Bijlagen bij
rapport S 241

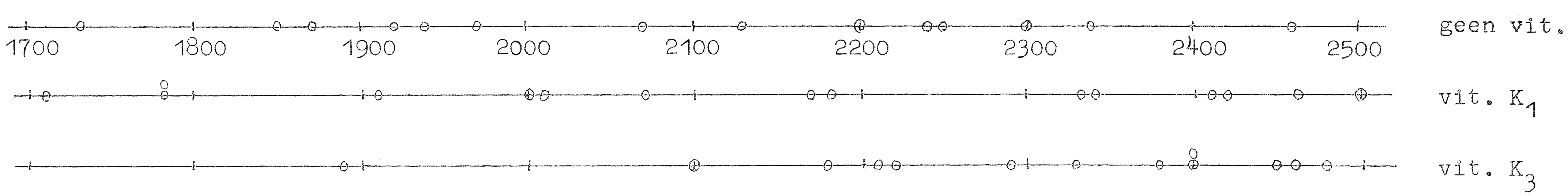
Onderzoek naar de invloed van
vitamine K bij pasgeboren kinderen (I)

door

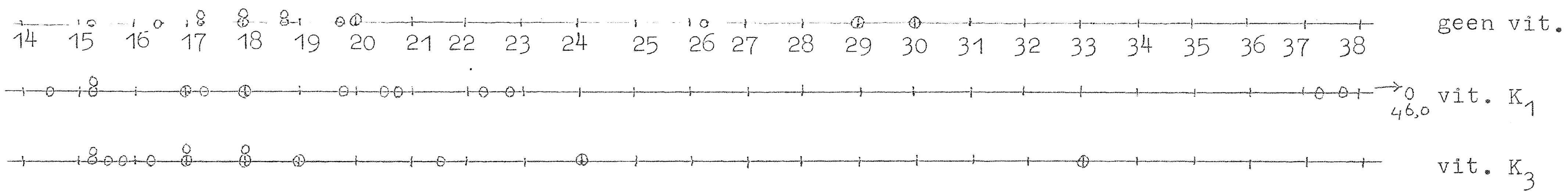
A.R.Bloemena

oktober 1958

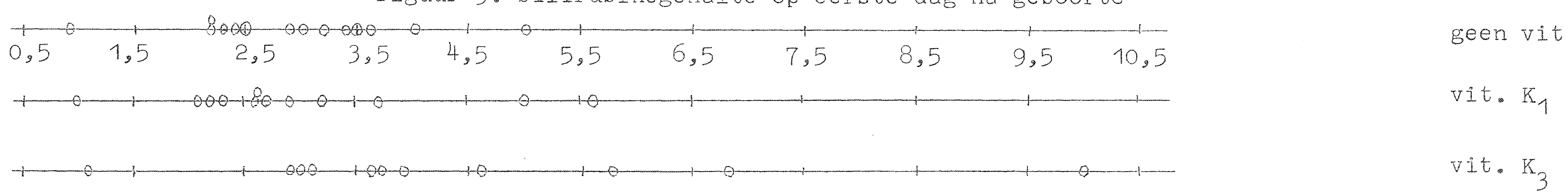
figuur 1: geboortegewicht



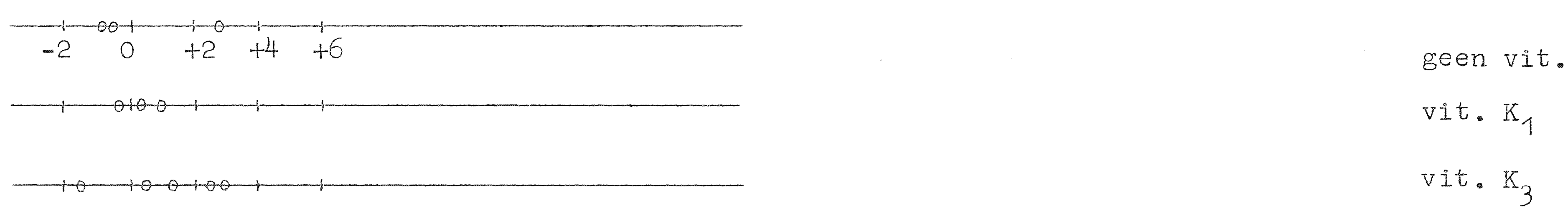
figuur 2: prothrombinetijd op eerste dag na geboorte



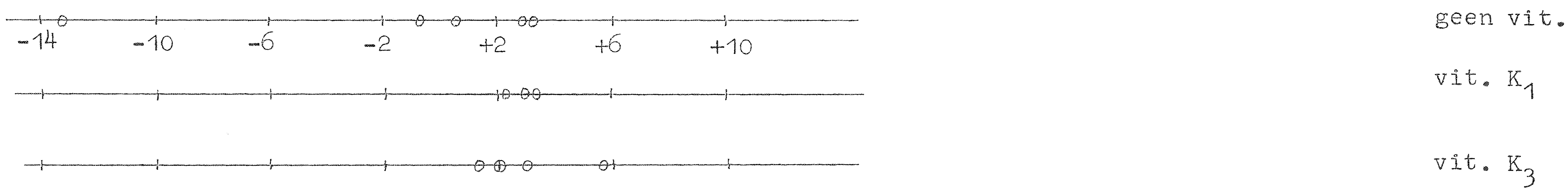
figuur 3: bilirubinegehalte op eerste dag na geboorte



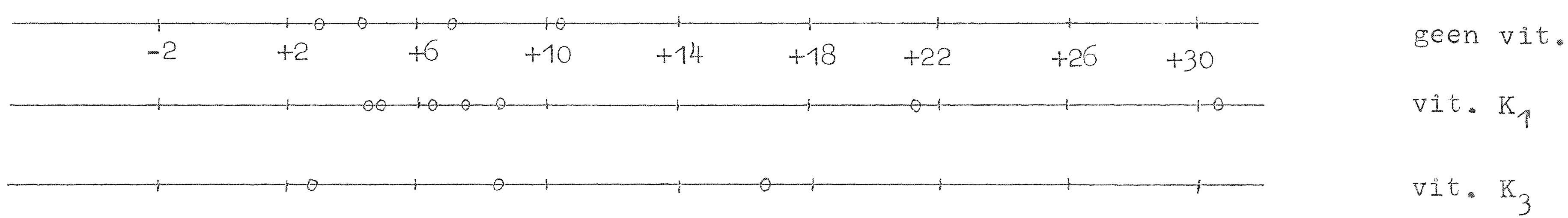
figuur 4: gemiddelde daling ptt tussen 1e en 3e dag na geboorte, groep A



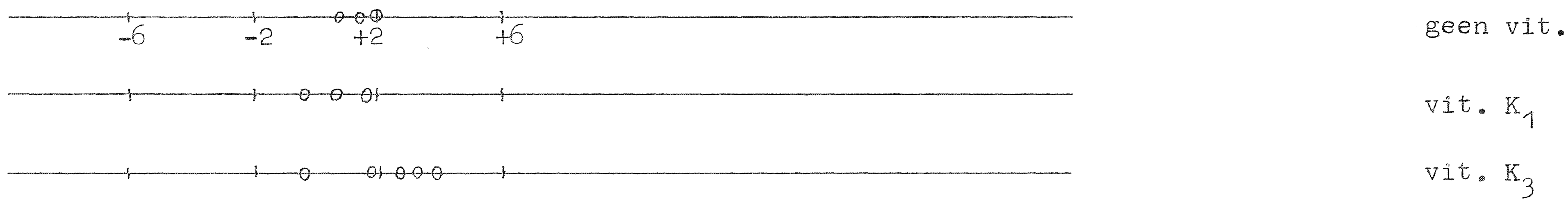
figuur 5: gemiddelde daling ptt tussen 1e en 3e dag na geboorte, groep B



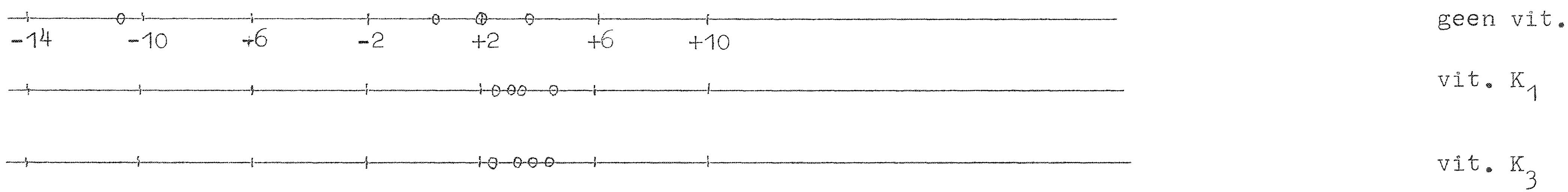
figuur 6: gemiddelde daling ptt tussen 1e en 3e dag na geboorte, groep C



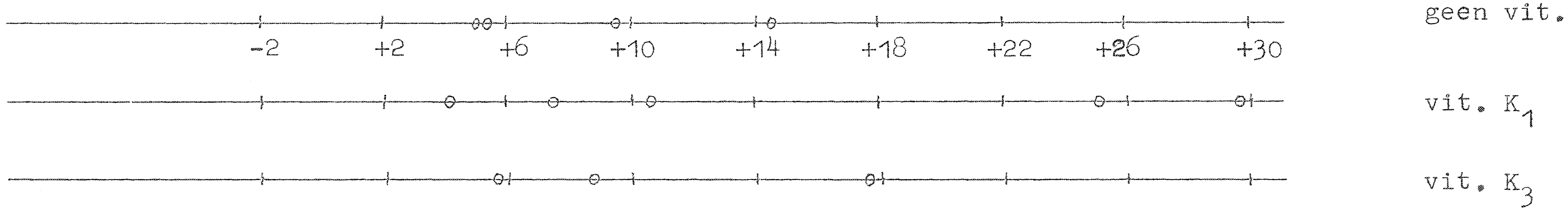
figuur 7: gemiddelde daling ptt 1e en 6e dag na geboorte, groep A



