

Enige Rangcorrelatieschema's *)

door Ph. van Elteren **) en W. F. van Peype ***)

Summary

In this paper the following rank correlation methods are discussed especially with respect to their application to psychological research:

1. *The method of m rankings (section 2),*
2. *Kendall's method of paired comparisons (section 3),*
3. *Durbin's method for balanced incomplete block designs (section 4),*
4. *A test for concordance in a particular incomplete design (section 5).*

For the last case the distribution of the test statistic under the hypothesis tested has been investigated using the results of a sampling experiment. Mean and variance of this distribution are derived in an appendix by H. Kesten and J. Th. Runnenburg.

1. Inleiding

In dit artikel worden rangcorrelatieschema's besproken welke kunnen worden gebruikt bij bepaalde psychologische onderzoeken. Wij denken hierbij aan experimenten, waarbij men een proefpersoon een aantal objecten volgens een bepaald criterium laat rangschikken. De objecten zijn dan bijvoorbeeld kaartjes, waarop een karaktertrek vermeld staat; het criterium de mate, waarin de proefpersoon die trekken voor zichzelf typerend acht. Of er staan op de kaartjes eigenschappen, die — volgens een tevoren verricht onderzoek — in meer of mindere mate van belang zijn voor het goed vervullen van de functie, die de proefpersoon ambieert; het criterium hun relatieve belangrijkheid volgens de proefpersoon. Uit het resultaat kunnen — na bewerking van een voldoende groot materiaal — eventueel voor bepaalde groepen personen typerende rangschikkingen geconstrueerd worden en aan de hand daarvan weer bepaalde conclusies worden getrokken betreffende de proefpersonen.

De betrouwbaarheid van dergelijke tests hangt af van de mate, waarin de verkregen volgorde consistent is bij herhaling van het experiment onder enigszins gewijzigde omstandigheden.

Om dit na te gaan kan men de proefpersoon bijvoorbeeld verschillende reeksen kaartjes voorleggen waarop dezelfde karaktertrekken maar met gewijzigde formuleringen vermeld zijn en hem dan ieder van deze reeksen laten

*) Rapport SP 47 van de Statistische Afdeling van het Mathematisch Centrum; dit rapport kwam tot stand naar aanleiding van onderzoeken van W. F. van Peype.

**) Medewerker van het Mathematisch Centrum, Amsterdam.

***) Nederlandse Stichting voor Psychotechniek, Utrecht.

rangschikken. Het resultaat wordt weergegeven in een schema waarbij de kolommen corresponderen met de vergeleken karaktertrekken en de rijen met de verschillende reeksen kaartjes. In dit artikel zullen wij, om in overeenstemming te komen met de gebruikelijke statistische terminologie, het woord object gebruiken voor het gemeenschappelijke element, waarop de waarnemingen in één kolom betrekking hebben (in het geval van ons voorbeeld zijn de karaktertrekken dus de objecten en niet de kaartjes.) Het schema wordt gevuld met de rangnummers, die aan de objecten in ieder der rangschikkingen worden toegekend; wij zullen een dergelijk schema een rangcorrelatieschema noemen. Wij zullen in dit artikel methoden bespreken, waarmee de consistentie van rangcorrelatieschema's kan worden onderzocht.

Bij de uitwerking van de te bespreken rangcorrelatiemethoden doen zich complicaties voor, indien men toelaat, dat de proefpersoon in twijfelgevallen in dezelfde rangschikking aan twee verschillende waarnemingen hetzelfde rangnummer toekent. Deze complicaties kunnen bij psychologische toepassingen in de regel vermeden worden door ongelijkheid der rangnummers dwingend voor te schrijven; wij zullen bij onze beschouwingen daarom van de mogelijkheid van gelijke rangnummers in dezelfde rangschikking afzien.

Indien men de proefpersoon in iedere rangschikking alle beschouwde objecten laat vergelijken, verkrijgt men een schema, dat geanalyseerd kan worden met de methode van m rangschikkingen. Wij zullen deze methode in het kort in par. 2 behandelen. Indien het aantal objecten te groot is, is deze methode bezwaarlijk, omdat de proefpersoon slechts een beperkt aantal objecten goed kan overzien; worden hem meer objecten ter vergelijking aangeboden, dan zal zijn oordeel van niet relevante factoren (b.v. van de volgorde, waarin hij de objecten beschouwt) gaan afhangen. Om aan dit bezwaar tegemoet te komen, kan men de proefpersoon steeds slechts een deelverzameling der objecten laten vergelijken. Men verkrijgt dan *onvolledige* schema's. Bij de opstelling van deze schema's tracht men alle objecten zoveel mogelijk gelijkelijk in de rangschikkingen te betrekken. Zo worden bijvoorbeeld in alle rangschikkingen evenveel objecten betrokken en ieder object in evenveel rangschikkingen.

Een eenvoudig voorbeeld van een onvolledig schema wordt gebruikt bij de methode van vergelijking in paren (Eng.: „Paired comparisons”). Hierbij worden in iedere rangschikking slechts 2 objecten met elkaar vergeleken; dit geschiedt evenveel malen met alle mogelijke paren objecten. Wij zullen deze methode in par. 3 behandelen.

Het aantal paren objecten dat men kan vormen stijgt zeer snel met het aantal objecten, zodat de methode van vergelijking in paren voor grote aantallen objecten zeer bewerkelijk wordt; men kan dit bezwaar o.a. ondervangen door meer dan 2 objecten per experiment te laten vergelijken. Indien men er

dan zorg voor draagt, dat ieder paar objecten in het schema even vaak in *dezelfde* rangschikking voorkomt, verkrijgt men een zogenaamd *evenwichtig* onvolledig schema. Op de analyse van deze schema's komen wij in par. 4 terug.

Niet bij ieder aantal objecten is een evenwichtig onvolledig schema mogelijk. Men kan dan trachten een schema op te stellen dat aan andere symmetrie-eisen voldoet. De analyse van dergelijke schema's wordt echter minder eenvoudig. Een van de schrijvers (W. F. van Peype) heeft een bepaald onvolledig schema onderzocht voor 14 objecten; een evenwichtig schema was hier niet mogelijk. Dit onderzoek werd de aanleiding tot dit artikel, par. 5 is eraan gewijd. Wij hebben om de consistentie van het schema te kunnen toetsen de waarschijnlijkheidsverdeling van een consistentiemaat nader onderzocht. Dit is zowel langs theoretische weg (par. 5.3) als door middel van een steekproef-experiment (par. 5.4—5.5) geschied. Een gedeelte van de theoretische berekeningen is uitgevoerd door J. T. h. R u n n e n b u r g en H. K e s t e n. De door hen gevolgde methode wordt in een appendix bij dit artikel nader uiteengezet.

