MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
STATISTISCHE AFDELING

Leiding: Prof. Dr D. van Dantzig
Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

Rapport S 115 (M 45)

Methoden voor het vergelijken, toetsen en schatten van onbekende kansen.

door

Constance van Eeden

1953

INHOUD

Inleiding.

Het toetsen van een hypothese in het algemeen.

Het toetsen van een hypothese H_0 berust op een aantal waarnemingen $x_1, x_2, \ldots\ldots, x_n$ van één of meer stochastische grootheden [1] of op enige groepen van waarnemingen.

Bij een toets behoort een toetsingsgrootheid $\underline{u}$, die een functie is van bovengenoemde waarnemingen.

Men berekent nu, onder de onderstelling dat H_0 juist is, de waarschijnlijkheidsverdeling van $\underline{u}$. Vervolgens kiest men een verzameling Z van mogelijke uitkomsten van $\underline{u}$, zodanig dat de kans dat $\underline{u}$ een in Z gelegen waarde aanneemt, onder de hypothese H_0, hoogstens gelijk is aan een gegeven getal α; Z wordt de kritieke zone genoemd en α de onbetrouwbaarheidsdrempel.

H_0 wordt nu verworpen als de bij het experiment gevonden waarde van $\underline{u}$ in Z ligt. Noemen we de kans dat dit gebeurt als H_0 juist is α', dan is α' de kans op ten onrechte verwerpen van H_0 als deze juist is; α' heet de onbetrouwbaarheid en deze is hoogstens gelijk aan α, de onbetrouwbaarheidsdrempel. De betekenis van α' en α is de volgende: α is de grootste kans op ten onrechte verwerpen van H_0, die de onderzoeker nog wil accepteren, α' de kans die hij werkelijk loopt om H_0 ten onrechte te verwerpen.

Indien de waarschijnlijkheidsverdeling van $\underline{u}$ continu is, is $\alpha' = \alpha$; is de verdeling van $\underline{u}$ discreet dan is in het algemeen $\alpha' < \alpha$.

Als uitkomst van een toets wordt vaak de overschrijdingskans k opgegeven. Dit is de kleinste α' (of α, dat maakt hiervoor geen verschil) waarvoor de gevonden waarde van $\underline{u}$ nog juist in de bij α behorende kritieke zone ligt. Werken we met een onbetrouwbaarheidsdrempel α dan wordt H_0 dus verworpen als $k \leq \alpha$ is.

De kritieke zone Z kan één- of tweezijdig gekozen worden. Tweezijdig wil zeggen: Z bestaat uit de waarden van $\underline{u}$ die ver van het gemiddelde afwijken. Rechts éénzijdig wil zeggen: Z bestaat uit de waarden van $\underline{u}$ die een grote positieve afstand tot het gemiddelde hebben.

----

1) Een stochastische grootheid is een grootheid die een waarschijnlijkheidsverdeling bezit; stochastische grootheden worden aangegeven met onderstreepte letters. Waarden aangenomen door een stochastische grootheid worden aangegeven met dezelfde letters, niet onderstreept.

De keuze van Z (één- of tweezijdig) hangt af van de alternatieve mogelijkheden, die men naast de getoetste hypothese in het onderzoek wenst te betrekken [2])

---

2) Nadere details hierover vindt men b.v. in [1]. In dit rapport volstaan wij, wat dit aspect betreft, met aanduidingen

DEEL I

1 1. De methode der dubbele dichotomie voor het vergelijken van twee kansen.

1 1.1. Inleiding

Wij beschouwen twee reeksen van onafhankelijke experimenten, waarbij ieder experiment één der mogelijke uitkomsten A of $\bar{A}$ (non-A) heeft.

Indien de eerste reeks uit n experimenten bestaat, waarvan er $\underline{a}$ de uitkomst A hebben gegeven en de tweede reeks uit m experimenten, waarvan er $\underline{b}$ de uitkomst A hebben gegeven, dan kunnen wij de resultaten als volgt samenvatten:

Tabel I

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

Hierin zijn $\underline{a}$, $\underline{b}$, $\underline{c}$, $\underline{d}$, $\underline{r}$ en $\underline{s}$ stochastisch, terwijl n en m (en dus ook $N = n + m$) gegeven getallen zijn.