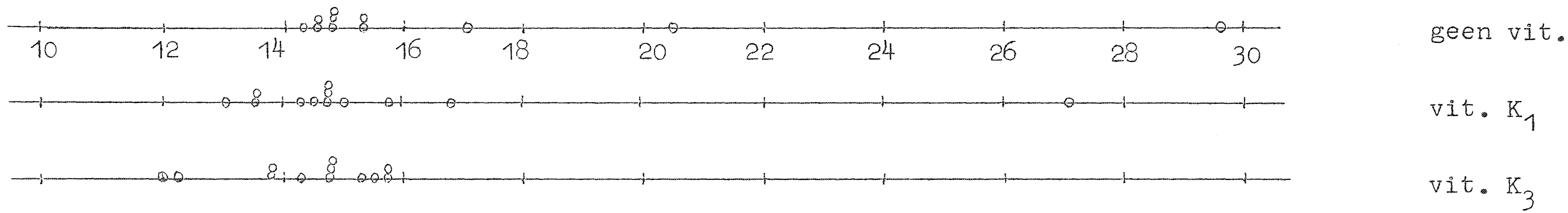
figuur 8: gemiddelde daling ptt 1e en 6e dag na geboorte, groep B



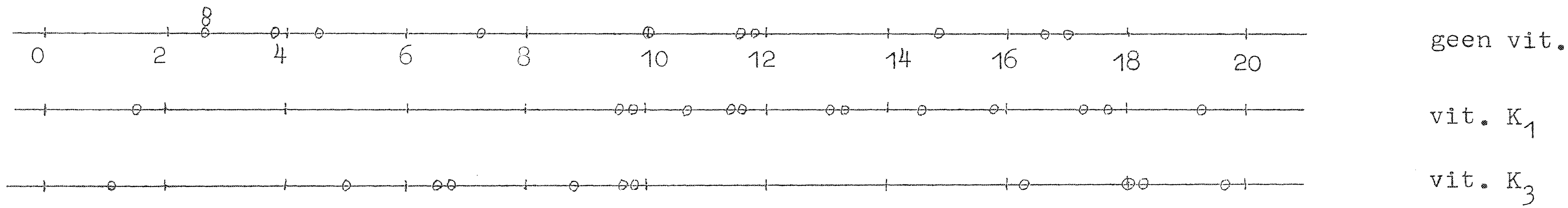
figuur 9: gemiddelde daling ptt 1e en 6e dag na geboorte, groep C



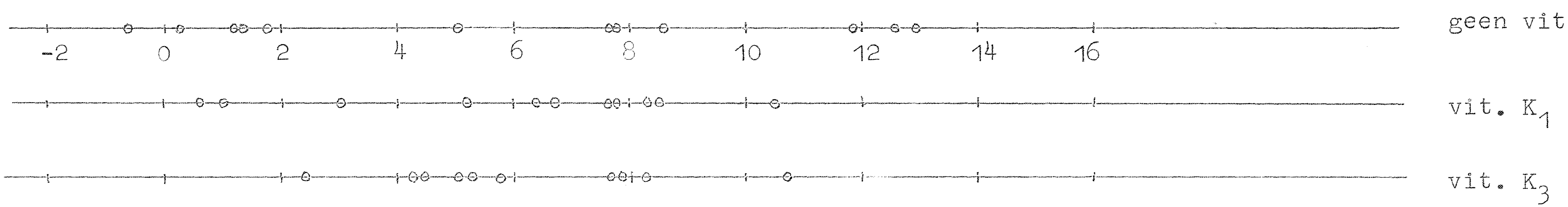
figuur 10: prothrombinetijd 6e dag na geboorte



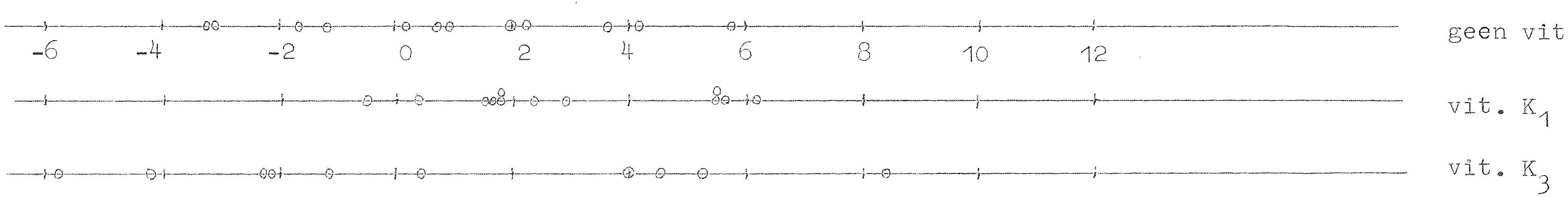
figuur 11: bilirubinegehalte 6e dag na geboorte



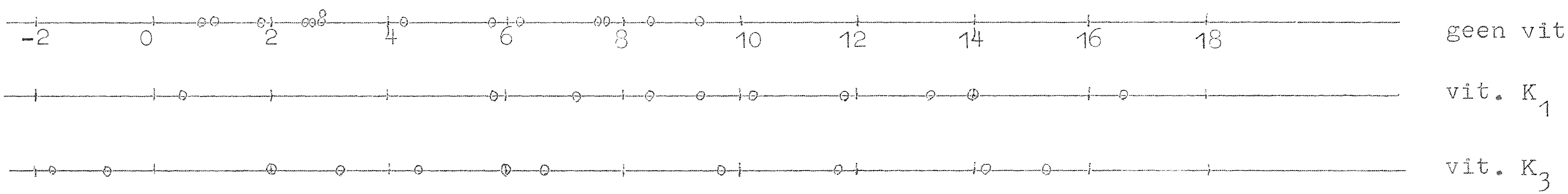
figuur 12: toename bilirubinegehalte tussen 1e en 3e dag



figuur 13: toename bilirubinegehalte tussen 3e en 6e dag



figuur 14: toename bilirubinegehalte tussen 1e en 6e dag



MATHEMATISCH CENTRUM
Amsterdam
Statistische Afdeling.

S47(M6)

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

De toetsing van een hypothese H_0 berust steeds op een aantal waarnemingen $x_1, x_2, \ldots, x_n$ van één of meer stochastische grootheden[2], of op enige groepen van waarnemingen (bv. twee steekproeven).

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

De toetsingsgrootheid wordt steeds zo gekozen, dat men, op grond van de onderstelling, dat H_0 juist is, de waarschijnlijkheidsverdeling van deze grootheid kan berekenen.

Vervolgens kiest men een verzameling Z van mogelijke uitkomsten van $\underline{u}$, en wel op zodanige wijze, dat de kans, dat $\underline{u}$ een in Z gelegen waarde aanneemt, onder de hypothese H_0, gelijk is aan een gegeven getal α, zodat Z dus van α afhankelijk is[3]. Z heet de kritieke zône van de toets, α de onbetrouwbaarheidsdrempel (Engels: level of significance). Voor α neemt men veelal de waarde 0,05 of 0,01.

Men verwerpt nu H_0 op grond van de waarnemingen $x_1, x_2, \ldots, x_n$, indien de bij deze waarnemingen behorende waarde van $\underline{u}$ in Z ligt. Dit wordt vaak uitgedrukt door te zeggen, dat het resultaat van het experiment "significant" is. De waarde van α moet dan echter worden vermeld. De kans, dat dit zal gebeuren, is, indien H_0 juist is, gelijk aan α. Derhalve is α de kans op ten onrechte verwerping van de juiste hypothese, ook de kans op een fout van de eerste soort genoemd. Indien men deze methode toepast, met α = 0,05 resp. 0,01, zal men in gemiddeld ongeveer één op 20 resp. op 100 van de gevallen, waarin de hypothese die men toetst juist is, deze toch verwerpen.

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

2) Een stochastische grootheid is een grootheid, die een waarschijnlijkheidsverdeling bezit, of, anders gezegd, een grootheid, die voor de elementen van een collectie(universum,populatie) gedefinieerd is en daarop allerlei waarden aanneemt. Stochastische grootheden worden aangegeven door onderstreepte letters.

3) Soms kan men slechts bereiken, dat deze kans $\leq \alpha$ is.

De toetsingstheorie biedt in het algemeen geen mogelijkheid om tot aanvaarding van een hypothese te komen. Indien een bepaalde hypothese $\mathcal{H}_0$ niet verworpen kan worden, is dit gewoonlijk met een hele verzameling van hypothesen tegelijk het geval. Niet-verwerpen staat dus niet gelijk met aanvaarden.

Wel zal men vaak in de loop van een statistische analyse bepaalde onderstellingen, die plausibel schijnen en voor de verdere analyse van nut zijn, toetsen, alvorens ze bij de verdere bewerking van het materiaal te gebruiken. Worden zij dan op grond van de toets niet verworpen, dan houdt dit in zo verre een rechtvaardiging van die onderstellingen in, dat een grote afwijking door de toets veelal wel zou zijn ontdekt. Indien men dan verder de onderstellingen gebruikt, verwaarloost men eventueel aanwezige afwijkingen van onbekende grootte, die echter niet zo groot zijn, dat zij door de toets zijn ontdekt.