2. Methoden van m rangschikkingen

2.1. *Elementaire vorm*

Zoals wij in par. 1 gezien hebben, kan de methode van m rangschikkingen gebruikt worden om de consistentie te onderzoeken van een rangcorrelatieschema, waarin n objecten m maal worden gerangschikt. De proefpersoon kent in iedere rangschikking aan de objecten de rangnummers $1, \dots, n$ toe naar opklimmende voorkeur. In schema 2.1 geven wij een voorbeeld van 6 rangschikkingen van 4 objecten.

Schema 2.1
Methode van m rangschikkingen

Rang- schikkingen	O_1	O_2	O_3	O_4
R_1	2	1	4	3
R_2	3	2	4	1
R_3	2	1	3	4
R_4	4	1	3	2
R_5	3	2	4	1
R_6	1	2	4	3
Kolomtotaal	15	9	22	14

De consistentie van het oordeel van de proefpersoon kan men afmeten aan de mate van overeenstemming tussen de rangschikkingen; indien deze overeenstemming groot is zullen de kolomtotaal van het schema sterk uiteenlopen.

Als maat voor de spreiding van deze kolomtotalen gebruikt men nu de som S van de kwadraten van hun gereduceerde waarden, dus de som van de kwadraten van de met hun gemiddelde verminderde kolomtotalen. In het geval van schema 2.1 is het gemiddelde der kolomtotalen gelijk aan 15, dus is

$$S = 0^2 + 6^2 + 7^2 + 1^2 = 86.$$

Om nu na te gaan welke betekenis men aan een gevonden waarde van $\underline{S}$ kan toekennen, past men daarop een toets toe. Indien de proefpersoon geen consistent oordeel heeft, zullen de rangnummers die hij toekent door niet relevante factoren worden bepaald; wij toetsen nu de hypothese H_0 dat dit volkomen aselekt¹⁾ geschiedt. Deze hypothese kan als volgt nader worden gespecificeerd:

1) In iedere rij van het schema zijn alle mogelijke manieren (in dit geval $n!$ in aantal) om de rangnummers toe te kennen even waarschijnlijk.

2) De verschillende rangschikkingen zijn onderling onafhankelijk.

Wij zullen ook bij de onvolledige schema's steeds een analoge hypothese toetsen.

De waarschijnlijkheidsverdeling van $\underline{S}$ onder de hypothese H_0 is voor kleine waarden van m en n te vinden in Kendall (1948); tevens vindt men daar benaderingen voor deze verdeling voor de niet getabelleerde gevallen.

Aangezien grote waarden van $\underline{S}$ wijzen op consistentie, zullen wij de hypothese H_0 verwerpen, als S groot is. De grenswaarde hangt af van de gekozen onbetrouwbaarheidsdrempel α . Men verwerpt H_0 als de kans dat $\underline{S}$ groter is dan de gevonden waarde onder de hypothese H_0 niet groter is dan α . In het geval van schema 2.1 is deze kans volgens de tabellen van Kendall (1948)

$$P[\underline{S} \geq 86] = 0,029.$$

Bij de veel gebruikte onbetrouwbaarheidsdrempel $\alpha = 0,05$ zal men dus H_0 verwerpen en besluiten dat er systematische overeenstemming tussen de rangschikkingen is.

Schema 2.1 geeft ons verder een schatting van de voorkeursvolgorde van de proefpersoon, nl. de grootte-volgorde van de kolomtotalen; het blijkt dan dat de proefpersoon de meeste voorkeur heeft voor object 3, vervolgens voor 1, dan voor 4 en de minste voorkeur voor object 2; hij heeft overigens weinig onderscheid gemaakt tussen de objecten 1 en 4.

De hier beschreven methode is afkomstig van M. Friedman (1937); uitvoerige bijzonderheden vindt men in M. G. Kendall (1948) en in Ph. van Elteren (1951). Een voorbeeld is in Statistica behandeld door J. Hemelrijk (1950).

¹⁾ Engels: at random.

2.2. Generalisatie

Een generalisatie van de methode van m rangschikkingen wordt behandeld in par. 4. Een verdergaande generalisatie is gegeven door A. B e n a r d en P h. v a n E l t e r e n (1953). Hierbij wordt toegelaten, dat een object per rangschikking een willekeurig aantal keren wordt beoordeeld, eventueel ook 0 keer.

Bij het voorbeeld van par. 1 komt de methode hierop neer, dat men de proefpersoon reeksen kaartjes voorlegt, op ieder waarvan een karaktertrek beschreven is; het is nu echter niet nodig, dat in iedere reeks juist één kaartje voor iedere karaktertrek voorkomt.

Men kan in dergelijke schema's, onder bepaalde voorwaarden, de hypothese H_0 toetsen, die o.a inhoudt, dat in iedere reeks alle mogelijke rangschikkingen der kaartjes even waarschijnlijk zijn. Daarbij wordt dan een benaderingsmethode gebruikt, die echter niet bruikbaar is voor experimenten, waarbij in een gering aantal rangschikkingen een betrekkelijk groot aantal objecten vergeleken wordt. Wij gaan hier niet dieper op deze generalisatie in.

3. Methode van vergelijking in paren

3.1. Voorbeeld

Een karakteristiek voorbeeld van toepassing van de methode van vergelijking in paren vindt men in K e n d a l l (1948), hoofdstuk 11. Men biedt dan een hond paarsgewijze 6 verschillende voedselsoorten aan en gaat na welke van de twee aangeboden voedselsoorten hij het eerst eet. In dit geval zijn schema's waarbij meer dan twee voedselsoorten tegelijk worden aangeboden niet bruikbaar. In totaal zijn hier $\frac{1}{2} \cdot 6 \cdot 5 = 15$ verschillende paren mogelijk. Het resultaat kan worden samengevat in een preferentieschema, dat gelijkt op de schema's waarin men de einduitslagen van sportcompetities weergeeft (zie schema 3.1.1).

Indien de hond bijvoorbeeld voedselsoort O_1 geprefereerd heeft boven voedselsoort O_2 (aan te duiden als $O_1 \rightarrow O_2$), dan wordt in het vakje (1,2) (rij 1, kolom 2) een cijfer 1 genoteerd; en in vakje (2,1) (rij 2, kolom 1) een cijfer 0. In het voorbeeld van schema 3.1.1 is $O_1 \rightarrow O_2$, $O_1 \rightarrow O_3$, $O_1 \rightarrow O_5$ en $O_1 \rightarrow O_6$; echter $O_4 \rightarrow O_1$. Een rij- resp. kolomtotaal van het schema is dus gelijk aan het aantal keren dat de betreffende voedselsoort boven andere geprefereerd is, resp. het omgekeerde. De som van het rij- en kolomtotaal behorende bij eenzelfde object moet dus steeds gelijk zijn aan het totale aantal vergelijkingen, waarbij dat object betrokken was, dus hier $6 - 1 = 5$.

De rijtotalen (of de kolomtotalen) van schema 3.1.1 geven ons de volgende schatting voor de preferentievolgorde van de hond: O_1 en O_3 werden vaker geprefereerd dan O_2 , O_5 en O_6 en deze weer vaker dan O_4 . Door toetsing kan worden onderzocht of hieraan enige waarde kan worden gehecht (zie par. 3.3).

