STATISTISCHE AFDELING
MATHEMATISCH CENTRUM
AMSTERDAM

Aug. ~~Juli~~ 1950.

Rapport S 42.

Door J. Hemelrijk en
Ph. van Elteren.

Over het toetsen van de onafhankelijkheid van twee 2 x 2-tabellen .

1. Inleiding.

Naar aanleiding van een onderzoek van T.N.O. - A.B.W. werd een door G.W. Snedecor, Statistical Methods, Iowa State College Press 1948, p. 202 en 203 vermelde toetsingsmethode onderzocht. Deze methode bleek afkomstig van M.S. Bartlett, Contingency table interactions, Jrn.Roy.Stat.Soc. Suppl. 2 (1935) p. 248-252. De conclusie was, dat deze toetsingsmethode fout is, terwijl anderzijds een correcte exacte toets, geschikt voor kleine aantallen, in het artikel van Bartlett beschreven is. Eén en ander wordt in de volgende paragrafen nader uiteengezet.

In het laatste gedeelte van dit rapport is een toetsingsmethode vermeld voor een hypothese, die van die van Bartlett verschilt, maar die bij dezelfde en aanverwante problemen wellicht eveneens relevant zal zijn.

2. Notaties.

Wij beschouwen een collectie, waarvan ieder element onderworpen is aan 3 alternatieven

$$A, \bar{A}, \quad B, \bar{B} \quad \text{en} \quad C, \bar{C} .$$

Het geheel is dus een categorisch systeem met 8 kenmerken:

(1).

$$\begin{array}{ll} A \wedge B \wedge C & A \wedge B \wedge \bar{C} \\ A \wedge \bar{B} \wedge C & A \wedge \bar{B} \wedge \bar{C} \\ \bar{A} \wedge B \wedge C & \bar{A} \wedge B \wedge \bar{C} \\ \bar{A} \wedge \bar{B} \wedge C & \bar{A} \wedge \bar{B} \wedge \bar{C} \end{array}$$

Dit systeem is op te vatten, als bestaande uit 2 dubbele dichotomieën van de kenmerken A en $\bar{A}$ en B en $\bar{B}$, waarbij in het ene geval ieder element ook het kenmerk C, in het andere geval het kenmerk $\bar{C}$ bezit.

(2).

Wij duiden de frequenties van deze kenmerken in de collectie aan met $p_{i,j,k}$, waarin

$i = 1$ als A geldt; $i = 0$ als $\bar{A}$ geldt
$j = 1$ " B " ; $j = 0$ als $\bar{B}$ "
$k = 1$ " C " ; $k = 0$ " $\bar{C}$ " .

(3).

Het resultaat van sommatie over i wordt aangeduid met $p_{.jk}$, dus een punt op de plaats van de eerst index. Op analoge wijze kan men sommatie over j en over k, alsmede over 2 der 3 indices aanduiden. Dus (mede omdat het kenmerkensysteem categorisch is):

$$P[B \wedge C] = p_{111} + p_{011} = p_{.11}$$

$$P[C] = p_{111} + p_{101} + p_{011} + p_{001} = p_{..1}$$

(4).

Voorwaardelijke waarschijnlijkheden duidt men aan als volgt:

$$P_C[A \wedge B] = p_{11\boxed{1}}$$

$$P_{B \wedge C}[A] = p_{1\boxed{11}}$$

(5).

Voor aantallen elementen met een bepaalde kenmerken combinatie gebruikt men op analoge wijze de notatie $n_{i,j,k}$, dus b.v.:

n_{101} = aantal elementen met kenmerk A, $\bar{B}$, C

$n_{..1} = n_{111} + n_{101} + n_{011} + n_{001}$ etc.

3. <u>De hypothese van Bartlett.</u>

Bij één dubbele dichotomie:

	A	$\bar{A}$
B	p_1	p_2
$\bar{B}$	p_3	p_4

waarin $p_1, \ldots, p_4$ de onvoorwaardelijke kansen van de verschillende kenmerkencombinaties zijn, wordt de grootheid

(6) $$\mathcal{J} = \frac{p_1}{p_2} : \frac{p_3}{p_4}$$

de <u>interactie</u> genoemd. De hypothese, die dan gewoonlijk getoetst wordt, is:

(7) $$\mathcal{J} = 1$$

Daar $p_1 = P[A \wedge B] = P_B[A] \cdot P[B]$ is en analoge uitdrukkingen voor p_2, p_3 en p_4 gelden, aequivalent met