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

In plaats van de onbetrouwbaarheidsdrempel α wordt vaak bij de uitslag van een toetsing de overschrijdingskans k opgegeven; dit is de kleinste waarde van α, waarbij in het betrokken geval, nog tot verwerping van $\mathcal{H}_0$ zou zijn overgegaan; anders gezegd: de kleinste α, waarvoor de gevonden waarde der toetsingsgrootheid nog juist in de (bij α behorende) kritieke zône Z ligt. Wordt dus de waarde k opgegeven en werkt men met onbetrouwbaarheidsdrempel α, dan wordt H_0 verworpen, indien $k \leq \alpha$ is.

Voor het onderscheid tussen één- en tweezijdige toetsing en de keuze tussen deze twee mogelijkheden vergelijke men bv. de tweede hieronder gegeven littèratuurplaats. Wij moeten hier volstaan met de opmerking, dat éénzijdige toetsing veelal eerder tot verwerping van $\mathcal{H}_0$ leidt, maar dat deze slechts onder bijzondere omstandigheden kan worden toegepast.

Litteratuur:

J.Neyman, First course in probability and statistics, New York, 1950, Chapter 5.

J.Hemelrijk en H.R. van der Vaart, Het gebruik van één- en tweezijdige overschrijdingskansen voor het toetsen van hypothesen, Statistica 4 (1950) p.54-66.

MATHEMATISCH CENTRUM,
2e Boerhaavestraat 49,
A m s t e r d a m - O.
Statistische Afdeling
Rapport S 47 (M 7)

De toets van Wilcoxon [1]

Deze methode dient tot het toetsen van de hypothese H_0, inhoudende, dat twee steekproeven $x_1, x_2, \dots, x_m$ en $y_1, y_2, \dots, y_n$ afkomstig zijn uit één collectie (ook wel populatie of universum genaamd).

De gegeven waarnemingsreeks kan dan worden samengevat als aangegeven in tabel I.

Tabel I

Voorkomende steekproef-waarden	Aantal malen dat z_i optreedt bij		totaal
	$\underline{y}$	$\underline{x}$	
z_1	$\underline{b}_1$	$\underline{a}_1$	$\underline{t}_1$
z_2	$\underline{b}_2$	$\underline{a}_2$	$\underline{t}_2$
.	.	.	.
.	.	.	.
.	.	.	.
z_k	$\underline{b}_k$	$\underline{a}_k$	$\underline{t}_k$
totaal	n	m	N

Hierin zijn $z_1, z_2, \dots, z_k$ de voorkomende steekproefwaarden, gerangschikt naar opklimmende grootte, n en m de steekproefgrootten, $t_1, t_2, \dots, t_k$ de grootten der groepen gelijke waarnemingen en $\underline{a}_i$[2] (resp. $\underline{b}_i$) het aantal malen dat de waarde z_i in de eerste (resp. tweede) steekproef optreedt; $\underline{a}_i$ kan dus de waarden $0, 1, \dots, t_i$ aannemen ($i = 1, 2, \dots, k$).

Voor het toetsen van de hypothese H_0 wordt gebruik gemaakt van een toetsingsgrootheid $\underline{W}$[2], die gelijk is aan twee maal

----

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.
2) Een stochastische grootheid is een grootheid, die een waarschijnlijkheidsverdeling bezit, hetgeen gewoonlijk blijkt uit het feit, dat bij herhaalde waarneming onder dezelfde omstandigheden verschillende uitkomsten verkregen worden. Stochastische grootheden worden door onderstreepte letters aangegeven; dezelfde letters, niet onderstreept, worden vaak gebruikt voor waarden, die zij aan kunnen nemen of aangenomen hebben.

het aantal paren waarnemingen (x_i, y_j), waarvoor $x_i > y_j$ is, vermeerderd met het aantal paren (x_k, y_ℓ) waarvoor $x_k = y_\ell$ is.[3)]

Het is duidelijk, dat W een kleine waarde zal aannemen, indien de x-waarden overwegend kleiner zijn dan de y-waarden, een grote, indien het omgekeerde het geval is en een intermediaire waarde, indien geen van deze beide situaties zich voordoet.

Indien de hypothese H_0 juist is, dus als de twee steekproeven uit dezelfde verdeling afkomstig zijn, kan de verdeling van de toetsingsgrootheid W berekend worden. Deze verdeling heeft de volgende eigenschappen:

1. De verdeling van W is discreet, want W kan alleen gehele waarden aannemen.
2. De verdeling van W is in het algemeen niet symmetrisch, behalve,

a) als $m = n$,

b) als $t_1 = t_k$, $t_2 = t_{k-1}$ enz.

3. De verwachting (het theoretische gemiddelde) μ, van de verdeling van W is gelijk aan het product der steekproefgrootten

(1) $$\mu = mn.$$

4. De variantie van de verdeling van W wordt gegeven door:

(2) $$\sigma^2 = \frac{mn\,(N^3 - D)}{3N(N-1)},$$

waarin $D = t_1^3 + t_2^3 + \ldots + t_k^3$.

Bevatten de beide steekproeven geen gelijke waarnemingen, dan wordt $D = N$ en dan is dus

$$N^3 - D = N^3 - N = N(N-1)(N+1)$$

en in dit geval wordt de variantie

(3) $\sigma^2 = \frac{1}{3}\,mn(N+1)$ (geen gelijke waarnemingen).

Wij zagen reeds dat de toetsingsgrootheid W een kleine waarde aan zal nemen, indien de x-waarden overwegend kleiner zijn dan de y-waarden een een grote waarde, indien het omgekeerde het geval is. Als tweezijdige kritieke zone nemen wij dus grote waarden van $|W - \mu|$.

---

3) Gewoonlijk wordt voor de toets van WILCOXON de toetsingsgrootheid $U = \frac{1}{2}W$ gebruikt. Om gebroken getallen te vermijden, wordt hier W ingevoerd (zie [4], [6]).

Als m en n groot zijn en onderling niet te veel verschillen en als bovendien de groepen gelijke waarnemingen niet te veel in omvang verschillen, kan men gebruik maken van het feit, dat de grootheid $\underline{W}$, als de hypothese H_0 juist is, bij benadering normaal verdeeld is met gemiddelde en variantie volgens (1) en (2).

De tweezijdige overschrijdingskans k, behorende bij een waarde W van $\underline{W}$ is dus

$$k = \frac{2}{\sqrt{2\pi}} \int_{\frac{|W-\mu|-1}{\sigma}}^{\infty} e^{-\frac{1}{2}u^2} \, du \qquad 4)$$

en kan in een tabel van de normale verdeling worden gevonden.

Voor $m \leq n \leq 10$, resp. voor $m + n \leq 40$, $n \geq 11$ en $m \leq n$ zijn tabellen van de eenzijdige overschrijdingskansen, resp. linker kritieke waarden beschikbaar (zie [6] tabel I en II). Deze tabellen gelden strikt genomen alleen als er geen gelijke waarnemingen zijn, maar zij geven voor het geval van weinig gelijke waarnemingen veelal een goede benadering.

Indien m en n klein zijn en bovendien onder de waarnemingen grote groepen gelijken voorkomen, kan men de overschrijdingskansen noch met behulp van de normale benadering van de verdeling van $\underline{W}$, noch met behulp van de genoemde tabellen bepalen. Hoe men in zulke gevallen dient te handelen, staat beschreven in [1].

<u>Opmerking</u>: Als er gelijke waarnemingen zijn, dan geldt

$$\frac{1}{3}\, m\, n\, (N+1) > \frac{m\, n\, (N^3 - D)}{3N(N-1)}.$$

Dus als wij, in de formule voor de variantie van $\underline{W}$, de gelijke waarnemingen niet in rekening brengen, vinden wij een te grote waarde voor deze variantie en dus een te grote waarde voor de overschrijdingskans. Is deze te grote waarde voor de overschrijdingskans $\leq \alpha$ dan wordt H_0 dus verworpen en hoeft men de berekening der derde machten niet uit te voeren.

----

4) De term -1 in de teller van $\frac{|W-\mu|-1}{\sigma}$ is de z.g. continuïteitscorrectie (zie [6], pag.16).

Literatuur

[1] van Eeden, Constance en Ir Doraline Wabeke, Handleiding voor de toets van WILCOXON (vervolg): Exacte behandeling als er gelijke waarnemingen zijn. Rapport S 176 (M 65A) van de Statistische Afdeling van het Mathematisch Centrum, Amsterdam, 1955.

[2] Hemelrijk, J., Note on WILCOXON's two sample test when ties are present, Ann.Math.Stat. 23 (1952), 133-135.

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

[4] van der Vaart, H.R., Gebruiksaanwijzing voor de toets van WILCOXON, Rapport S 32 (M 4) van de Statistische Afdeling van het Mathematisch Centrum, Amsterdam, 1950.

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

[6] Wabeke, Ir Doraline en Constance van Eeden, Handleiding voor de toets van WILCOXON, Rapport S 176 (M 65) van de Statistische Afdeling van het Mathematisch Centrum, Amsterdam 1955.

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

Mei 1956

Statistische Afdeling
Rapport S 190 (M 13)

De rangcorrelatietoets van KENDALL[1)]

1. Doel en toepasbaarheid

De door M.G. KENDALL ontwikkelde rangcorrelatietoets is toepasbaar op de volgende situatie:

De stochastische grootheden $\underline{x}$[2)] en $\underline{y}$ bezitten een simultane verdeling en (x_i, y_i) $(i=1,2,\ldots,n)$ zijn n onderling onafhankelijke waarnemingsparen van deze stochastische grootheden.

In dit memorandum wordt een methode beschreven om op grond van dit waarnemingsmateriaal de hypothese H_0 te toetsen dat $\underline{x}$ en $\underline{y}$ onderling onafhankelijk verdeeld zijn. Over de vorm van de verdeling van $\underline{x}$ en $\underline{y}$ behoeft hierbij niets ondersteld te worden.