De gegevens van schema 3.1.1 kunnen ook door rangnummers worden uit-

Schema 3.1.1.

Preferenties van een hond voor 6 voedselsoorten

	O_1	O_2	O_3	O_4	O_5	O_6	Rijtotalen
O_1	×	1	1	0	1	1	4
O_2	0	×	0	1	1	0	2
O_3	0	1	×	1	1	1	4
O_4	1	0	0	×	0	0	1
O_5	0	0	0	1	×	1	2
O_6	0	1	0	1	0	×	2
Kolomtotalen	1	3	1	4	3	3	15

gedrukt in 15 rangschikkingen van telkens 2 objecten (voedselsoorten). Dit is geschied in schema 3.1.2; het geprefereerde object van een paar kreeg rangnummer 2, het andere rangnummer 1, het rangnummer van een object in een rij is hier dus 1 meer dan het aantal objecten (0 of 1) waarboven het in die rij geprefereerd wordt.

Schema 3.1.2.

Rangnummerschema afgeleid uit schema 3.1.1.

		Objecten					
		O_1	O_2	O_3	O_4	O_5	O_6
Rangschikkingen	R_1	2	1	—	—	—	—
	R_2	2	—	1	—	—	—
	R_3	1	—	—	2	—	—
	R_4	2	—	—	—	1	—
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
	.	.	.	.	.	.	.
Kolomtotalen		9	7	9	6	7	7

De som s_j der rangnummers van object O_j is 5 groter dan het aantal keren a_j dat het boven andere geprefereerd is, omdat ieder object in $6 - 1 = 5$ rangschikkingen voorkomt. In het algemeen zal bij n objecten gelden:

$$(3.1.1) \quad s_j = a_j + n - 1$$

De getallen a_j zijn de rijtotalen van schema 3.1.1. Betrekkingen van het type (3.1.1) zullen wij ook bij de andere te behandelen methoden tegenkomen.

3.2. Consistentiemaat

Men kan bij de methode van vergelijking in paren de consistentie van het oordeel van de proefpersoon, evenals bij de methode van m rangschikkingen

meten door de som S van de kwadraten der gereduceerde kolomtotalen van het rangnummerschema te bepalen. Deze is immers een maat voor de overeenstemming tussen de rijen. In schema 3.1.2, waar het gemiddelde der kolomtotalen $7\frac{1}{2}$ is, vindt men

$$S = (1\frac{1}{2})^2 + (\frac{1}{2})^2 + (1\frac{1}{2})^2 + (1\frac{1}{2})^2 + (\frac{1}{2})^2 + (\frac{1}{2})^2 = 7\frac{1}{2}.$$

De grootheid S wordt maximaal als bijvoorbeeld steeds $O_i \rightarrow O_j$ voor $i < j$. Dan geldt:

$$(3.2.1) \quad S_{max} = \left(\frac{n-1}{2}\right)^2 + \left(1 - \frac{n-1}{2}\right)^2 + \dots + \\ + \left(n-1 - \frac{n-1}{2}\right)^2 = \frac{1}{12} n (n^2 - 1).$$

In schema's waarvoor $S = S_{max}$, zijn de preferenties tussen ieder drietal objecten consistent. Indien daarin b.v. geldt $O_1 \rightarrow O_2$ en $O_2 \rightarrow O_3$, dan vindt men ook $O_1 \rightarrow O_3$. Als $S < S_{max}$ is, dan bevat het schema *inconsistente* drietallen. Dit zijn drietallen, die een cyclische preferentievolverde vertonen, b.v. $O_1 \rightarrow O_2$, $O_2 \rightarrow O_3$; $O_3 \rightarrow O_1$. Schema 3.1.1 bevat vijf dergelijke drietallen: $O_1 O_2 O_4$, $O_1 O_3 O_4$, $O_1 O_5 O_4$, $O_1 O_6 O_4$, $O_2 O_5 O_6$. Het aantal d der inconsistente drietallen hangt als volgt samen met de waarde S (vgl. K e n n d a l l (1948), hoofdstuk 11):

$$(3.2.2) \quad d = \frac{1}{2}(S_{max} - S)$$

Voor een schema met $n = 6$ objecten geldt volgens (3.1.1)

$$S_{max} = \frac{1}{2}6(6^2 - 1) = 17\frac{1}{2};$$

uit formule (3.2.2) volgt dus voor schema 3.1.1 waarin $S = 7\frac{1}{2}$ is:

$$d = \frac{1}{2}(17\frac{1}{2} - 7\frac{1}{2}) = 5$$

welk aantal wij ook door rechtstreekse telling hebben gevonden.

Uit formule (3.2.2) volgt dat het aantal inconsistente drietallen toeneemt naarmate S kleiner wordt. Hieruit blijkt in welk opzicht S een maat voor de consistentie van het schema is.

3.3. Toets voor de consistentie

Men kan bij de methode van vergelijking in paren ook de hypothese H_0

toetsen, dat de preferenties volkomen aselekt zijn. De preciese formulering van deze hypothese wordt:

1. Voor ieder paar objecten hebben de beide preferentierichtingen een kans $\frac{1}{2}$,
2. De verschillende preferenties zijn onderling onafhankelijk.

De verdeling van d onder de hypothese H_0 is voor $n \leq 7$ getabelleerd in K e n d a l l (1948), tabel 9; voor grote waarden van n worden benaderingsmethoden gegeven. Het kritieke gebied is wederom van de vorm $S \geq S_\alpha$ of $d \leq d_\alpha$ (d_α gedefinieerd als de grootste waarde waarvoor $P[d \leq d_\alpha] \leq \alpha$ is). In overeenstemming daarmee wordt de overschrijdingskans van een gevonden waarde d van $\underline{d}$ gedefinieerd door $P[\underline{d} \leq d]$. Bij schema 3.1.1 vindt men met behulp van de tabel van K e n d a l l (1948) ($n = 6$):

$$P[\underline{d} \leq 5] = 0,509.$$

Dit schema geeft dus geen aanwijzing voor consistentie.

Voor schema's, waarbij dezelfde paren objecten meer dan éénmaal vergeleken worden, verwijzen wij naar K e n d a l l (1948). Voor vele psychologische doeleinden is het echter bezwaarlijk de talrijke paren objecten te vergelijken. Er is gezocht naar methoden waarbij slechts een gedeelte van de mogelijke paren objecten in het schema wordt opgenomen. Hierover is echter nog weinig bekend (zie K e n d a l l (1954)). Daarom zullen wij ons in de volgende paragrafen beperken tot methoden waarbij weliswaar niet alle, maar toch meer dan 2 objecten tegelijk gerangschikt worden.

4. Evenwichtige onvolledige schema's

4.1. Definitie, voorbeeld

Een evenwichtig onvolledig rangcorrelatieschema bezit de volgende eigenschappen:

1. In iedere rangschikking worden evenveel objecten vergeleken.
2. Alle objecten worden even vaak beschouwd.
3. Alle paren objecten komen even vaak in eenzelfde rangschikking voor.