Stel nu dat bij ieder experiment van de eerste reeks de kans op A gelijk is aan p (en de kans op $\bar{A}$ dus gelijk aan $1-p$) en dat bij ieder experiment van de tweede reeks de kans op A gelijk is aan p' (en de kans op $\bar{A}$ dus $1-p'$).

De hypothese die we willen toetsen luidt dan:

$$H_0 : \quad p = p' .$$

1.1.2. De toetsingsmethode

Als toetsingsgrootheid wordt hier de grootheid $\underline{a}$ gebruikt, dus het aantal malen A in de eerste reeks, en wel wordt speciaal gelet op de waarden, die $\underline{a}$ aan zou kunnen nemen bij de gegeven waarden van n en m <u>en bij de gevonden waarde van $\underline{r}$</u>.

Dit laatste is noodzakelijk, omdat men alleen op die wijze de waarschijnlijkheidsverdeling van $\underline{a}$, als H_0 juist is, kan berekenen. De grootste waarde die deze grootheid $\underline{a}$ aan kan nemen, bij gegeven n, m en r, is: $\min.(n,r)$. De kleinste waarde is $n - \min.(s,n)$. De kans dat $\underline{a}$ een waarde a aanneemt, onder de aanname dat de hypothese H_0 juist is en onder de voorwaarde, dat $\underline{r}$ de bij het experiment gevonden waarde r aanneemt, wordt gegeven door:

$$(1) \qquad P[\underline{a} = a \mid \underline{r} = r ; H_0] = \frac{\binom{n}{a} \cdot \binom{m}{b}}{\binom{N}{r}} .$$

Voor de verschillende waarden a, die $\underline{a}$ aan kan nemen, kunnen we deze kans berekenen. Hierbij kunnen wij gebruik maken van tabellen der binomiaalcoëfficiënten, die o.a. te vinden zijn in [2].

Voorbeeld 1

Ter vergelijking van twee geneesmiddelen voor een bepaalde ziekte werden 11 patiënten met het eerste en 16 patiënten met het tweede middel behandeld. Vervolgens werd bij ieder der patiënten nagegaan of de ziekte aanmerkelijk werd verlicht. De resultaten zijn samengevat in tabel II:

Tabel II

geneesmiddel	succes	geen succes	totaal
1	6	5	11
2	14	2	16
totaal	20	7	27

We hebben dus: $a = 6$, $n = 11$, $m = 16$, $r = 20$ en de waarden die $\underline{a}$ aan kan nemen bij deze waarden van n, m en r zijn: 4, 5, 6, 7, 8, 9, 10, 11.
Berekenen wij met formule (1) de waarschijnlijkheidsverdeling van $\underline{a}$ bij de gevonden waarde (20) van $\underline{r}$, dan vinden we:

Tabel III

a	$P[\underline{a} = a \mid \underline{r} = 20; H_0]$ $(n = 11, m = 16)$
4	0,0004
5	0,0083
6	0,0624
7	0,2081
8	0,3382
9	0,2705
10	0,0992
11	0,0129

In figuur 1 is horizontaal a uitgezet en verticaal $P[\underline{a} = a \mid \underline{r} = 20; H_0]$:

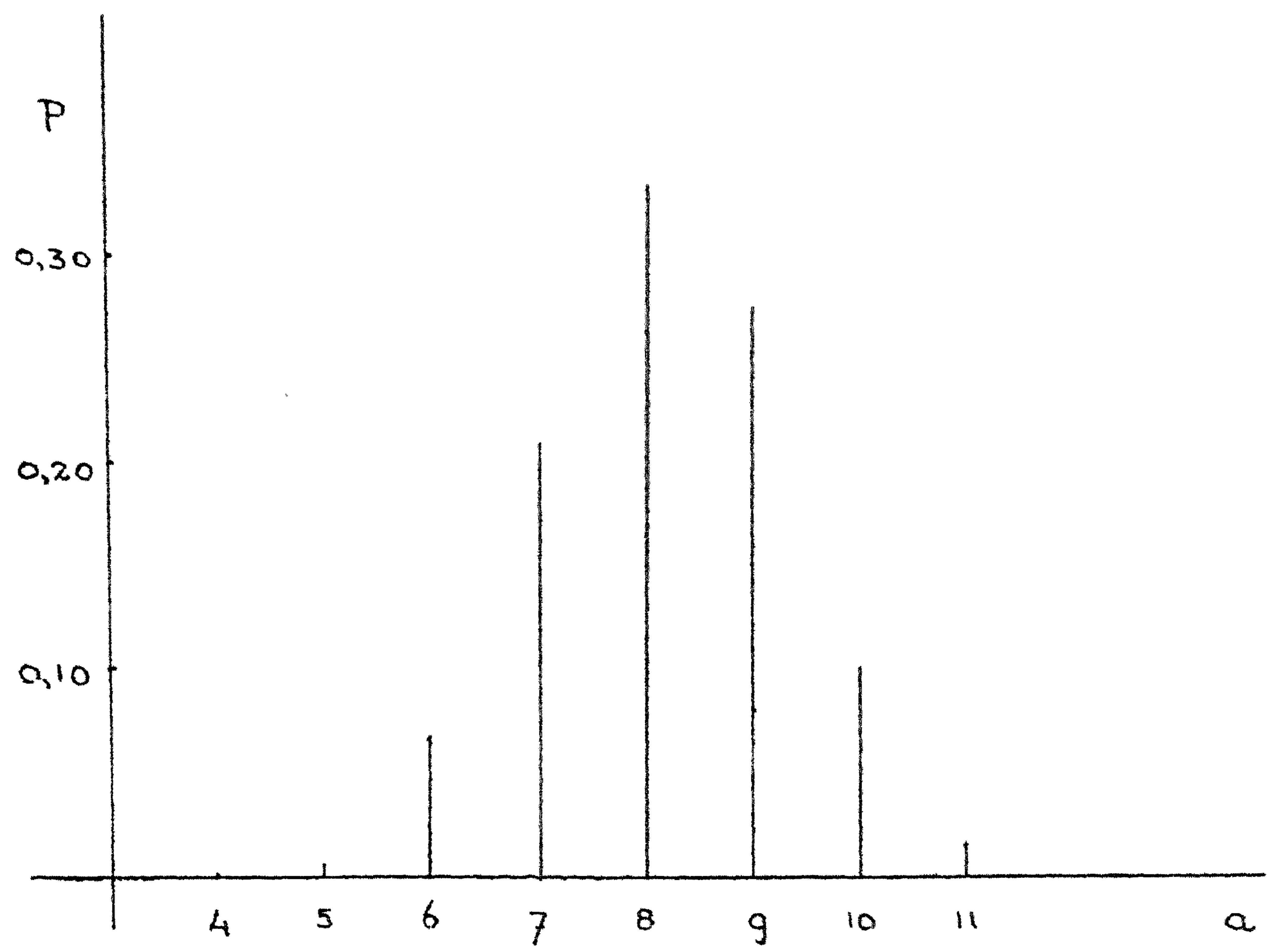

fig. 1 $P[\underline{a} = a \mid \underline{r} = 20; H_0]$, $n = 11$, $m = 16$

Het gemiddelde en de variantie (spreidingskwadraat) van $\underline{a}$ worden gegeven door

(2) $$\mathcal{E}(\underline{a} \mid \underline{r} = r; H_0) = \frac{nr}{N},$$

(3) $$\sigma^2(\underline{a} \mid \underline{r} = r; H_0) = \frac{nmrs}{N^2(N-1)}.$$

Wil men als andere mogelijkheden beschouwen $p \neq p'$ dan kiest men een tweezijdige kritieke zone Z. Om deze kritieke zone te vormen worden de waarden van $\underline{a}$ met de kleinste waarschijnlijkheden bij elkaar gezocht totdat de gekozen onbetrouwbaarheidsdrempel het toevoegen van een nieuwe waarde verhindert.

De overschrijdingskans k wordt gevonden door alle waarschijnlijkheden bij elkaar te zoeken die niet groter zijn dan de waarschijnlijkheid van de gevonden waarde van $\underline{a}$.