2. De toetsingsgrootheid

KENDALL's toetsingsgrootheid $\underline{S}$ wordt als volgt gedefiniëerd. De waarnemingsparen (x_i, y_i) en (x_j, y_j) leveren ieder tot S een bijdrage

+1 als $(x_i - x_j)(y_i - y_j) > 0$,
 0 als $(x_i - x_j)(y_i - y_j) = 0$,
-1 als $(x_i - x_j)(y_i - y_j) < 0$.

S is nu de algebraïsche som van de bijdragen van alle paren (x_i, y_i) en (x_j, y_j), waarvoor $i < j$ is.
Wij zullen de berekening van de toetsingsgrootheid $\underline{S}$ aan de hand van twee voorbeelden nader toelichten.

Voorbeeld 1:

Stel wij hebben de volgende zes waarnemingsparen (x_i, y_i)

i	1	2	3	4	5	6
x_i	0,11	0,12	0,10	0,11	0,15	0,13
y_i	3,4	3,0	3,2	3,5	3,5	3,5

De grootheid S berekenen wij als volgt: Het waarnemingspaar (x_1, y_1) levert met de paren $(x_2, y_2), \ldots, (x_6, y_6)$ resp. de bijdragen -1, +1, 0, +1, +1. De totale bijdrage van het paar (x_1, y_1) tot S is dus -1+1+1+1 = +2. Op dezelfde manier vinden wij voor de bijdragen

---

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.
2) Stochastische grootheden worden onderscheiden van getallen (b.v. van de waarden die zij bij een experiment aannemen) door hun symbolen te onderstrepen.

van de paren $(x_2, y_2), \ldots, (x_6, y_6)$ tot S:

(x_2, y_2): $-1-1+1+1 = 0$
(x_3, y_3): $+1+1+1 = 3$
(x_4, y_4): $0 + 0 = 0$
(x_5, y_5): 0

De grootheid S is dus $2+0+3+0+0 = +5$.

Voorbeeld 2:

Stel dat x een grootheid is die slechts de drie waarden 1, 2 en 3 aanneemt en y een die slechts de vier waarden 1, 2, 3 en 4 aanneemt. Bij 30 onderling onafhankelijke waarnemingsparen van deze grootheden wordt het volgende resultaat verkregen:

		x			
		1	2	3	totaal
y	1	6	2	0	8
	2	1	4	2	7
	3	1	3	2	6
	4	1	1	7	9
totaal		9	10	11	30

In deze tabel staan aantallen waarnemingen vermeld. Er zijn b.v. 6 paren (x_i, y_i) met $x_i = y_i = 1$, 3 paren (x_i, y_i) met $x_i = 2$, $y_i = 3$ enz. Verder zien wij dat er onder de 30 x-waarnemingen 9× een 1, 10× een 2 en 11× een 3 voorkomt; onder de 30 y-waarnemingen komt 8× een 1, 7× een 2, 6× een 3 en 9× een 4 voor.

In een geval zoals dit (dus een geval met veel gelijke waarnemingen) kan men het eenvoudigst de grootheid $2S$ berekenen, dus de som van de bijdragen van alle paren (x_i, y_i) en (x_j, y_j) zowel voor $i < j$ als voor $i > j$.

De bijdrage tot $2S$ voor ieder der 4 waarnemingsparen (x_i, y_i) met $x_i = y_i = 2$ b.v. vinden wij als volgt: er zijn 2+7+6 paren (x_j, y_j) met $(x_i - x_j)(y_i - y_j) > 0$, nl. de 2 paren met $x_j = y_j = 3$, de 7 paren met $x_j = 3$, $y_j = 4$ en de 6 paren met $x_j = y_j = 1$. Verder zijn er 2 paren (x_j, y_j) met $(x_i - x_j)(y_i - y_j) < 0$, nl. het paar met $x_j = 1$, $y_j = 3$ en het paar met $x_j = 1, y_j = 4$. De bijdrage van ieder van deze 4 paren tot $2S$ is dus $15-2 = 13$; tezamen geven zij dus een bijdrage $4 \times 13 = 52$.

Op deze wijze kan men voor ieder der vakjes in de bovenstaande tabel de bijdrage tot $2S$ berekenen. Dit geeft:

	1	2	3
1	114	16	0
2	11	52	4
3	0	33	22
4	-13	4	119

Dus

$$2S = 114 + 11 - 13 + 16 + 52 + 33 + 4 + 4 + 22 + 119 = 362$$

$$S = 181.$$

3. Kritieke zônes en overschrijdingskansen

Als de hypothese H_0 juist is, dan is de verwachting van $\underline{S}$ gelijk aan 0. Is H_0 onjuist, dan zal $\underline{S}$ in het algemeen grote positieve waarden aannemen als $\underline{x}$ en $\underline{y}$ positief gecorreleerd zijn en grote negatieve waarden als $\underline{x}$ en $\underline{y}$ negatief gecorreleerd zijn. De tweezijdige kritieke zône bestaat daarom uit grote waarden van $|S|$, de linkseenzijdige uit grote negatieve en de rechtseenzijdige uit grote positieve waarden van S.

Tabellen van overschrijdingskansen kan men vinden in [1] (pag. 141) voor $n = 4$ t/m 10 en in [2] (tabel I en II) voor $n =$ 4 t/m 40. Bovendien vindt men in [2] (tabel III) de linker kritieke waarden van S voor de onbetrouwbaarheidsdrempels $\alpha = 0{,}005$; 0,01; 0,025; 0,05 en 0,10 en $n = 4$ t/m 40. Deze tabellen gelden strikt genomen alleen voor het geval dat er noch bij de x_i noch bij de y_i gelijke waarnemingen voorkomen, maar zij geven een goede benadering voor het geval van weinig gelijke waarnemingen.

Voor grote waarden van n kan men veelal gebruik maken van de benadering met de normale verdeling. Men berekent daartoe de variantie van $\underline{S}$ onder de hypothese H_0 als volgt: stel dat de waarnemingen van $\underline{x}$ (resp van $\underline{y}$) uiteenvallen in g (resp. h) groepen gelijke waarnemingen en dat de aantallen waarnemingen in deze groepen $t_1, t_2, \ldots, t_g$ (resp. $u_1, u_2, \ldots, u_h$) zijn. Dan is

$$\sum_{i=1}^{g} t_i = \sum_{j=1}^{h} u_j = n.$$

In voorbeeld 1 is $g=5, h=4, t_1=t_2=t_3=t_4=1, t_5=2$ en $u_1 = u_2 = u_3 = 1, u_4 = 3$; in voorbeeld 2 is: $g=3, h=4, t_1=9, t_2=10, t_3=11$ en $u_1=8, u_2=7, u_3=6, u_4=9$.

De variantie van $\underline{S}$ wordt dan gevonden uit de formule

$$\sigma^2\{\underline{S}\} = \frac{2\{n^3-\sum_{i=1}^{g} t_i^3 - 3(n^2-\sum_{i=1}^{g} t_i^2)\}\{n^3-\sum_{j=1}^{h} u_j^3 - 3(n^2-\sum_{j=1}^{h} u_j^2)\} + 9(n-2)(n^2-\sum_{i=1}^{g} t_i^2)(n^2-\sum_{j=1}^{h} u_j^2)}{18n(n-1)(n-2)} .$$

Als er noch onder de x_i noch onder de y_i gelijke waarnemingen voorkomen dan is

$$t_i = 1 \quad \text{voor iedere } i$$
$$u_j = 1 \quad \text{voor iedere } j$$
$$g = h = n .$$

De variantie van $\underline{S}$ wordt dan

$$\sigma^2\{\underline{S}\} = \frac{1}{18}\, n(n-1)(2n+5) .$$

Een tabel van $\sigma\{\underline{S}\}$ voor dit geval vindt men voor n = 40 t/m 100 in [2] (tabel IV).

De grootheid $\frac{\underline{S}}{\sigma\{\underline{S}\}}$ is nu, voor grote waarden van n en als, zowel bij de x_i als bij de y_i, de groepen gelijke waarnemingen niet te veel in grootte verschillen, bij benadering normaal verdeeld met gemiddelde 0 en spreiding 1. De overschrijdingskans kan dan dus in een tabel der normale verdeling worden opgezocht. Hierbij past men gewoonlijk een continuïteitscorrectie 1 toe, d.w.z. als rechteroverschrijdingskans neemt men het oppervlak de normale verdeling rechts van $\frac{S-1}{\sigma\{\underline{S}\}}$, als linkeroverschrijdingskans het oppervlak links van $-\frac{S+1}{\sigma\{S\}}$ en als tweezijdige overschrijdingskans het oppervlak links van $-\frac{|S|-1}{\sigma\{S\}}$ plus het oppervlak rechts van $\frac{|S|-1}{\sigma\{S\}}$. Als er bij de x_i òf bij de y_i, doch niet bij beide, tweetallen of drietallen gelijken voorkomen, dus als b.v.

$$t_i = 1 \quad \text{voor iedere } i$$
$$u_j \leq 3 \quad \text{voor iedere } j ,$$

dan kan men als $n \leq 10$ is gebruik maken van exacte tabellen van SILLITTO [4] .

In gevallen waarin men geen gebruik kan maken van de exacte tabellen en waar de normale benadering niet van toepassing is, moet men de exacte verdeling zelf berekenen. Hierop zullen wij niet verder ingaan. Voorbeeld 1 is zo'n geval. Bij voorbeeld 2 kan men echter gebruik maken van de normale benadering. Hier was

$$S = 181$$
$$n = 30$$
$$n^3 - \sum_{i=1}^{g} t_i^3 = 22146 \qquad n^2 - \sum_{i=1}^{g} t_i^2 = 898$$
$$n^3 - \sum_{j=1}^{h} u_j^3 = 23190 \qquad n^2 - \sum_{j=1}^{h} u_j^2 = 670 .$$

dus

$$\sigma^2\{S\} = \frac{2.22146.23190 + 9.28.598.670}{18.30.29.28} = \frac{1128097800}{438480} = 2572{,}75$$

$$\sigma\{S\} = 50{,}72$$

$$\frac{S-1}{\sigma\{S\}} = 3{,}55 .$$

In een tabel van de normale verdeling vindt men voor de tweezijdige overschrijdingskans 0,0004.