Wij gebruiken de volgende notatie:

m = aantal rangschikkingen

n = aantal objecten

k = aantal objecten per rangschikking

l = aantal malen dat ieder object wordt gerangschikt

λ = aantal malen dat eenzelfde paar objecten in verschillende rangschikkingen voorkomt.

Als voorbeeld geven wij schema 4.1.1 met $m = n = 13$, $k = l = 4$, $\lambda = 1$. In het schema zijn de rangnummers van de objecten bij rangschikking naar toenemende preferentie vermeld.

Schema 4.4.1.

Evenwichtig onvolledig rangcorrelatieschema $m = n = 13$, $k = l = 4$, $\lambda = 1$

Rangschikkingen	Objecten												
	O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}	O_{11}	O_{12}	O_{13}
R_1	1	2	—	3	—	—	—	—	—	4	—	—	—
R_2	—	1	3	—	2	—	—	—	—	—	4	—	—
R_3	—	—	3	1	—	2	—	—	—	—	—	4	—
R_4	—	—	—	1	2	—	3	—	—	—	—	—	4
R_5	4	—	—	—	1	2	—	3	—	—	—	—	—
R_6	—	1	—	—	—	3	4	—	2	—	—	—	—
R_7	—	—	1	—	—	—	4	2	—	3	—	—	—
R_8	—	—	—	1	—	—	—	2	3	—	4	—	—
R_9	—	—	—	—	1	—	—	—	3	4	—	2	—
R_{10}	—	—	—	—	—	1	—	—	—	4	3	—	1
R_{11}	1	—	—	—	—	—	2	—	—	—	4	3	—
R_{12}	—	1	—	—	—	—	—	2	—	—	—	3	4
R_{13}	2	—	1	—	—	—	—	—	3	—	—	—	4
	8	5	8	6	6	8	13	9	11	15	15	12	13

Als consistentiemaat kan ook hier S gebruikt worden. Men vindt: $S = 140$; de maximale waarde van S wordt bij een dergelijk schema:

$$(4.1.1) \quad S_{max} = \frac{1}{12} \lambda^2 n(n^2 - 1)$$

dus in dit geval: $S_{max} = 182$.

Voor schema's met $\lambda = 1$ kan men inconsistente drietallen definiëren zoals gedaan is in par. 3.2. Drietallen objecten, die tezamen in een rij voorkomen kunnen hier niet inconsistent zijn.

Niettemin bestaat tussen d en S ook hier de relatie (3.2.2). Dit kan onmiddellijk worden afgeleid uit het bewijs van Kendall (1948), Hoofdstuk 11, omdat ook hier S gelijk is aan de som van de kwadraten der gereduceerde rijtotalen van een preferentieschema van het type 3.1.1 dat men uit een schema van het type 4.1.1 kan afleiden. In ons voorbeeld vindt men:

$$d = \frac{1}{2}(182 - 140) = 21.$$

De verdelingen van $\underline{d}$ en $\underline{S}$ zijn echter niet dezelfde als bij paarsgewijze vergelijking.

4.2. Toetsing

J. Durbin (1951) heeft twee benaderingsmethoden gegeven om H_0 te toetsen met behulp van een evenwichtig onvolledig rangcorrelatieschema.

De verdeling van $\underline{S}$ onder H_0 kan alleen voor kleine schema's exact berekend worden. Durbin stelt voor om de verdeling van $\underline{W} = \underline{S}/S_{max}$ te benaderen

met een Bêta-verdeling, die hetzelfde gemiddelde en dezelfde variantie heeft als $\underline{W}$. Dit geschiedt naar analogie van een dergelijke benadering bij de methode van m rangschikkingen. De Bêta-verdeling is getabelleerd door K. Pearson (1934); de parameters p en q die in deze tabel voorkomen worden hier:

$$(4.2.1) \quad p = \frac{nl \left(1 - \frac{k+1}{\lambda(n+1)} \right)}{2 \left(\frac{nl}{n-1} - \frac{k}{k-1} \right)} - \frac{k+1}{\lambda(n+1)}$$

en:

$$(4.2.2) \quad q = \left(\frac{\lambda(n+1)}{k+1} - 1 \right) p.$$

Wij vinden in het geval van schema 4.1.1: $p = 5,21$ en $q = 1,8$ $p = 9,39$. Verder is $W = \frac{140}{182} = \frac{10}{13} = 0,769$. Bij toepassing van de Bêta-benadering is de overschrijdingskans van deze waarde van W ongeveer 0,0004. Men kan deze overschrijdingskans ook bepalen, door over te gaan op de F -verdeling; indien immers $\underline{W}$ een Bêta-verdeling heeft met parameters p en q , heeft:

$$(4.2.3) \quad F = \frac{q}{p} \cdot \frac{\underline{W}}{1 - \underline{W}}$$

een F -verdeling met $2p$ en $2q$ vrijheidsgraden (zie bijvoorbeeld D. van Dantzig (1947—1949) Hoofdstuk VI, par. 3).

Voor grote waarden van l kan men gebruik maken van het feit dat:

$$(4.2.4) \quad \underline{\chi^2} = \frac{12 \underline{S}}{n \lambda (k+1)}$$

onder de hypothese H_0 asymptotisch voor $l \rightarrow \infty$ een χ^2 -verdeling met $n-1$ vrijheidsgraden heeft. Dit is af te leiden uit een algemene stelling van A. Benard en P. h. van Elteren (1953) (zie par. 2.2); de benadering is reeds eerder aangegeven door Durbin (1951). Aangezien l in ons voorbeeld echter slechts 4 is, verdient de andere benadering vermoedelijk de voorkeur.

Aan de eisen, die in de definitie in par. 4.1 gesteld zijn, voldoen, als bijzondere gevallen, de schema's behorende bij de methode van m rangschikkingen ($k = n$, $\lambda = l = m$) en bij de methode van vergelijking in paren ($k = 2$). Voor $3 \leq k \leq n-1$ is slechts in bepaalde gevallen een evenwichtig onvolledig schema mogelijk.

Een aantal evenwichtige onvolledige schema's vindt men in W. G. Cochran and G. M. Cox (1950), Hoofdstuk 11. Voor niet te uitgebreide schema's

is vooral $\lambda = 1$ van belang. Het blijkt dat dan voor $n \leq 20$ zulke schema's mogelijk zijn voor:

n	7	13	13	15	19
k	3	3	4	3	3
l	3	6	4	7	9

5. Een niet evenwichtig onvolledig schema

5.1. Definitie van het schema, voorbeeld

Wij behandelen in deze paragraaf een schema voor 14 objecten dat voldoet aan de eisen 1 en 2 vermeldt in par. 4.1 doch niet aan eis 3. Dit is voor $n = 14$ ook niet mogelijk. Iedere rangschikking heeft betrekking op 6 objecten, terwijl elk object 3 maal gerangschikt wordt; de objecten van ieder paar komen éénmaal tezamen in dezelfde rangschikking voor, behalve de objecten van 7 paren, die belangrijker zijn voor het onderzoek waarop dit voorbeeld betrekking heeft en die driemaal vergeleken worden. Twee objecten van een dergelijk paar noemen wij „toegevoegde” objecten. Uit schema 5.1.1 blijkt welke objecten in iedere rangschikking worden vergeleken, de paren toegevoegde objecten zijn hierin $O_1 O_2, O_3 O_4, O_5 O_6, \dots, O_{13} O_{14}$. De in het schema voorkomende cijfers zijn wederom rangnummers van de objecten naar toenemende preferentie; de betekenis der onder het schema toegevoegde grootheden s_j , $\mathbf{s}_j$, $\mathbf{u}_j$ en $\mathbf{a}_j$ zal in de volgende paragrafen nader worden toegelicht.