(7*) $$P_B[A] = P_{\bar{B}}[A].$$

Noemen wij nu de interacties van de 2 x 2-tabel van de elementen, die het kenmerk C resp. het kenmerk $\bar{C}$ bezitten: $\mathcal{J}_C$ resp. $\mathcal{J}_{\bar{C}}$, dan is de door Bartlett gestelde hypothese:

$\mathcal{J}_c = \mathcal{J}_{\bar{c}}$, of, uitgedrukt in de onvoorwaardelijke waarschijnlijkheden p_{ijk} :

(8) $$\frac{p_{111}}{p_{011}} : \frac{p_{101}}{p_{001}} = \frac{p_{110}}{p_{010}} : \frac{p_{100}}{p_{000}}$$

Deze hypothese kan ook in voorwaardelijke waarschijnlijkheden worden uitgedrukt. Daar nl.

$$p_{ijk} = p_{..k}\, p_{ij\boxed{k}}$$

is, is (8) aequivalent met

(9) $$\frac{p_{11\boxed{1}}}{p_{01\boxed{1}}} : \frac{p_{10\boxed{1}}}{p_{00\boxed{1}}} = \frac{p_{11\boxed{0}}}{p_{01\boxed{0}}} : \frac{p_{10\boxed{0}}}{p_{00\boxed{0}}}$$

hetgeen op analoge wijze verder herleid kan worden tot

(10) $$\frac{p_{1\boxed{11}}}{p_{0\boxed{11}}} : \frac{p_{1\boxed{01}}}{p_{0\boxed{01}}} = \frac{p_{1\boxed{10}}}{p_{0\boxed{10}}} : \frac{p_{1\boxed{00}}}{p_{0\boxed{00}}}$$

De aequivalentie van (8), (9) en (10) heeft tot gevolg, dat de exacte toets van Bartlett niet alleen toepasbaar is, indien de randtotalen stochastisch zijn en de onvoorwaardelijke waarschijnlijkheden bestaan, maar ook indien bepaalde randtotalen gegeven zijn door de aard van het experiment, of door de experimentator gekozen zijn. Bartlett gaat nl. wel uit van de onvoorwaardelijke waarschijnlijkheden (dus van (8) als hypothese), maar zijn toets is een voorwaardelijke onder de voorwaarde, dat de randtotalen de bij het experiment gevonden waarden aannemen.

4. <u>Exacte toets van Bartlett</u>.

In punt (4) van zijn artikel schetst Bartlett een exacte toets van zijn hypothese, die voor kleine steekproeven toegepast kan worden.

Hij gaat uit van de a priori kans $p(n_{ijk})$ om in een steekproef met uitgebreidheid N , n_{111} elementen met kenmerk $A \wedge B \wedge C$, n_{101} elementen met kenmerk $A \wedge \bar{B} \wedge C$ etc. te vinden. Deze kans is

(11) $$p(n_{ijk}) = P\left[\underline{n}_{111} = n_{111}, \ldots, \underline{n}_{000} = n_{000} \,\middle|\, \sum n_{ijk} = N\right] =$$

$$= \frac{N!}{\prod\limits_{ijk} n_{ijk}!} \prod_{ijk} p_{ijk}^{n_{ijk}}$$

Deze kans heeft betrekking op een 7-dimensionale steekproefruimte. Immers iedere steekproef wordt gekarakteriseerd door 8 grootheden n_{ijk} , waartussen slechts één relatie

$$\sum_{ijk} n_{ijk} = N$$ bestaat.

Men kan opmerken, dat na substitutie van

(12) $\quad n_{ijk} = m_{ijk} + (-1)^{i+j+k}.x$

(11) overgaat in:

(13) $\quad p(n_{ijk}) = \dfrac{N!}{\prod\limits_{ijk} n_{ijk}!} \prod\limits_{ijk} p_{ijk}^{m_{ijk}}$

immers er geldt:

(14) $\quad (p_{111}\, p_{100}\, p_{010}\, p_{001})^{-x} \cdot (p_{110}\, p_{101}\, p_{011}\, p_{000})^{x} = 1$

krachtens (8), d.i. de te toetsen hypothese.

Omdat de verdeling gekarakteriseerd door (11) 7-dimensionaal is, construeert Bartlett een toets, onder de voorwaarde dat alle randtotalen vast zijn.