Nemen we $\alpha = 0{,}05$ dan bestaat de tweezijdige kritieke zone Z in ons voorbeeld uit de waarden $a = 4$; 5 en 11. De kans α' dat $\underline{a}$ in Z valt is dan, onder de hypothese H_0, (zie tabel III) gelijk aan: 0,0004 + 0,0083 + 0,0129 = 0,0216. Zouden wij de waarde $a = 6$ ook nog bij Z trekken, dan zou de waarde α overschreden worden.
We zien dat α' in dit geval veel kleiner dan α, zelfs klei-

ner dan $\frac{1}{2}\alpha$, is. Het alleen opgeven van α zou dus misleidend zijn.
De overschrijdingskans wordt (voor de in het voorbeeld voor a gevonden waarde 6):

$$k = 0,0004 + 0,0083 + 0,0129 + 0,0624 = 0,0840.$$

De hypothese H_0 wordt dus (bij een onbetrouwbaarheidsdrempel van 0,05) niet verworpen.

Wil men als andere mogelijkheden naast $p = p'$ alleen beschouwen $p > p'$ (dit kan men b.v. doen als men weet dat p zeker niet kleiner is dan p'), dan kiest men een rechts éénzijdige kritieke zone Z_r, d.w.z. men kiest een kritieke zone die alleen bestaat uit grote waarden van a. De overschrijdingskans k wordt nu gevonden door de waarschijnlijkheden van alle waarden van a die niet kleiner zijn dan de gevonden waarde bij elkaar te tellen. Analoog voor een linkeréénzijdige kritieke zone Z_ℓ.

In ons voorbeeld bestaat een rechts éénzijdige kritieke zone Z_r (bij een onbetrouwbaarheidsdrempel 0,05) uit de waarde $a = 11$. De kans α' dat a in Z_r valt is, onder de hypothese H_0, gelijk aan 0,0129.
Een links éénzijdige kritieke zone Z_ℓ met onbetrouwbaarheidsdrempel 0,05 bestaat uit de waarden $a = 4$ en 5; in dit geval is $\alpha' = 0,0087$.

1.1.3. Benadering voor de verdeling van a voor grote waarden van N.

Voor grote waarden van N maken wij gebruik van het feit dat a bij benadering normaal verdeeld is met gemiddelde $\frac{nr}{N}$ en spreiding $\sqrt{\frac{nmrs}{N^2(N-1)}}$ (zie formules (2) en (3)). Daar wij hiermee een discrete verdeling benaderen met een continue passen wij de z.g. continuïteitscorrectie toe, d.w.z. wij nemen voor a een waarde die $\frac{1}{2}$ dichter bij het gemiddelde ligt.

Toetst men tweezijdig dan is de overschrijdingskans bij benadering 2 maal het oppervlak van de normale verdeling met gemiddelde 0 en spreiding 1, dat rechts van

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

ligt (de verticale strepen geven aan, dat de absolute waarde van $ad-bc$ genomen moet worden, d.w.z. $ad-bc$ zelf als dit > 0 is en anders $bc-ad$).

Beschouwt men als alternatieve mogelijkheden naast $p = p'$ alleen $p > p'$, dan is de overschrijdingskans bij benadering het oppervlak rechts van

$$\frac{ad-bc-\frac{1}{2}N}{\sqrt{\frac{nmrs}{N-1}}}\ .$$

Dit is dan de rechteréénzijdige overschrijdingskans. De linker-overschrijdingskans (behorend bij linkszijdige toetsing, als $p < p'$ de enige toegelaten alternatieve mogelijkheid is) is bij benadering gelijk aan het oppervlak links van

$$\frac{ad-bc+\frac{1}{2}N}{\sqrt{\frac{nmrs}{N-1}}}\ .$$

Voorbeeld 2.

Bij 52 mannen en 109 vrouwen werd nagegaan of zij een bepaalde stof (phenyl-tio-carbamide) konden proeven [3]. Het doel van dit onderzoek was na te gaan of er een verschil is tussen mannen en vrouwen wat betreft het al of niet proeven van deze stof.

De resultaten staan vermeld in tabel IV:

Tabel IV

	proevers	niet proevers	totaal
mannen	35	17	52
vrouwen	69	40	109
totaal	104	57	161

We hebben dus: $a=35$, $b=69$, $c=17$, $d=40$

$n=52$, $m=109$, $r=104$, $s=57$, $N=161$.