Schema 5.1.1.

Niet evenwichtig onvolledig schema

		Objecten													
		O_1	O_2	O_3	O_4	O_5	O_6	O_7	O_8	O_9	O_{10}	O_{11}	O_{12}	O_{13}	O_{14}
Rangschikkingen	R_1	2	3	1	4	—	—	5	6	—	—	—	—	—	—
	R_2	—	—	4	2	1	3	—	—	5	6	—	—	—	—
	R_3	—	—	—	—	1	2	5	6	—	—	4	3	—	—
	R_4	—	—	—	—	—	—	5	3	4	6	—	—	1	2
	R_5	2	3	—	—	—	—	—	—	5	4	1	6	—	—
	R_6	—	—	4	6	—	—	—	—	—	—	2	5	3	1
	R_7	3	4	—	—	1	6	—	—	—	—	—	—	5	2
s_j		7	10	9	12	3	11	15	15	14	16	7	14	9	5
$\mathbf{s}_j = s_j - 10\frac{1}{2}$		$-3\frac{1}{2}$	$-\frac{1}{2}$	$-1\frac{1}{2}$	$1\frac{1}{2}$	$-7\frac{1}{2}$	$1\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-3\frac{1}{2}$	$3\frac{1}{2}$	$-1\frac{1}{2}$	$-\frac{1}{2}$
$\mathbf{u}_j$		$-1+1$	0	0	$-1+1$	0	0	0	0	0	0	0	0	0	0
$\mathbf{a}_j = \mathbf{s}_j - \mathbf{u}_j$		$-2\frac{1}{2}$	$-1\frac{1}{2}$	$-1\frac{1}{2}$	$1\frac{1}{2}$	$-6\frac{1}{2}$	$1\frac{1}{2}$	$4\frac{1}{2}$	$4\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-3\frac{1}{2}$	$3\frac{1}{2}$	$-1\frac{1}{2}$	$-\frac{1}{2}$

5.2. Consistentiematen

Naar analogie van de andere behandelde schema's kan men hier als consistentiemaat de som S van de kwadraten van de gereduceerde kolomtotalen $\mathbf{s}_j$ kiezen, dus:

$$S = (3\frac{1}{2})^2 + (\frac{1}{2})^2 + \dots + (5\frac{1}{2})^2 = 213\frac{1}{2}.$$

Men kan echter ook het aantal d der inconsistente drietallen tellen. Daartoe is het nodig om voor 2 toegevoegde objecten, welke driemaal vergeleken worden, één preferentie-volgorde te definiëren; het ligt voor de hand hiervoor de preferentie te kiezen, die in minstens twee van de drie gevallen voorkomt; volgens deze definitie geldt in schema 5.1.1:

$$O_2 \rightarrow O_1, O_4 \rightarrow O_3, O_6 \rightarrow O_5, O_8 \rightarrow O_7, O_{10} \rightarrow O_9, O_{12} \rightarrow O_{11} \text{ en } O_{13} \rightarrow O_{14}$$

Tussen d en S bestaat nu echter niet meer een relatie van het type (3.2.2). Wel bestaat deze relatie tussen d en de som A der kwadraten van de gereduceerde rijtotalen van een preferentieschema van het type 3.1.1 dat uit schema 5.1.1 kan worden afgeleid. Het rijtotaal a_j van het dan verkregen schema is gelijk aan het aantal objecten waarboven O_j geprefereerd wordt. Het rangnummer dat object O_j in een bepaalde rangschikking heeft is 1 groter dan het aantal objecten waarboven O_j in die rangschikking wordt geprefereerd; s_j is dus 3 groter dan het aantal objecten waarboven O_j in het gehele schema wordt geprefereerd; indien O_j 3 maal boven zijn toegevoegde is geprefereerd, is deze preferentie in s_j driemaal geteld; in a_j echter slechts éénmaal; voor een dergelijk object geldt dus:

$$s_j = a_j + 5$$

Indien wij een grootheid u_j invoeren, gedefinieerd door:

$u_j = 2$ als O_j driemaal boven het toegevoegde object wordt geprefereerd
 $u_j = 1$ als dit twee- of éénmaal gebeurt en
 $u_j = 0$ als dit geen enkele maal gebeurt,
 vinden wij:

$$(5.2.1) \quad s_j = a_j + u_j + 3$$

In schema 5.1.1 is bijvoorbeeld $u_j = 2$ voor $j = 2$ en 6; $u_j = 0$ voor $j = 1$ en 5 en $u_j = 1$ voor de overige waarden van j .

Indien wij de gereduceerde waarden van s_j , a_j en u_j aangeven met $\mathbf{s}_j$, $\mathbf{a}_j$ en $\mathbf{u}_j$ ($\mathbf{s}_j = s_j - \frac{1}{14} \sum_i s_i = s_j - 10\frac{1}{2}$ enz.), vinden wij:

$$(5.2.2) \quad \mathbf{s}_j = \mathbf{a}_j + \mathbf{u}_j \quad \text{of} \quad \mathbf{a}_j = \mathbf{s}_j - \mathbf{u}_j.$$

Voor twee toegevoegde objecten O_i en O_j geldt steeds:

$$(5.2.3) \quad u_i + u_j = 2$$

en hieruit volgt, dat $\frac{1}{14} \sum_i u_i = 1$ is, dus $\mathbf{u}_j = u_j - 1$. Men kan nu door toepassing van formule (5.2.3) de gereduceerde rijtotalen $\mathbf{a}_j$ van het preferentie-

schema direct uit de gereduceerde kolomtotalen s_j van het rangnummer-schema afleiden; dit is in de laatste regel van schema 5.1.1 gedaan. De som A van de kwadraten der a_j wordt nu:

$$A = (2\frac{1}{2})^2 + (1\frac{1}{2})^2 + \dots + (5\frac{1}{2})^2 = 195\frac{1}{2}.$$