Hij beschouwt dus alle steekproeven waarvoor behalve $\sum\limits_{ijk} n_{ijk} = N$ geldt:

$$\left.\begin{aligned} \sum_i n_{ijk} &= n_{.jk} \\ \sum_j n_{ijk} &= n_{i.k} \\ \sum_k n_{ijk} &= n_{ij.} \end{aligned}\right\}$$

waarbij $n_{.jk}$, $n_{i.k}$ en $n_{ij.}$ constanten zijn voor $i, j, k =$ 0 en 1

(Voor deze randtotalen worden dan, bij toepassing van de toetsingsmethode, de bij het experiment gevonden randtotalen genomen).

Dit is mogelijk door behalve N nog 6 andere onafhankelijke randtotalen te geven: bv.:

(15) $\quad n_{1.1} = c_{1.1}\,;\ n_{0.1} = c_{0.1}\,;\ n_{.11} = c_{.11}$

$\qquad\ n_{1.0} = c_{1.0}\,;\ n_{.10} = c_{.10}\,;\ n_{11.} = c_{11.}$.

Door de randtotalen vast te houden, bepalen wij 6 onafhankelijke relaties tussen de n_{ijk}. Alle steekproeven met gegeven randtotalen liggen dus in een ééndimensionale lineaire deelruimte R van de oorspronkelijke steekproefruimte.

Tot R behoort een eindig aantal steekproeven, omdat alle n_{ijk} geheel en nietnegatief zijn.

Indien de c's klein zijn, blijft een zeer klein aantal mogelijkheden voor n_{111} over; bij iedere n_{111} is met behulp van de gegeven c's en N, precies één stelsel n_{ijk}

te vinden, zodat wij een ééndimensionale verdeling overhouden. Ieder punt van R is bepaald door n_{111} van de corresponderende steekproef.

(16).
Stellen wij nu $\mathcal{E}_R \underline{n}_{ijk} = a_{ijk}$, dan geldt

1. De a_{ijk} voldoen aan relaties analoog aan die vermeld onder (15)

$$a_{101} + a_{111} = c_{1.1}$$
$$a_{001} + a_{011} = c_{0.1} \qquad \text{etc.}$$

2. Iedere steekproef, die behoort tot R, kan worden gekarakteriseerd door:

(17) $$n_{ijk} = a_{ijk} + (-1)^{i+j+k} \cdot x$$

<u>Bewijs</u>: Als $n_{111} = a_{111} - x$, geldt,(wegens:

$$n_{111} + n_{110} = a_{111} + a_{110} = c_{11.})$$
$$n_{110} = a_{110} + x$$

etc.

3. Als een steekproef kan worden gekarakteriseerd door (17) behoort zij tot R.

$$n_{111} + n_{110} = a_{111} - x + a_{110} + x = a_{111} + a_{110} = c_{11.} \qquad \text{etc.}$$

Het is dus volkomen aequivalent of een steekproef behoort tot R, dan wel gekarakteriseerd kan worden door n_{ijk} die voldoen aan (17).

Omdat (17) overeenkomt met (12), kunnen we schrijven voor een punt dat behoort tot R:

(18) $$p(n_{ijk}) = \frac{N!}{\prod_{ijk} n_{ijk}!} \prod_{ijk} p_{ijk}^{a_{ijk}} = \frac{C}{\prod_{ijk} n_{ijk}!}$$

waarin C onafhankelijk van de n_{ijk} [15]. Stellen we

(19) $$p_R(n_{ijk}) = P[\underline{n}_{ijk} = n_{ijk} \mid R]$$

dan geldt:

(20) $$p_R(n_{ijk}) = \frac{p(n_{ijk})}{P[R]}$$

waarin P [R] , de kans is, dat een steekproef tot R behoort, dus een bepaalde constante. Waaruit volgt:

(21) $$p_R(n_{ijk}) = \frac{C'}{\prod_{ijk} n_{ijk}!}$$ C' onafh. van de n_{ijk}

Wij hebben reeds gezien, dat, als de gegeven aantallen klein zijn, het aantal in aanmerking komende waarden voor n_{111} , dus voor alle n_{ijk} , beperkt is. In dat geval is het practisch mogelijk de voorwaardelijke verdeling, gekarakteriseerd door (21) (C' te bepalen uit $\sum p_R = 1$), te berekenen. In dat geval bezitten we hierin dus een mogelijkheid om de hypothese (8) te toetsen.

5. χ^2-toets van Bartlett.