4. De rangcorrelatiecoëfficiënt τ

Als maat voor de correlatie in de rij waarnemingsparen $(x_1, y_1), \dots, (x_n, y_n)$ heeft KENDALL de coëfficiënt τ gedefiniëerd, die +1 is als de volgorden in de twee rijen volledig overeenstemmen en -1 als deze volgorden volkomen tegengesteld zijn. De definitie van τ is

$$\tau = \frac{2S}{\left\{n^2 - \sum_{i=1}^{g} t_i^2\right\}^{\frac{1}{2}} \left\{n^2 - \sum_{j=1}^{h} u_j^2\right\}^{\frac{1}{2}}} .$$

Als er in geen van beide rijen gelijken voorkomen, dan wordt deze formule

$$\tau = \frac{2S}{n(n-1)} .$$

Literatuur

[1] KENDALL, M.G., Rank Correlation Methods, London 1948.

[2] KAARSEMAKER, L. en A. VAN WIJNGAARDEN, Tables for use in Rank Correlation, Report R 73 of the Computation Department of the Mathematical Centre (1952) en Statistica 7 (1953), p. 41-54.

[3] HEMELRIJK, J., Kendall's rangcorrelatiecoëfficiënt τ, Hoofdstuk I van de Cursus "Parametervrije Methoden", Rapport S 59 van het Mathematisch Centrum (1951).

[4] SILLITTO, G.P., The distribution of Kendall coefficient of rank correlation in rankings containing ties Biometrika 34 (1947), p. 36-40.

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

Statistische Afdeling
S 53 (M 23).

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

Wij beschouwen twee reeksen van onafhankelijke experimenten, waarbij ieder experiment van de ene reeks één van de twee resultaten A of $\bar{A}$ (non-A) heeft en ieder experiment van de tweede reeks één van de beide resultaten B of $\bar{B}$ (hierbij kan A=B zijn). Daarbij wordt ondersteld, dat bij ieder der experimenten van de ene reeks de kans op A gelijk aan p_1 (en dus de kans op $\bar{A}$ gelijk aan $1-p_1$) is en bij ieder der experimenten van de tweede reeks de kans op B gelijk aan p_2 (en dus de kans op $\bar{B}$ gelijk aan $1-p_2$). De te toetsen hypothese luidt nu:

$$H_o : p_1 = p_2.$$

Indien de eerste reeks uit n en de tweede reeks uit m waarnemingen bestaat, waaronder a (resp. b) maal A (resp. B) voorkomt, kunnen deze gegevens in de volgende 2 x 2-tabel worden samengevat:

	A resp. B	$\bar{A}$ resp. $\bar{B}$	totaal
eerste reeks	a	c	n
tweede reeks	b	d	m
totaal	r	s	N

Als toetsingsgrootheid wordt a, het aantal malen A in de eerste reeks waarnemingen, gebruikt. Indien H_o, juist is bezit deze grootheid onder de voorwaarde, dat r de bij het experiment gevonden waarde aanneemt, de volgende waarschijnlijkheidsverdeling: de kans, dat een bepaalde waarde a aangenomen wordt, is gelijk aan:

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

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

---

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

lijkheid van de gevonden waarde zijn (bij éénzijdige toetsing echter gelijk aan de som van de waarschijnlijkheden van alle waarden die groter of gelijk aan de gevondene, of van alle waarden, die kleiner of gelijk aan de gevondene zijn). Deze exacte toetsingsmethode voor H_0 is afkomstig van R.A.FISHER.

Indien n en m zo groot zijn, dat deze exacte berekening te omslachtig wordt, maakt men gebruik van de volgende benadering:

Gemiddelde en spreiding van de grootheid a zijn (indien H_0 juist is):

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

Men gebruikt dan in plaats van de exacte waarschijnlijkheidsverdeling van a de normale verdeling met hetzelfde gemiddelde en dezelfde spreiding en in plaats van de gevonden waarde van a neemt men het getal dat $\frac{1}{2}$ dichter bij het gemiddelde ligt dan deze gevonden waarde (dit laatste is de z.g. "continuïteitscorrectie", die bij toenemende n en m weldra verwaarloosd kan worden). Bij positieve $a - \frac{nr}{N}$ berekent men dus:

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

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

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

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

Literatuur.

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

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

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
Amsterdam-O.
Statistische Afdeling
S 73 (M 30a)

Twee parametervrije toetsen voor k steekproeven.[1]

Gegeven zijn k onafhankelijke steekproeven van k stochastische grootheden:

$x_{1,1}, x_{1,2},\dots,x_{1,n_1}$ van de grootheid $\underline{x}_1$ [2],

$x_{2,1}, x_{2,2},\dots,x_{2,n_2}$ van de grootheid $\underline{x}_2$,

----

$x_{k,1}, x_{k,2},\dots,x_{k,n_k}$ van de grootheid $\underline{x}_k$.

De uitgebreidheden van de steekproeven zijn dus $n_1,n_2,\dots,n_k$; de eerste index van de waarneming $x_{i,j}$ geeft aan, uit welke steekproef deze waarneming afkomstig is, terwijl de tweede index het nummer der waarneming binnen die steekproef aangeeft.

De hypothese H_0, die wij wensen te toetsen, luidt dat de waarnemingen van dezelfde stochastische grootheid afkomstig zijn.
Anders uitgedrukt: H_0 houdt in, dat de stochastische grootheden $\underline{x}_1, \underline{x}_2,\dots,\underline{x}_k$ onderling onafhankelijk verdeeld zijn en alle dezelfde waarschijnlijkheidsverdeling bezitten.
Om H_0 te toetsen kunnen we gebruik maken van de volgende twee toetsingsmethoden.

1. De H-toets.

De berekening van de toetsingsgrootheid $\underline{H}$, welke bij de toets gebruikt wordt, geschiedt op de volgende wijze (zie KRUSKAL en WALLIS [1] en [2], RIJKOORT [3] en TERPSTRA [4]).
Alle waarnemingen worden naar opklimmende grootte gerangschikt en vervolgens van een rangnummer voorzien. Indien alle waarnemingen verschillend zijn, zijn dit de rangnummers 1 tot en met $n = n_1+n_2+\dots+n_k$. Indien t waarnemingen gelijk aan elkaar zijn, wordt aan elk der t waarnemingen eenzelfde rangnummer toegekend. Dit rangnummer is dan gelijk aan het gemiddelde der rangnummers, welke de waarnemingen gekregen zouden

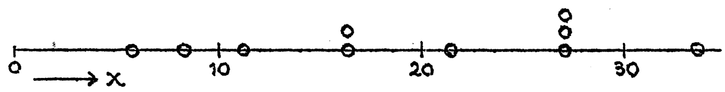
----

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

2) Stochastische grootheden worden door onderstreepte letters aangeduid.

hebben, indien ze alle verschillend waren.

Voorbeeld:



rangnummers: 1, 2, 3, $4\frac{1}{2}$, $4\frac{1}{2}$, 6, 8, 8, 8, 10.

Stelt nu $\underline{R}_i$ de som van de rangnummers in de i^e steekproef voor en zijn alle waarnemingen onderling verschillend, dan wordt de toetsingsgrootheid $\underline{H}$ gedefiniëerd door

$$\underline{H} = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{\underline{R}_i^2}{n_i} - 3(n+1).$$

Komen onder alle n waarnemingen tezamen r groepen van gelijken voor, waarbij de eerste groep uit t_1 waarnemingen bestaat, de tweede groep uit t_2 waarnemingen etc., dan wordt bovenstaande uitdrukking voor $\underline{H}$ gedeeld door de factor

$$N = \left[1 - \{(t_1-1)t_1(t_1+1) + \ldots + (t_r-1)t_r(t_r+1)\}\right] / (n^3-n).$$

Indien alle waarnemingen verschillend zijn is $N = 1$.
In het algemeen geldt dus:

$$\underline{H} = \left[\frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{\underline{R}_i^2}{n_i} - 3(n+1)\right] / N.$$

Voor het geval van twee steekproeven ($k = 2$) is $\underline{H}$ gelijk aan de genormeerde toetsingsgrootheid van WILCOXON [6] voor twee steekproeven. De H-toets kan dus beschouwd worden als een generalisatie van de toets van WILCOXON.

Beschouwen wij nu de verzameling van alle bij het experiment mogelijke uitkomsten, dan bezit $\underline{H}$ op deze verzameling een waarschijnlijkheidsverdeling. Indien de hypothese H_0, inhoudende dat alle waarnemingen uit dezelfde verdeling afkomstig zijn, juist is, dan bezit $\underline{H}$ voor grote waarden van $n_1, n_2, \ldots, n_k$ bij benadering een χ^2-verdeling met k-1 vrijheidsgraden.