A wordt maximaal, als de grootheden a_j een rekenkundige rij met verschil 1 vormen; dit is bijvoorbeeld het geval indien in schema 5.1.1 van links naar rechts gaande in iedere rij de rangnummers in de volgorde 1, 2, 3, 4, 5, 6 staan; men vindt dan:

$$A_{max} = 2 \{ (6\frac{1}{2})^2 + (5\frac{1}{2})^2 + \dots + (\frac{1}{2})^2 \} = 227\frac{1}{2}$$

zodat men tenslotte voor d in schema 5.1.1 vindt:

$$(5.2.5) \quad d = \frac{1}{2}(A_{max} - A) = \frac{1}{2}(227\frac{1}{2} - 195\frac{1}{2}) = 16.$$

De grootte d is maximaal als $A = \sum_{j=1}^{14} a_j^2$ minimaal is. De absolute waarde van a_j is minstens $\frac{1}{2}$ en het is mogelijk het schema zodanig te construeren, dat de absolute waarden van alle a_j gelijk zijn aan $\frac{1}{2}$. Dit bereikt men door bijvoorbeeld in R_1 de rangnummers 1, 5, 6, 2, 3, 4 in deze volgorde toe te kennen aan de objecten $O_1, O_2, O_3, O_4, O_7, O_8$ in R_2 aan $O_3, O_4, O_5, O_6, O_9, O_{10}$ en verder cyclisch (men begint in R_i bij het object O_{2i-1}). Dan geldt voor de objecten met even nummers $s_j = \frac{1}{2}$ en voor de objecten met oneven nummers $s_j = -\frac{1}{2}$, terwijl alle $u_j = 0$ zijn. De minimale waarde van A wordt dus:

$$A_{min} = 14(\frac{1}{2})^2 = 3\frac{1}{2},$$

zodat

$$(5.2.6) \quad d_{max} = \frac{1}{2}(A_{max} - A_{min}) = \frac{1}{2}(227\frac{1}{2} - 3\frac{1}{2}) = 112.$$

In de volgende paragrafen zullen wij nu nagaan in hoeverre het mogelijk is om op d als toetsingsgrootte een consistentietoets voor het schema te baseren.

5.3. Verwachting en spreiding van $\underline{d}$

Om een toets te kunnen baseren op de grootte $\underline{d}$ is het nodig de verdeling van $\underline{d}$ onder de hypothese H_0 te onderzoeken. Gezien het grote aantal mogelijkheden is het ondoenlijk de verdeling van $\underline{d}$ onder H_0 exact te berekenen. Wel kan men de verwachting en de spreiding van $\underline{d}$ bepalen. Voor de afleiding hiervan verwijzen wij naar de appendix van dit artikel, geschreven door H. Kesten en J. T. h. R u n n e n b u r g. Men vindt: $E\underline{d} = 63$ en $\sigma^2(\underline{d}) = 216,98$, dus: $\sigma(\underline{d}) = 14,73$.

5.4. Steekproefexperiment

Om de vorm van de verdeling van $\underline{d}$ te onderzoeken is een steekproefexperi-

ment verricht. Hierbij werden door loting 200 schema's van het type 5.1.1 vervaardigd, d.w.z. men bepaalt de rangnummers in iedere rij met behulp van een tabel van aselechte getallen (random numbers, zie bijvoorbeeld M. G. Kendall and B. Babington Smith (1946)). Bij ieder der 200 schema's werd nu d bepaald volgens de in par. 5.2 aangegeven methode. Het resultaat is weergegeven in tabel 5.4.1.

Wij kunnen dit resultaat beschouwen als een steekproef van 200 waarnemingen uit de verdeling van $\underline{d}$ onder de hypothese H_0 . Het gemiddelde van deze steekproef is:

$$d = \frac{1}{200} \sum_{i=1}^{200} d_i = \frac{1}{200} \{1,21 + 1,22 + \dots + 1,93\} = 61,8$$

en vertoont dus een redelijke overeenstemming met het gemiddelde $Ed = 63$ van de verdeling.

Tabel 5.4.1
Resultaat steekproefexperiment

d	Aant.	d	Aant.	d	Aant.	d	Aant.	d	Aant.	d	Aant.	d	Aant.	d	Aant.	d	Aant.
21	1	31	3	41	3	51	1	61	7	71	6	81	3	91	1		
22	1	32	1	42	5	52	9	62	3	72	4	82	4	92	0		
23	1	33	0	43	3	53	5	63	10	73	2	83	2	93	1		
24	0	34	2	44	1	54	3	64	7	74	2	84	2	94	0		
25	0	35	2	45	0	55	4	65	5	75	5	85	2	95	0		
26	1	36	1	46	4	56	4	66	6	76	6	86	3	96	0		
27	0	37	1	47	3	57	1	67	3	77	4	87	0	97	0		
28	0	38	1	48	3	58	2	68	4	78	4	88	3	98	0		
29	1	39	0	49	6	59	2	69	6	79	4	89	0	99	0		
30	0	40	0	50	4	60	3	70	5	80	3	90	1	100	0		

De variantie van de steekproefwaarden wordt

$$s'^2 = \frac{1}{199} \sum_{i=1}^{200} (d_i - \bar{d})^2 = 236.$$

Ook deze waarde komt vrij goed overeen met de variantie $\sigma^2(\underline{d}) = 216,98$ van de verdeling.

In de volgende paragraaf behandelen wij enige mogelijkheden om aan de experimentele frequentieverdeling van $\underline{d}$ een waarschijnlijkheidsverdeling aan te passen, daarbij gebruik makend van de theoretische resultaten.

5.5. Aanpassing van een waarschijnlijkheidsverdeling

De volgende mogelijkheden zijn door ons onderzocht:

a) *Een normale verdeling*

Als gemiddelde en spreiding van de aangepaste verdeling zijn de theoretische berekende waarden: $E\bar{d} = 63$ en $\sigma(\bar{d}) = 14,73$ gekozen. Hierbij werd de χ^2 -toets voor aanpassing (zie bijvoorbeeld P. G. H o e l (1946), 6e druk (1951), hoofdstuk 10) toegepast met de volgende klasseindeling:

$$d \leq 35; 36 \leq d \leq 40 \text{ enz. t/m } 86 \leq d \leq 90; d \geq 91,$$

dit zijn in totaal 13 klassen. Voor de berekening van de klassefrequenties der aangepaste verdeling werd een continuïteitscorrectie ter grootte van een halve eenheid in d toegepast. Men vindt $\chi^2 = 23,7$ bij 12 vrijheidsgraden, de overschrijdingskans hiervan is iets groter dan 0,02.

Wij hebben ook de aanpassing van een normale verdeling met gemiddelde $\bar{d} = 61,8$ en spreiding $s' = 15,36$ geschat uit de steekproef onderzocht (zie in fig. 5.5.1 de kromme $N(\bar{d}, s')$). Men vindt dan bij dezelfde klasseindeling als boven $\chi^2 = 20,3$ hier met 10 vrijheidsgraden omdat twee parameters ontleend zijn aan de steekproef. De overschrijdingskans is hier iets kleiner dan 0,03.

Beide resultaten leiden tot de conclusie dat de verdeling van $\bar{d}$ onder de hypothese H_0 niet goed door een normale verdeling wordt benaderd.

b) *Een Pearsonkromme* (zie bijvoorbeeld D. v a n D a n t z i g (1947—1949), Hoofdstuk III, 3)).