Voor grote steekproeven, beveelt Bartlett een χ^2-toets aan. Daarbij gaat hij er van uit, dat er tussen de verwachtingen a_{ijk} , de relatie

(22) $$a_{111}\, a_{001}\, a_{100}\, a_{010} = a_{110}\, a_{000}\, a_{101}\, a_{011}$$

bestaat.

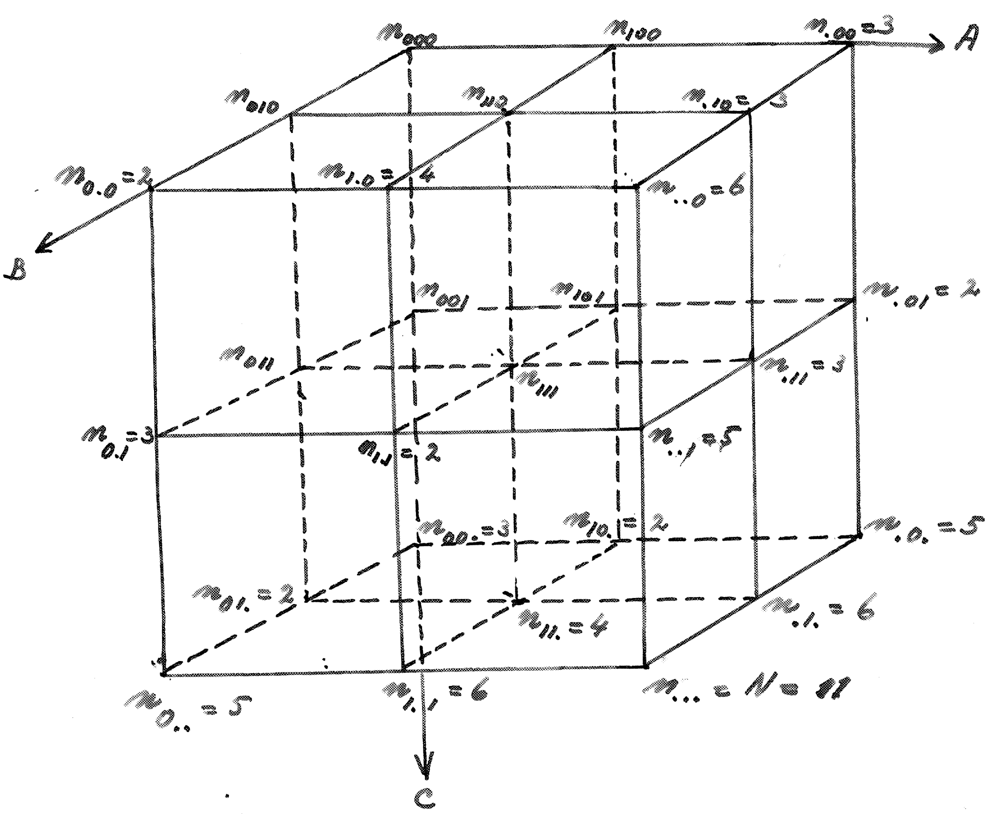
Met een tegen-voorbeeld kunnen wij gemakkelijk aantonen dat deze relatie niet juist is.

Dit tegen-voorbeeld is het volgende:

$n_{1.1} = 2$ $n_{0.1} = 3$ $n_{.11} = 3$

N = 11

$n_{1.0} = 4$ $n_{.10} = 3$ $n_{11.} = 4$.



De gegeven randtotalen laten voor n_{111} alleen de waarden 1 en 2 toe:

Er geldt als: $n_{111} = 1$:

$n_{101} = 1$, $n_{011} = 2$, $n_{001} = 1$

$n_{110} = 3$, $n_{100} = 1$, $n_{010} = 0$ $n_{000} = 2$

Dan: (zie (21))

$$P[\underline{n}_{111} = 1] = \frac{c}{1!\,1!\,2!\,1!\,3!\,1!\,0!\,2!} = \frac{c}{24}$$

als $n_{111} = 2$

$n_{101} = 0$ $n_{011} = 1$ $n_{001} = 2$

$n_{110} = 2$ $n_{100} = 2$ $n_{010} = 1$ $n_{000} = 1$

$$P[\underline{n}_{111} = 2] = \frac{c}{2!\,0!\,1!\,2!\,2!\,2!\,1!\,1!} = \frac{c}{16}$$

$\frac{c}{24} + \frac{c}{16} = 1$ dus $c = \frac{48}{5}$

$P[\underline{n}_{111} = 1] = \frac{2}{5}$ $P[\underline{n}_{111} = 2] = \frac{3}{5}$