Indien de hypothese H_0 niet vervuld is, bezit $\underline{H}$ gemiddeld grotere waarden, dan indien H_0 wel vervuld is. De hypothese H_0 wordt daarom verworpen, indien de uit de steekproef bepaalde waarde van $\underline{H}$ groter is dan een kritieke waarde H_α, welke gegeven wordt door

$$P[\underline{H} \geq H_\alpha | H_0] = P[\chi^2_{k-1} \geq H_\alpha | H_0] = \alpha.$$

Hierin is α de vooraf gekozen onbetrouwbaarheidsdrempel: indien de gevonden waarde H groter is dan H_α wordt de hypothese H_0 verworpen met een onbetrouwbaarheid α.
De waarde H_α kan gemakkelijk bepaald worden uit tabellen en nomogrammen van de χ^2-verdeling, evenals de overschrijdingskans van de gevonden waarde H, welke gedefinieerd wordt door:

$$P[\underline{H} \geq H | H_0] = P[\chi^2_{k-1} \geq H | H_0].$$

Een benadering voor kleine steekproeven.

Indien de steekproeven niet voldoende groot zijn, wordt de exacte verdeling van $\underline{H}$ niet meer goed benaderd door de χ^2-verdeling.
Voor het geval van 3 steekproeven met $n_i \leq 5$ $(i=1,2,3)$ zijn de kritieke waarden van $\underline{H}$, corresponderende met de onbetrouwbaarheden $\alpha = 0{,}10$, $0{,}05$ en $0{,}01$, exact berekend (zie [2] en [3]).
Aan de hand van deze uitkomsten is aangetoond, dat de exacte verdeling van $\underline{H}$ voor kleine steekproeven zeer goed benaderd wordt door de F-verdeling. Deze verdeling wordt gekarakteriseerd door 2 grootheden ν_1 en ν_2, de zgn. aantallen graden van vrijheid (zie b.v. [5]).
De uit de steeproefwaarden te berekenen grootheid F wordt dan gegeven door

$$\underline{F} = \frac{H(M-k+1)}{(k-1)(M-H)},$$

waarin

$$M = \frac{n^3 - \sum_{i=1}^{k} n_i^3}{n(n+1)}.$$

Verder worden ν_1 en ν_2 gegeven door

$$\nu_1 = 2(k-1)\,\frac{(k-1)(M-k+1)-\upsilon}{M\upsilon}$$

en

$$\nu_2 = 2(M-k+1)\,\frac{(k-1)(M-k+1)-\upsilon}{M\upsilon} = \frac{M-k+1}{k-1}\cdot\nu_1,$$

waarin

$$\upsilon = 2(k-1) - \frac{2[3k^2-6k+n(2k^2-6k+1)]}{5n(n+1)} - \frac{6}{5}\sum_{i=1}^{k}\frac{1}{n_i}.$$

De kritieke waarden voor $\underline{F}$ en de overschrijdingskans van de gevonden waarde F kunnen worden bepaald uit tabellen en nomogrammen van de F-verdeling. Grote waarden van F leiden tot verwerping van H_0.

2. De T^2-toets.

De bij deze toets gebruikte toetsingsgrootheid $\underline{T}^2$ wordt als volgt berekend (TERPSTRA [4]).
Evenals bij de H-toets, worden allereerst alle waarnemingen naar opklimmende grootte gerangschikt (voor gelijke waarnemingen geldt hetzelfde als onder 1), en wordt voor iedere steekproef de som der rangnummers $\underline{R}_i$ $(i = 1,2,\ldots,k)$ bepaald.
Bovendien wordt nog ieder tweetal van steekproeven onderling vergeleken. Dit gaat het gemakkelijkst door eerst de steekproeven willekeurig te nummeren.
De $h^{\underline{e}}$ en $j^{\underline{e}}$ steekproef $(h \leq j)$ worden dan vergeleken door de n_h+n_j waarnemingen uit de twee steekproeven naar opklimmende grootte te rangschikken en van een rangnummer te voorzien (voor gelijke waarnemingen geldt weer hetzelfde als onder 1). Voor de $h^{\underline{e}}$ steekproef wordt vervolgens de som der rangnummers bepaald, welk aantal we $\underline{R}_h^{(j)}$ noemen.
De toetsingsgrootheid $\underline{T}^2$ wordt dan gedefinieerd door

$$\underline{T}^2 = 12 \sum_{h<j} \frac{\underline{\tilde{U}}_{h,j}^2}{n_h n_j} - \frac{12}{n+1} \sum_{i=1}^{k} \frac{\underline{\tilde{U}}_i^2}{n_i},$$

waarin $\qquad \underline{\tilde{U}}_{h,j} = \underline{R}_h^{(j)} - \tfrac{1}{2} n_h(n_h+n_j+1),$

en $\qquad \underline{\tilde{U}}_i = \underline{R}_i - \tfrac{1}{2} n_i(n+1).$ 3)

De toetsingsgrootheid $\underline{T}^2$ kan ook geschreven worden als

$$\underline{T}^2 = 12 \sum_{h<j} \frac{\underline{\tilde{U}}_{h,j}^2}{n_h n_j} - n\underline{H},$$

----

3) De grootheden $\underline{\tilde{U}}_i$ en $\underline{\tilde{U}}_{h,j}$ staan in verband met de toetsingsgrootheid $\underline{U}$ van WILCOXON [6], zoals ze gedefinieerd is door MANN en WHITNEY [7]. Beschouwen we 2 steekproeven van de uitgebreidheden n en m, dan is $\underline{U}$ het aantal keren, dat een waarneming uit de tweede steekproef kleiner is dan een waarneming uit de eerste steekproef (het zgn. aantal inversies) en is $\underline{\tilde{U}} = \underline{U} - \frac{1}{2}nm$. De grootheden $\underline{\tilde{U}}_i$ en $\underline{\tilde{U}}_{h,j}$ hebben betrekking op de i^e steekproef en de overige steekproeven tezamen, resp. op de $h^{\underline{e}}$ en de $j^{\underline{e}}$ steekproef.

waarin H gegeven wordt door de teller van de algemene uitdrukking onder 1e.
Voor het geval van 2 steekproeven ($k = 2$) is T^2 gelijk aan de genormeerde toetsingsgrootheid U van WILCOXON [6] voor 2 steekproeven.

Indien de hypothese H_0, inhoudende dat alle waarnemingen uit dezelfde verdeling afkomstig zijn, juist is, bezit de toetsingsgrootheid T^2 voor grote waarden van $n_1, n_2, \ldots, n_k$ bij benadering een χ^2-verdeling met $\frac{k(k-1)}{2}$ vrijheidsgraden.
Indien H_0 niet vervuld is, bezit T^2 gemiddeld grotere waarden dan onder H_0.
De hypothese H_0 wordt daarom verworpen indien de gevonden waarde T^2 groter is dan een kritieke waarde T^2_α, welke gegeven wordt door

$$P[T^2 \geq T^2_\alpha \mid H_0] = P[\chi^2_{\frac{k(k-1)}{2}} \geq T^2_\alpha \mid H_0] = \alpha.$$

De onbetrouwbaarheid van de uitspraak is dan gelijk aan α.
De overschrijdingskans van de gevonden waarde T^2 wordt gegeven door

$$P[T^2 \geq T^2 \mid H_0] = P[\chi^2_{\frac{k(k-1)}{2}} \geq T^2 \mid H_0].$$

Opmerkingen

1e. Uit het voorgaande blijkt, dat de T^2-toets bewerkelijker is dan de H-toets. Dit zal in vele gevallen echter geen bezwaar zijn, daar de T^2-toets eventueel bestaande verschillen tussen de k steekproeven vermoedelijk scherper aantoont dan de H-toets, d.w.z. dat hij een groter onderscheidend vermogen bezit.

2e. Dat de grootheid T^2 voor grote waarden van n_1, $n_2, \ldots, n_k$ bij benadering een χ^2-verdeling bezit, is strikt genomen alleen bewezen voor het geval, dat alle waarnemingen verschillend zijn. Indien echter niet te veel gelijken voorkomen, kan bovenstaande methode zonder bezwaar toegepast worden.

3e. Een onderzoek naar de aard van de verdeling van T^2 voor kleine steekproeven is nog gaande. Vermoedelijk zal, evenals bij de H-toets, de exacte verdeling van T^2 goed benaderd kunnen worden met behulp van de F-verdeling.

3. Voorbeeld van de H-toets en de T^2-toets.

Gegeven zijn de volgende 5 steekproeven, elk van de uitgebreidheid 10.

1.

R_1	1	2	4	6	9	13	16	20	23	28
$R_1^{(2)}$	1	2	4	6	8	10	12	14	15	17
$R_1^{(3)}$	1	2	3	4	6	8	10	12	14	16
$R_1^{(4)}$	1	2	3	4	5	7	8	10	12	14
$R_1^{(5)}$	1	2	3	4	5	6	7	8	9	11

2.

R_2	3	5	8	11	15	19	24	29	33	36
$R_2^{(3)}$	1	2	4	6	8	10	12	14	16	18
$R_2^{(4)}$	1	2	3	4	6	8	10	12	14	16
$R_2^{(5)}$	1	2	3	4	5	6	7	9	11	12

3.

R_3	7	10	14	17	21	27	31	34	39	43
$R_3^{(4)}$	1	2	4	5	7	10	12	13	16	18
$R_3^{(5)}$	1	2	3	4	5	7	8	10	12	15

4.

R_4	12	18	22	25	30	35	38	41	44	46
$R_4^{(5)}$	1	2	3	4	6	8	10	12	14	16

5.