Voor de aanpassing van een Pearson-kromme heeft men 4 momenten van de verdeling nodig. Wij kennen alléén de verwachting en de spreiding van $\bar{d}$ exact; het derde en vierde moment worden geschat uit de steekproef; men heeft daartoe meerdere mogelijkheden, omdat men al dan niet gebruik kan maken van het feit, dat $E\bar{d}$ en $\sigma(\bar{d})$ exact bekend zijn. Wij geven hieronder een drietal schattingen van het gereduceerde 3e moment $\mu_3 = E(\bar{d} - E\bar{d})^3$ van $\bar{d}$. Het zijn *zuivere* schattingen, waarbij dus de verwachting gelijk is aan μ_3 :

$$1e \frac{200}{199 \cdot 198} \sum_{i=1}^{200} (d_i - \bar{d})^3 = -1309 \text{ (zuiver experimenteel),}$$

$$2e \frac{1}{200} \sum_{i=1}^{200} (d_i - E\bar{d})^3 = -2137 \text{ (met } E\bar{d} = 63 \text{ in plaats van } \bar{d}),$$

$$3e \frac{1}{200} \sum_{i=1}^{200} d_i^3 - 3E\bar{d}\sigma^2(\bar{d}) - (E\bar{d})^3 = -12787 \text{ (met } E\bar{d} \text{ resp. } \sigma^2(\bar{d}) \text{ in plaats van } \bar{d} \text{ resp. } s').$$

Deze schattingen lopen dermate sterk uiteen, dat het weinig zin heeft op één ervan de aanpassing van een Pearson-kromme te baseren.

c) *Een χ^2 -verdeling*

K e n d a l l (1948) geeft in hoofdstuk 12 een χ^2 -benadering voor de ver-

deling van $\underline{d}$ bij de methode van vergelijking in paren (zie par. 3). De benadering berust op een aanpassing met 3 momenten (gemiddelde, variantie, 3e moment). Gezien de boven geconstateerde verschillen tussen de voor het derde moment in aanmerking komende schattingen is deze methode hier niet bruikbaar.

d) *Een Bêta-verdeling*

Hierbij maken wij gebruik van de volgende parameters van de verdeling van $\underline{d}$, die wij in de voorafgaande paragrafen hebben berekend:

$$d_{min} = 0; d_{max} = 112; E\underline{d} = 63; \sigma^2(\underline{d}) = 216,98.$$

Bij deze aanpassing wordt dus geen gebruik gemaakt van de experimenteel gevonden momenten. De variabele $\underline{x} = \frac{\underline{d}}{112}$ varieert tussen 0 en 1 en heeft een verwachting $E\underline{x} = \frac{63}{112} = 0,562$ en variantie $\sigma^2(\underline{x}) = \frac{216,98}{112^2} = 0,0173$.

Een stochastische variabele die een Bêta-verdeling heeft met parameters p en q varieert eveneens tussen 0 en 1 en heeft een verwachting $\frac{p}{p+q}$ en variantie $\frac{pq}{(p+q)^2(p+q+1)}$.

Als benadering voor de verdeling van $\underline{x}$ komt dus een Bêta-verdeling in aanmerking, waarvan de parameters p en q voldoen aan:

$$\frac{p}{p+q} = 0,562 \text{ en } \frac{pq}{(p+q)^2(p+q+1)} = 0,0173.$$

Hieruit volgt: $p = 7,44$ en $q = 5,79$

Nu bevat de tabel van P e a r s o n (1934) alleen de Bêta-verdelingen met parameterwaarden die een geheel veelvoud van $\frac{1}{2}$ zijn; om een hinderlijke interpolatie (in drie dimensies) te vermijden hebben wij de experimentele frequentie-verdeling van $\underline{x}$ (af te leiden uit tabel 5.4.1) vergeleken met een drietal Bêta-verdelingen met „naburige” parameterwaarden. Wij hebben wederom de χ^2 -toets voor aanpassing toegepast. De klasseindeling was hetzelfde als bij de normale verdeling met dien verstande, dat hier de twee hoogste klassen zijn samengenomen. Ook hier is een continuïteitscorrectie ter grootte van $\frac{1}{2}$ toegepast. De resultaten vindt men in tabel 5.5.1; gemiddelde en variantie van de aangepaste Bêta-verdelingen zijn ter vergelijking met $E\underline{d}$ en $\sigma^2(\underline{d})$ in de tabel opgenomen.

Tabel 5.5.1

	Verdeling van $\underline{d}$	Aangepaste Bêta-verdelingen		
p	7,44	7,5	7	7,5
E	5,79	6	6	5,5
μ	0,562	0,556	0,538	0,577
σ	0,0173	0,0170	0,0178	0,0174
χ^2	—	17,7	18,3	25,9
k	—	0,09	0,076	0,007

$$\mu = E \frac{\underline{d}}{112} \text{ resp. } \frac{p}{p+q}; \sigma^2 = \sigma^2 \left(\frac{\underline{d}}{112} \right) \text{ resp. } \frac{pq}{(p+q)^2(p+q+1)};$$

k = overschrijdingskans van χ^2 .

Als men de overschrijdingskans k als criterium gebruikt, vertoont de $B(7,5; 6)$ -verdeling (de Bêta-verdeling met parameters $p = 7,5$ en $q = 6$) de beste aanpassing.

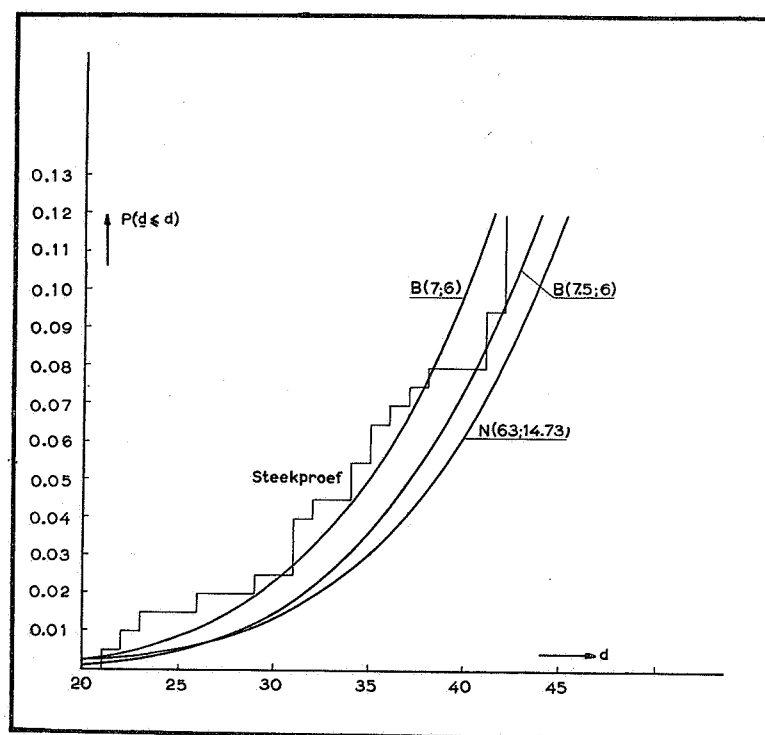
Als men voor de aanpassing van de Bêta-verdeling uitgaat van de uit de steekproef geschatte momenten $\bar{d} = 61,8$ en $s'^2 = 235,97$, vindt men: $p = 6,70$ en $q = 5,44$.