We vinden dus:

$$a_{111} = \mathcal{E}_R\, \underline{n}_{111} = \frac{2}{5}\cdot 1 + \frac{3}{5}\cdot 2 = \frac{8}{5}$$

waarna we, gebruik makend van 1e onder (16), vinden:

$a_{101} = \frac{2}{5}$ $a_{011} = \frac{7}{5}$ $a_{001} = \frac{8}{5}$

$a_{110} = \frac{12}{5}$ $a_{100} = \frac{8}{5}$ $a_{010} = \frac{3}{5}$ $a_{000} = \frac{7}{5}$

Nu geldt: $a_{111}\, a_{001}\, a_{100}\, a_{010} = \frac{8.8.8.3}{5^4} \cancel{\frac{3000}{7^4}} = \frac{64}{5^4}$

en $a_{110}\, a_{101}\, a_{011}\, a_{000} = \cancel{\frac{10.4.11.11}{7^4}}\ \frac{12.2.7.7}{5^4} = \frac{49}{5^4}$.

Dus de relatie (22) geldt **niet** voor dit voorbeeld:

Derhalve is de χ^2-toets en de andere door Bartlett aangegeven toets, in deze vorm niet correct.

6. Toetsing van de hypothese, dat twee dubbele dichotom-ieën beschouwd kunnen worden als onafhankelijke experimenten van hetzelfde verschijnsel.

Wij beschouwen weer de twee dubbele dichotomieën van § 2: de dichotomieën zijn dus A en $\bar{A}$ en B en $\bar{B}$, terwijl de twee gevallen worden onderscheiden doordat in het éne geval alle elementen het kenmerk C en in het andere geval het kenmerk

$\bar{C}$ bezitten.

Schematisch kunnen wij dus de 2 x 2-tabellen als volgt aangeven:

C

	A	$\bar{A}$
B	n_{111}	n_{011}
$\bar{B}$	n_{101}	n_{001}

$\bar{C}$

	A	$\bar{A}$
B	n_{110}	n_{010}
$\bar{B}$	n_{100}	n_{000}

Indien nu deze twee 2 x 2-tabellen onafhankelijke experimenten omtrent hetzelfde verschijnsel zijn, zal voldaan zijn aan de volgende betrekkingen:

(23) $P_{B \wedge C}[A] = P_{B \wedge \bar{C}}[A]$

(24) $P_{\bar{B} \wedge C}[A] = P_{\bar{B} \wedge \bar{C}}[A]$

(25) $P_{A \wedge C}[B] = P_{A \wedge \bar{C}}[B]$

(26) $P_{\bar{A} \wedge C}[B] = P_{\bar{A} \wedge \bar{C}}[B]$

voorzover de genoemde voorwaardelijke waarschijnlijkheden gedefinieerd zijn.

Deze hypothesen kunnen achtereenvolgens getoetst worden door toepassing van het χ^2-criterium (of een andere methode) op de volgende 2 x 2-tabellen:

B

	A	$\bar{A}$
C	n_{111}	n_{011}
$\bar{C}$	n_{110}	n_{010}

$\bar{B}$

	A	$\bar{A}$
C	n_{101}	n_{001}
$\bar{C}$	n_{100}	n_{000}

A

	B	$\bar{B}$
C	n_{111}	n_{101}
$\bar{C}$	n_{110}	n_{100}

$\bar{A}$

	B	$\bar{B}$
C	n_{011}	n_{001}
$\bar{C}$	n_{010}	n_{000}

Op deze wijze wordt een gedifferentieerd overzicht verkregen over de (eventuele) invloed van de kenmerken C en $\bar{C}$ op de voorwaardelijke waarschijnlijkheden van A en B. Bij de interpretatie van resultaten zal men echter voorzichtig moeten zijn, daar de 4 toetsingen onderling afhankelijk zijn.

Opmerking: de hier gestelde hypothesen gaan veel verder dan de hypothese van Bartlett, daar deze laatste slechts inhoudt, dat een bepaalde dubbelverhouding van voorwaardelijke waarschijnlijkheden door de kenmerken C en $\bar{C}$ niet wordt beïnvloed. Daar de hypothese van Bartlett vervuld is, als de in deze paragraaf gestelde hypothese juist is, volgt uit verwerping van de eerstgenoemde ook die van de laatstegenoemde, echter niet andersom.