R_5	26	32	37	40	42	45	47	48	49	50

De steekproeven zijn op de in paragraaf 1 en 2 gegeven wijzen van verschillende rangnummers voorzien. Uit het schema vinden we:

R_i	$R_h^{(j)}$	$\tilde{U}_{h,j}$	R_i^2	$\tilde{U}^2_{h,j}$	R_i^2/n_i	$\tilde{U}^2_{h,j}/n_h n_j$
$R_1=122$			$R_1^2=14884$		1488,4	
	$R_1^{(2)}=89$	$\tilde{U}_{1,2}=-16$		256		2,56
	$R_1^{(3)}=76$	$\tilde{U}_{1,3}=-29$		841		8.41
	$R_1^{(4)}=66$	$\tilde{U}_{1,4}=-39$		1521		15,21
	$R_1^{(5)}=56$	$\tilde{U}_{1,5}=-49$		2401		24,01
$R_2=183$			$R_2^2=33489$		3348,9	
	$R_2^{(3)}=91$	$\tilde{U}_{2,3}=-14$		196		1,96
	$R_2^{(4)}=76$	$\tilde{U}_{2,4}=-29$		841		8,41
	$R_2^{(5)}=60$	$\tilde{U}_{2,5}=-45$		2025		20,25

R_i	$R_h^{(j)}$	$\tilde{U}_{h,j}$	R_i^2	$\tilde{U}_{h,j}^2$	R_i^2/n_i	$\tilde{U}_{h,j}^2/n_h n_j$
$R_3=243$			$R_3^2=59049$		5904,9	
	$R_3^{(4)}=88$	$\tilde{U}_{3,4}=-17$		289		2,89
	$R_3^{(5)}=67$	$\tilde{U}_{3,5}=-38$		1444		14,44
$R_4=311$			$R_4^2=96721$		9672,1	
	$R_4^{(5)}=76$	$\tilde{U}_{4,5}=-29$		841		8,41
$R_5=416$			$R_5^2=173056$		17305,6	

Hieruit volgt:

$$\sum_{i=1}^{5} \frac{R_i^2}{n_i} = 37719{,}90.$$

en

$$\sum_{h<j\leq 5} \frac{\tilde{U}_{h,j}^2}{n_h n_j} = 106{,}55.$$

We vinden dus

$$H = \frac{12}{50 \times 51}\, 37719{,}90 - 3 \times 51 = 177{,}51 - 153 = 24{,}51$$

en

$$T^2 = 12 \times 106{,}55 - 50 \times 24{,}51 = 53{,}10.$$

De grootheden $\underline{H}$ en $\underline{T}^2$ bezitten onder de hypothese H_0 bij benadering χ^2-verdelingen met resp. $5-1 = 4$ en $\frac{5 \times 4}{2} = 10$ graden van vrijheid.
Voor de berekende waarden van H en T^2 vinden we in een χ^2-tabel of nomogram overschrijdingskansen, welke beide kleiner zijn dan 1.10^{-4}.
De steekproeven zijn dus volgens beide toetsingsmethoden duidelijk systematisch verschillend en H_0 moet verworpen worden.

4. Literatuur.

[1] W.H.KRUSKAL, A non-parametric test for the several sample problem, Ann. Math. Stat. 23, no. 4 (1952).

[2] W.H.KRUSKAL and W.A.WALLIS, Use of ranks in one criterion analysis of variance, Journ. Am. Stat. Ass., 47 (1952), pp. 538-621.

[3] P.J.RIJKOORT, A generalization of Wilcoxon's test, Kon. Ned. Ak. v. Wet. Proc. A 55, Indagationes Mathematicae 14 (1952), pp. 394-404.

[4] T.J.TERPSTRA, A non-parametric k-sample-test and its connection with the H-test, Report S 92 (VP 2), Mathematisch Centrum, Amsterdam.

[5] H.CRAMÈR, Mathematical Methods of Statistics, 1946.

[6] F.WILCOXON, Individual comparisons by ranking methods, Biometrics Bull. 1, 80-83 (1945).

[7] H.B.MANN and D.R.WHITNEY, On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Stat. 18, 50-60 (1947).

5. Tabellen en nomogrammen van de χ^2-verdeling.

1. M.G.KENDALL, The Advanced Theory of Statistics, I, 1947, p. 444-446.
2. H.CRAMÈR, Mathematical Methods of Statistics, 1946, p. 559.
3. Statistica 1 (1946), p. 109.

6. Tabellen en nomogrammen van de F-verdeling.

1. Maxime Merrington and Catherine M.Thompson, Tables of percentage points of the inverted Beta (F) distribution, Biometrika 33 (1943), 73-88.

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
A M S T E R D A M - O.

Statistische Afdeling
S 73 (M 34)

Toets voor de hypothese dat een regressiecoëfficiënt nul is, wanneer de afwijkingen van de regressielijn normaal verdeeld zijn met gelijke spreidingen[1].

Ondersteld wordt:

1° De grootheid ξ kan zonder fouten worden waargenomen, terwijl de fouten[2] in de waarnemingen voor de grootheid η onderling onafhankelijk normaal[3] verdeeld zijn met gemiddelde 0 en onbekende, maar steeds dezelfde, spreiding σ.

2° Tussen de grootheden η en ξ bestaat een lineair verband:

$$\eta = \beta \xi + \alpha,$$

waarin de grootheden β en α onbekende parameters zijn.

3° Bij iedere waarneming x_i $(i = 1, \ldots, h)$ van ξ wordt een groepje, onderling onafhankelijke, waarnemingen $y_{i1}, \ldots, y_{in_i}$ voor η verricht.

Gevraagd wordt op grond van het waarnemingsmateriaal:

$$\begin{array}{l} (x_1,y_{11}),\ (x_1,y_{12}),\ \ldots,\ (x_1,y_{1n_1}) \\ \vdots \\ (x_h,y_{h1}),\ (x_h,y_{h2}),\ \ldots,\ (x_h,y_{hn_h}) \end{array}$$

de hypothese te toetsen dat de regressiecoëfficiënt β gelijk aan nul is:

$$H_0:\ \beta = 0.$$

Om de hypothese H_0 te toetsen, worden schattingen b, a en s^2 voor β, α en σ^2 berekend met behulp van de methode der kleinste kwadraten, welke in dit geval overeenkomt met de methode

---

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

2) Onder fouten vallen in dit geval, behalve meetfouten, ook toevallige afwijkingen van andere aard, b.v. physiologische afwijkingen.

3) De stochastische grootheid $\underline{x}$ bezit een normale verdeling met gemiddelde μ en spreiding σ, indien voor iedere a geldt:

$$P\left[\underline{x} \leq a\right] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{1}{2}\frac{(u-\mu)^2}{\sigma^2}}\, du.$$

der meest aannemelijke schattingen (Eng: method of maximum likelihood).

Dit geeft:

$$b = \frac{\sum_{i=1}^{h} n_i(x_i - \bar{x})y_i}{\sum_{i=1}^{h} n_i(x_i - \bar{x})^2},$$

$$a = \bar{y} - b\bar{x},$$

$$s^2 = \frac{\sum_{i=1}^{h} \sum_{j=1}^{n_i} (y_{ij} - bx_i - a)^2}{N - 2} =$$

$$= \frac{\sum_{i=1}^{h} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 + \sum_{i=1}^{h} n_i(y_i - bx_i - a)^2}{N - 2},$$

waarin:

$$N = \sum_{i=1}^{h} n_i; \qquad \bar{x} = \frac{\sum_{i=1}^{h} n_i x_i}{N};$$

$$\bar{y}_i = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i} \quad \text{en} \quad \bar{y} = \frac{\sum_{i=1}^{h} n_i y_i}{N}.$$

Beschouwen we niet alleen de waarnemingen y_{ij} ($j = 1, \ldots, n_i$; $i = 1, \ldots, h$), maar de verzameling van alle mogelijke waarden, die waargenomen kunnen worden, dan bezitten zij op deze verzameling een waarschijnlijkheidsverdeling, welke afhankelijk is van de (exact waargenomen) waarden $x_1, \ldots, x_h$. De schattingen b, a en s^2 zijn dan eveneens stochastisch, hetgeen door onderstreping van de letters wordt aangegeven.

De grootheden $\underline{b}$ en $\underline{s}$ zijn onderling onafhankelijk verdeeld, hetgeen b.v. door Mood [1] , pp. 292 - 294, bewezen wordt. De grootheid $\underline{b}$ bezit een normale verdeling met gemiddelde β en spreiding:

$$\frac{\sigma}{\sqrt{\sum_{i=1}^{h} n_i(x_i - \bar{x})^2}},$$

terwijl de grootheid

$$\frac{(N-2)\underline{s}^2}{\sigma^2}$$

verdeeld is volgens een χ^2-verdeling met (N-2) vrijheidsgraden. Hieruit volgt dat de stochastische grootheid

$$\underline{t} = \frac{(\underline{b} - \beta)\sqrt{\sum_{i=1}^{h} n_i(x_i - \bar{x})^2}}{\underline{s}}$$

verdeeld is volgens de verdeling van Student met (N-2) vrijheidsgraden.

Onder de hypothese H_0: $\beta = 0$ gaat deze grootheid over in

$$\underline{t} = \frac{b\sqrt{\sum_{i=1}^{h} n_i(x_i - \bar{x})^2}}{\underline{s}},$$

die dus onder de hypothese H_0 weer bovengenoemde Studentverdeling bezit.

Is $\beta \neq 0$, dus de hypothese niet juist, dan zullen verder van nul gelegen waarden van $\underline{t}$ grotere waarschijnlijkheid bezitten, dan dan wanneer $\beta = 0$. Voor toetsing van de hypothese gebruikt men daarom in het twee-zijdige geval (als zowel $\beta > 0$ als $\beta < 0$ kan optreden) een twee-zijdige kritieke zône:

$$|\underline{t}| \geq t_0,$$

waarbij t_0, behorende bij een bepaalde onbetrouwbaarheid α, opgezocht kan worden in tabellen, vermeld in onderstaande litteratuurlijst.

Indien slechts alternatieve mogelijkheden van de vorm $\beta > 0$ of $\beta < 0$ worden toegelaten, gebruikt men een eenzijdige kritieke zône:

$$\underline{t} > t_1 \text{ resp. } \underline{t} < -t_1,$$

waarin t_1 in dezelfde tabellen kan worden opgezocht als t_0, door te zoeken in een tabel voor de tweezijdige toets met onbetrouwbaarheidsdrempel 2α.

