\underline{n}_1)\} = \sigma^2\{(\tfrac{1}{2} + \eta)\underline{n}_1 + \mathcal{E}\underline{m}'\} \approx$$
$$\approx (\tfrac{1}{2} + \eta)^2 \sigma^2_{\underline{n}_1} = n\,\theta(1-\theta)(\tfrac{1}{2} + \eta)^2$$

en volgens (34):

$$(37)\quad \mathcal{E}\{\sigma^2(\underline{u}/\underline{n}_1)\} \leqq \tfrac{1}{4}\mathcal{E}\left(n + \underline{n}_1 + \sqrt{n\underline{n}_1}\right) \leqq \tfrac{1}{4}n\left(1 + \theta + \sqrt{\theta}\right)$$

daar $\mathcal{E}\sqrt{\underline{n}_1} \leqq \sqrt{\mathcal{E}\underline{n}_1}$ is.

Dus volgt, met behulp van (35):

---

4) Hierin is $\sigma^2\{\mathcal{E}(\underline{u}/\underline{n}_1)\}$ het spreidingskwadraat, genomen over de verdeling van $\underline{n}_1$, van de stochastische grootheid $\mathcal{E}(\underline{u}/\underline{n}_1)$, die verkregen wordt door in $\mathcal{E}(\underline{u}/n_1)$ de grootheid n_1 weer als stochastisch te beschouwen. De formule wordt als volgt bewezen:

$$\sigma^2_{\underline{u}} = \mathcal{E}(\underline{u} - \mathcal{E}\underline{u})^2 = \mathcal{E}\left[\mathcal{E}\{(\underline{u} - \mathcal{E}\underline{u})/\underline{n}_1\}^2\right] =$$
$$= \mathcal{E}\left[\mathcal{E}\{(\underline{u} - \mathcal{E}(\underline{u}/\underline{n}_1))^2/n_1\} + \{\mathcal{E}(\underline{u}/\underline{n}_1) - \mathcal{E}\underline{u}\}^2\right] =$$
$$= \mathcal{E}\{\sigma^2(\underline{u}/\underline{n}_1)\} + \sigma^2\{\mathcal{E}(\underline{u}/\underline{n}_1)\}$$

(38)
$$\begin{cases} \sigma_{\underline{u}}^2 \leq n\left\{\theta(1-\theta)(\tfrac{1}{2}+\eta)^2+\tfrac{1}{4}(1+\theta+\sqrt{\theta})\right\} = h_1^2 = O(n) \\ \text{terwijl} \\ \mathcal{E}\underline{u} = \mathcal{E}\underline{k}_x + \mathcal{E}\underline{m}' = \mathcal{E}(\tfrac{1}{2}+\eta)\underline{n}_1 + \mathcal{E}\underline{m}' = n\theta(\tfrac{1}{2}+\eta)+\mathcal{E}\underline{m}' \end{cases}$$

is, waarin wij $\mathcal{E}\underline{m}'$ volgens (31) verder wel kunnen verwaarlozen.

Voor $\underline{v} = \underline{n}_1 - \underline{u}$ hebben wij

(39)
$$\begin{cases} \mathcal{E}\underline{v} = \mathcal{E}\underline{n}_1 - \mathcal{E}\underline{u} = n\theta(\tfrac{1}{2}-\eta) \\ \sigma_{\underline{v}}^2 = (\sigma_{\underline{n}_1} + \sigma_{\underline{u}})^2 = h_2^2 = O(n) \end{cases}$$

Beschouwen wij nu de simultane waarschijnlijkheidsverdeling van $\underline{u}$ en $\underline{v}$ onder hypothese H_0, dan zijn $\underline{u}$ en $\underline{v}$ onderling onafhankelijk verdeeld volgens binomiale verdelingen met gemiddelde $\frac{1}{4}n$ spreidingen $\sqrt{\frac{1}{8}n}$.

De grootheden

(40)
$$\underline{u}' = \frac{\underline{u} - \frac{1}{4}n}{\sqrt{\frac{1}{8}n}} \qquad en \qquad \underline{v}' = \frac{\underline{v} - \frac{1}{4}n}{\sqrt{\frac{1}{8}n}}$$

zijn dus asymptotisch normaal (en onafhankelijk) verdeeld met gemiddelden 0 en spreidingen 1 , zodat de begrenzing van de kritieke zône Z , na reductie met behulp van de door (40) gegeven transformatie, voor $n \to \infty$ nadert tot een cirkel, met middelpunt $u'=v'=0$ en straal ϱ_α , [illegible]afhankelijk van α .