Kiest men nu als benadering de $B(6,5; 5,5)$ -verdeling, dan vindt men bij de bovenvermelde klasseindeling $\chi^2 = 16,8$; de overschrijdingskans hiervan is, als men rekening houdt met het feit, dat twee parameters zijn aangepast (dus hier slechts 9 vrijheidsgraden) $k \approx 0,05$; gebruikt men de $(7; 5,5)$ -verdeling als benadering, dan vindt men $\chi^2 = 17,1$ met $k \approx 0,05$. Ook langs deze weg wordt dus geen betere benadering verkregen.

Om na te gaan welke benadering in de praktijk het beste voldoet, hebben wij een diagram gemaakt van de kansen $P[\underline{d} \leq d]$ geschat uit de steekproef en volgens enige der behandelde aangepaste verdelingen. Wij hebben hiertoe alleen de krommen genomen die aangepast zijn aan de theoretisch berekende momenten $E\underline{d} = 63$ en $\sigma(\underline{d}) = 14,73$, aangezien deze in de praktijk gebruikt zouden worden. Krommen aangepast aan de momenten, die uit de steekproef geschat zijn, zullen in de regel een geflatteerd goede overeenkomst met de steekproef vertonen.

Het diagram is gegeven in fig. 5.5.1; men vindt hierin de aangepaste normale verdeling ($N(63; 14,73)$) en de aangepaste Bêta-verdelingen ($B(7,5; 6)$ en $B(7; 6)$). In het diagram zijn alleen kleine waarden van $\underline{d}$ opgenomen. Deze waarden zijn voor de praktijk het belangrijkste, omdat de kleine waarden van $\underline{d}$ het critieke gebied vormen. In dit gebied blijkt de $B(7;6)$ -verdeling het beste te voldoen, hoewel volgens het χ^2 -criterium de $B(7,5; 6)$ -verdeling de voorkeur verdiende. Deze sluit blijkbaar voor de grote waarden van $\underline{d}$ beter

aan. Vergeleken met de resultaten van de steekproef schijnt ook de $B(7; 6)$ -verdeling de overschrijdingskansen voor lage waarden van d nog te onderschatten. Het verdient dus aanbeveling om als critieke waarde bij een onbetrouwbaarheidsdrempel 0,05 hoogstens $d = 35$ te kiezen volgens de $B(7,6)$ -benadering. De waarde $d = 16$ die wij in ons voorbeeld vonden (zie (5.2.5)) wijst dus in ieder geval op een duidelijke consistentie.



Figuur 5.5.1.

Verdelingsfunctie van d volgens steekproef en enige aangepaste verdelingen ($d \leq 40$).

6. Conclusies

Wij hebben in par. 5 van dit artikel een rangcorrelatieschema besproken, dat voortgekomen was uit een compromis tussen de praktische mogelijkheden van bepaalde experimenten en de voorwaarden, die gesteld moeten worden om een statistische analyse mogelijk te maken. Het vereiste zeer veel werk om een, nog niet geheel bevredigende, benadering voor de verdeling van de voor dit schema voorgestelde toetsingsgrootte op te stellen.

Hieruit blijkt eens te meer, dat het aanbeveling verdient de proefopzet, indien enigszins mogelijk te doen aansluiten bij één der verder ontwikkelde statistische methoden, zoals deze in de paragrafen 2, 3 en 4 besproken zijn.

Het is in dit verband wenselijk dat men het aantal te vergelijken objecten zo klein mogelijk houdt.

Opmerking

Wij willen er hierbij nog op wijzen, dat in het geval van het in par. 1 beschreven voorbeeld, schema's, waarbij men ieder paar objecten slechts éénmaal rechtstreeks vergelijkt (methode van vergelijking in paren, evenwichtig onvolledig schema met $\lambda = 1$) aanbeveling verdienen indien men slechts over één serie kaartjes (met één omschrijving per object) beschikt. Bij herhaalde rechtstreekse vergelijking van dezelfde twee objecten (zoals dat voorgeschreven is bij de methode van m rangschikkingen en bij de evenwichtige onvolledige schema's met $\lambda > 1$) bestaat de mogelijkheid dat de proefpersoon de eenmaal uitgesproken preferentie doelbewust herhaalt, hetgeen dus zou betekenen, dat onafhankelijkheid der rangschikkingen verloren zou gaan. In een dergelijk geval zou het resultaat niet alleen door de consistentie van het oordeel, doch ook door het geheugen van de proefpersoon beïnvloed worden. Bij toepassing van de methode van m rangschikkingen verdient het daarom de voorkeur de proefpersoon m verschillende series kaartjes (met omschrijvingen van dezelfde n objecten!) voor te leggen. Ook bij onvolledige schema's is het beter om bij iedere rangschikking met een afzonderlijke serie te werken.

Literatuur:

- A. Benard and Ph. van Elteren (1953), A generalization of the method of m rankings, Proc. Kon. Ned. Ak. v. Wet. A 56 (1953), p. 358—369. *Indagationes Mathematicae* 15, (1953) p. 358—369.
- G. Cochran and G. M. Cox (1950), *Experimental Designs*, Wiley and Sons, New York, Chapman and Hall, London.
- D. van Dantzig (1947—1949), *Kadercursus Mathematische Statistiek*, Mathematisch Centrum, Amsterdam.
- J. Durbin (1951), Incomplete blocks in ranking experiments, *The British Journal of Psychology, Statistical Section* 4 (1951), p. 85—90.
- Ph. van Elteren (1951), *Methode van m rangschikkingen*, Mathematisch Centrum, Rapport S 59, Hoofdstuk II.
- M. Friedman (1937), The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *Journ. of the American Statistical Association* 32 (1937) p. 675—699.
- J. Hemelrijk (1950), Rangcorrelatie en de schattingsproef van Varangot, *Statistica* 4 (1950) p. 216—225.
- P. G. Hoel (1956), *Introduction to Mathematical Statistics*, Wiley and Sons, New York, Chapman and Hall, London.
- M. G. Kendall (1948), *Rank correlation methods*, Ch. Griffin & Co. London.
- M. G. Kendall (1954), Further contributions to the theory of paired comparisons, *Biometrika* 41 (1954) p. 43—62.
- M. G. Kendall and B. Babington Smith (1946), *Tables of Random Sampling Numbers*, Cambridge University Press.
- K. Pearson (1934), *Tables of the incomplete Beta-function*, Cambridge University Press.

8222 NC