Het bij de tabellen vermelde aantal vrijheidsgraden (vaak aangegeven door ν of n-1 of n') is in dit geval gelijk aan N-2.

Opmerking: Er kan bewezen worden (zie [5]), dat de hier beschreven toets voldoet aan het, door J. Neyman ingevoerde, λ-principe (zie [6]).

Litteratuur:

[1] A.M. Mood, Introduction to the theory of statistics, Mc Graw-Hill, New York, Toronto, London, 1950, hoofdstuk 13 (regressietheorie) en p. 425 (tabel).

[2] H. Cramér, Mathematical methods of statistics, Princeton Univ. Press, Princeton, 1946.

[3] M.G. Kendall, The advanced theory of statistics I, Griffin and Co., London, 1947, pp. 440 - 441 (tabel).

[4] C.W. Emmens, Principles of biological assay, Chapman and Hall Ltd, London, 1948, pp. 18 - 19 (tabel).

[5] H.B. Mann, Analysis and design of experiments, Dover Publications, New York, 1949, pp. 41 - 42.

[6] J. Neyman, First course in probability and statistics, Henry Holt and Co., New York, 1950, pp. 338 - 343.

MATHEMATISCH CENTRUM,
2e Boerhaavestraat 49,
Amsterdam - O.
Statistische Afdeling
S 168 (M 58)

Enige in de covariantieanalyse gebruikte toetsen. [1] [2]

Wij gaan uit van k groepen stochastische grootheden

$$\begin{array}{l} y_{11}, \dots, y_{1n_1}, \\ y_{21}, \dots, y_{2n_2}, \\ \vdots \\ y_{k1}, \dots, y_{kn_k}. \end{array}$$

De verwachtingen van deze grootheden voldoen aan de volgende relaties:

$$(1) \qquad \begin{cases} \mathcal{E} y_{1j} = \alpha_1 + \beta_1 x_{1j} & (j = 1, \dots, n_1), \\ \mathcal{E} y_{kj} = \alpha_k + \beta_k x_{kj} & (j = 1, \dots, n_k). \end{cases}$$

Hierin stellen α_i en β_i $(i = 1, \dots, k)$ onbekende parameters voor, terwijl de waarden x_{ij} $(i = 1, \dots, k\,;\, j = 1, \dots n_i)$ gegeven zijn. Verder wordt ondersteld, dat alle y's onderling onafhankelijk en normaal verdeeld zijn met dezelfde spreiding σ. Hierin is dus begrepen het geval waarin x_{ij} een waarneming is van de stochastische grootheid $\underline{x}_{ij}$, waarbij y_{ij} en $\underline{x}_{ij}$ een tweedimensionale normale verdeling hebben. Onder de voorwaarde dat $\underline{x}_{ij}$ de waarde x_{ij} heeft aangenomen, is y_{ij} dan immers normaal verdeeld met als gemiddelde een lineaire functie van x_{ij}. Wij stellen nu toetsen op voor de volgende hypothesen:

(A) $\qquad \beta_1 = \dots = \beta_k$

met als toegelaten hypothesen willekeurige α's en β's.

(B) $\qquad \alpha_1 = \dots = \alpha_k$ en $\beta_1 = \dots = \beta_k$

met als toegelaten hypothesen willekeurige α's maar gelijke β's en

(C) $\qquad \alpha_1 = \dots = \alpha_k$ en $\beta_1 = \dots = \beta_k$

met als toegelaten hypothesen willekeurige α's en β's.

---

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

2) De eerste van de in dit memorandum besproken toetsen wordt ook in memorandum S 73 (M 37) behandeld met gebruik van een enigszins andere notatie.

Met (A) toetsen wij dus de evenwijdigheid van de k regressielijnen uitgaande van het model (1). Met (B) toetsen wij of de k lijnen samenvallen, aangenomen dat ze evenwijdig zijn en met (C) toetsen wij direct of wij slechts met één regressielijn te maken hebben, uitgaande van (1).

Wij voeren de volgende afkortingen in:

$$C_{xxi} = \sum_{j=1}^{n_i}\left(x_{ij} - \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}\right)^2 = \sum_{j=1}^{n_i} x_{ij}^2 - \frac{\left(\sum_{j=1}^{n_i} x_{ij}\right)^2}{n_i},$$

$$C_{yyi} = \sum_{j=1}^{n_i}\left(y_{ij} - \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}\right)^2 = \sum_{j=1}^{n_i} y_{ij}^2 - \frac{\left(\sum_{j=1}^{n_i} y_{ij}\right)^2}{n_i},$$

$$C_{xyi} = \sum_{j=1}^{n_i}\left(x_{ij} - \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}\right)\left(y_{ij} - \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}\right) =$$

$$= \sum_{j=1}^{n_i} x_{ij} y_{ij} - \frac{\sum_{j=1}^{n_i} x_{ij} \sum_{j=1}^{n_i} y_{ij}}{n_i},$$

$$C_{xxW} = \sum_{i=1}^{k} C_{xxi},$$

$$C_{yyW} = \sum_{i=1}^{k} C_{yyi},$$

$$C_{xyW} = \sum_{i=1}^{k} C_{xyW},$$

$$C_{xxT} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}^2 - \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}\right)^2}{\sum_{i=1}^{k} n_i},$$

$$C_{yyT} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{\left(\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}\right)^2}{\sum_{i=1}^{k} n_i},$$

$$C_{xyT} = \sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij} y_{ij} - \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij} \sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij}}{\sum_{i=1}^{k} n_i}.$$

De som van de kwadraten van de afwijkingen van de y waarden van de geschatte regressielijnen binnen de afzonderlijke groepen, gesommeerd over de groepen is

$$\underline{S}_1 = \sum_{i=1}^{k} \left(\underline{C}_{yyi} - \frac{\underline{C}_{xyi}^2}{C_{xxi}} \right).$$

Als het model (1) geldt, heeft $\underline{S}_1/\sigma^2$ een χ^2- verdeling met $\sum_{i=1}^{n_i} - 2k$ vrijheidsgraden.

Wij kunnen ook zo goed mogelijk evenwijdige regressielijnen in de verschillende groepen schatten. De som van de kwadraten van de afwijkingen van deze lijnen is

$$\underline{S}_2 = \sum_{i=1}^{k} \underline{C}_{yyi} - \frac{\underline{C}_{xyW}^2}{C_{xxW}}.$$

Deze som gedeeld door σ^2 heeft een χ^2-verdeling met $\sum_{i=1}^{n_i} - k - 1$ vrijheidsgraden als hypothese (B) of (C) waar is.

Als hypothese (A) geldt, heeft de som van kwadraten van afwijkingen van één lijn

$$\underline{S}_3 = \underline{C}_{yyT} - \frac{\underline{C}_{xyT}^2}{C_{xxT}},$$

gedeeld door σ^2, een χ^2-verdeling met $\sum_{i=1}^{k} n_i - 2$ vrijheidsgraden.

Voor de hypothese (A) gebruiken wij de toetsingsgrootheid

$$\underline{F}_A = \frac{\sum_{i=1}^{k} n_i - 2k}{k-1} \cdot \frac{\underline{S}_2 - \underline{S}_1}{\underline{S}_1},$$

die onder de hypothese (A) een F-verdeling bezit met $k-1$ en $\sum_{i=1}^{k} n_i - 2k$ vrijheidsgraden.

Als toetsingsgrootheid voor de hypothese (B) dient

$$\underline{F}_B = \frac{\sum_{i=1}^{k} n_i - k - 1}{k-1} \cdot \frac{\underline{S}_3 - \underline{S}_2}{\underline{S}_2},$$

welke grootheid een F-verdeling met $k-1$ en $\sum_{i=1}^{k} n_i - k - 1$ vrijheidsgraden heeft, als (B) waar is.

Hypothese (C) toetsen wij tenslotte met

$$\underline{F}_C = \frac{\sum_{i=1}^{k} n_i - 2k}{2(k-1)} \cdot \frac{\underline{S}_3 - \underline{S}_1}{\underline{S}_1},$$

een F verdeelde grootheid met $2(k-1)$ en $\sum_{i=1}^{k} n_i - 2k$ vrijheidsgraden, onder (C).

Als de betreffende hypothesen niet vervuld zijn, hebben de bijbehorende F's geen F-verdeling, maar zijn grotere F waarden meer waarschijnlijk. Als kritieke zône gebruikt men daarom $F \geqq F_0$, waarin F_0 behorende bij een bepaalde onbetrouwbaarheidsdrempel ε opgezocht kan worden in tabellen van de F-verdeling met het juiste aantal vrijheidsgraden.

Literatuur over de achtergrond van bovengenoemde toetsen kan men vinden in MANN [4] en KENDALL [3]. Het rekenschema is grotendeels ontleend aan DIXON en MASSEY [1]. Al deze boeken bevatten tabellen van de F-verdeling. Deze verdeling is het uitvoerigst getabelleerd in FISHER en YATES [2] .

Literatuur

[1] W.J. DIXON en F.J. MASSEY, Introduction to Statistical Analysis, Mc Graw-Hill Book Company, Inc., New York, Toronto, London, 1951, Chapter 12.

[2] R.A. FISHER en F. YATES, Statistical Tables for Biological, Agricultural and Medical research, 3d ed., Oliver & Boyd, London 1949, Table V.

[3] M.G. KENDALL, The advanced theory of statistics, Vol II, 2nd ed., Griffin, London, 1948, pp 237-246.

[4] H.B. MANN, Analysis and Design of Experiments, Dover Publication, Inc., New York 1949.