Daar

$$\frac{1}{\sqrt{2\pi}} \int_{x^2+y^2 \geq k^2} e^{-\frac{1}{2}(x^2+y^2)}\, dx\, dy = e^{-\frac{1}{2}k^2}$$

is de bij een onbetrouwbaarheidsdrempel α behorende waarde

$$\varrho_\alpha = \sqrt{-2\ln\alpha}$$

Nemen wij dus n voldoende groot en $\alpha' = \alpha + \varepsilon$ met $\varepsilon > 0$, dan behoren de gebieden

(41) $\quad |\underline{u}'| \geq \varrho_{\alpha'} \quad en \; |\underline{v}'| \geq \varrho_{\alpha'}$

tot Z. Uitgedrukt in $\underline{u}$ en $\underline{v}$ is dit:

(41') $\quad |\underline{u} - \frac{1}{4}n| \geq \varrho_{\alpha'}\sqrt{\frac{n}{8}} \qquad en \qquad |\underline{v} - \frac{1}{4}n| \geq \varrho_{\alpha'}\sqrt{\frac{n}{8}}$

Indien $\theta \neq \frac{1}{2}$ of $\eta \neq 0$ (of beide), is hetzij $\theta(\frac{1}{2}+\eta) \neq \frac{1}{4}$, hetzij $\theta(\frac{1}{2}-\eta) \neq \frac{1}{4}$ (of beide).

Zij b.v.

(42) $\quad \theta(\frac{1}{2}+\eta) > \frac{1}{4}$

dan is voor voldoende grote n

(43) $$\theta n\left(\tfrac{1}{2}+\eta\right) > \tfrac{1}{4}n + c_{\alpha'}\sqrt{\tfrac{n}{8}}$$

zodat, volgens de ongelijkheid van Bienaymé - Tchebychoff geldt:

(44) $$P\left[\underline{u} \geqq \tfrac{1}{4}n + c_{\alpha'}\cdot\sqrt{\tfrac{n}{8}} \mid \theta,\eta\right] =$$

$$= P\left[\underline{u} - \theta n\left(\tfrac{1}{2}+\eta\right) < n\left\{\tfrac{1}{2}\left(\theta-\tfrac{1}{2}\right)+\theta\eta\right\} - c_{\alpha'}\sqrt{\tfrac{n}{8}} \,/\, \theta,\eta\right]$$

$$\geqq 1 - \frac{h_1^2}{\left[n\left\{\tfrac{1}{2}\left(\theta-\tfrac{1}{2}\right)+\theta\eta\right\} - c_{\alpha'}\sqrt{\tfrac{n}{8}}\right]^2}$$

waarin h_1^2 (zie (38)) $O(n)$ is en de noemer voor voldoende grote n $> a n^2$, met constante $a > 0$, wordt. Wij hebben dus een ondergrens afgeleid voor het onderscheidingsvermogen, die voor $n \to \infty$ tot 1 nadert. Dit kan voor de van (42) verschillende mogelijkheden op analoge wijze worden gedaan. De stelling is hiermee bewezen[5].

5. Opmerkingen.

De hier beschreven toets is er één uit een groep van toetsen voor dezelfde hypothese, die vollediger beschreven wordt in

J.Hemelrijk, A family of parameterfree tests for symmetry with respect to a given point, I, Proc.Kon.Ned.Ak.v.Wet. 53 (1950) p.945-955, Indagationes Mathematicae 12 (1950) p.340-350, II verschijnt binnenkort.

De toets is daarin gegeven voor een algemenere hypothese dan de hier als H_0 vermelde; de onderstelling van continuïteit is niet gemaakt, evenmin als de onderstelling, dat alle z_i dezelfde verdeling bezitten. Alleen de onafhankelijkheidsonderstelling is daar gehandhaafd. Om de bewijzen eenvoudig te houden hebben wij hier een speciaal geval behandeld. Een uitvoerige publicatie, waarin ook verdere bruikbaarheidsstellingen worden gegeven, staat op stapel.

---

5) Verder kan men gemakkelijk bewijzen, dat de voorwaarde $\vartheta \neq \frac{1}{2}$ of $\eta \neq 0$ niet alleen voldoende, maar ook noodzakelijk is voor de bruikbaarheid.