Dit geeft $ad-bc=227$, dus $ad-bc$ is positief (en dus gelijk aan $|ad-bc|$) en we krijgen dus

$$\frac{|ad-bc|-\frac{1}{2}N}{\sqrt{\frac{nmrs}{N-1}}} = 0{,}32\ .$$

In een tabel van de normale verdeling vinden we voor de tweezijdige overschrijdingskans $k = 2 \times 0{,}3745 = 0{,}75$ en de hypothese, dat er geen verschil is tussen mannen en vrou-

----

3) Zie: Sexual and racial variations in ability to taste phenyl-tio-carbamide, with some data on the inheritance, by William C.Boyd and Lyle G.Boyd, Ann. Eugen. 8 (1936), p. 46-51.

wen wat betreft het al of niet proeven van phenyl-tio-carbamide, wordt op grond van dit experiment dus niet verworpen.

1.2. Het combineren van een aantal dubbele dichotomieën.

1.2.1. Inleiding.

We zullen hier beginnen met een voorbeeld:

Het bovengenoemde onderzoek naar het verschil tussen mannen en vrouwen wat betreft het al of niet proeven van phenyl-tio-carbamide werd o.a. verricht in de stad San Sebastian. In deze stad kan men twee bevolkingsgroepen onderscheiden, nl. de Basken en de niet-Basken. Het onderzoek werd nu uitgevoerd voor ieder van deze twee groepen apart en de resultaten staan vermeld in de tabellen V en VI:

Tabel V
Basken

	proevers	niet proevers	totaal
mannen	24	5	29
vrouwen	49	20	69
totaal	73	25	98

Tabel VI
niet-Basken

	proevers	niet proevers	totaal
mannen	17	8	25
vrouwen	35	14	49
totaal	52	22	74

De hypothese die we nu willen toetsen is, dat er noch bij de Basken, noch bij de niet-Basken een verschil is tussen mannen en vrouwen wat betreft het proeven van phenyl-tio-carbamide.

1.2.2. Beschrijving van de methode.

In het algemene geval hebben wij dus een aantal tabellen van de vorm:

	A	$\bar{A}$	totaal
eerste reeks	a_i	c_i	n_i
tweede reeks	b_i	d_i	m_i
	r_i	s_i	N_i

waarbij $i = 1, 2, \ldots, \ell$.

Stel nu dat bij de i^e proef de kansen op A bij de eerste en tweede reeks resp. p_i en p_i' zijn. De hypothese die we dan willen toetsen is dat voor iedere i geldt: $p_i = p_i'$ De p_i s mogen echter onderling verschillen.

De alternatieve hypothesen houden uiteraard in, dat voor minstens één der tabellen geldt: $p_i \neq p_i'$. Indien men nu weet, dat alle eventuele verschillen tussen p_i en p_i' hetzelfde teken bezitten, dus dat $p_i > p_i'$ voor alle i met $p_i \neq p_i'$ òf $p_i < p_i'$ voor al deze gevallen, dan berekent men de grootheid:

$$\underline{x} = \frac{\sum_{i=1}^{\ell}\left(\underline{a}_i - \frac{n_i r_i}{N_i}\right)}{\sqrt{\sum_{i=1}^{\ell}\frac{n_i m_i r_i s_i}{N_i^2(N_i-1)}}}.$$

Deze grootheid $\underline{x}$ is onder de hypothese H_0 bij benadering normaal verdeeld met gemiddelde nul en spreiding 1

De overschrijdingskans kan weer in een tabel der normale verdeling worden opgezocht.

Een continuïteitscorrectie is hier moeilijk aan te brengen en wordt daarom gewoonlijk buiten beschouwing gelaten.

Beschouwt men als alternatieve mogelijkheden naast $p_i = p_i'$, $p_i \neq p_i'$ (waarbij dus het teken van $p_i - p_i'$ niet voor alle i hetzelfde behoeft te zijn) dan berekent men:

$$\underline{\chi}^2 = \sum_{i=1}^{\ell}\frac{(\underline{a}_i N_i - n_i r_i)^2}{\frac{n_i m_i r_i s_i}{N_i - 1}} = \sum_{i=1}^{\ell}(N_i-1)\frac{(\underline{a}_i \underline{d}_i - \underline{b}_i \underline{c}_i)^2}{n_i m_i r_i s_i}.$$

De grootheid $\underline{\chi}^2$ heeft, onder de hypothese H_0, bij benadering een χ^2-verdeling met ℓ graden van vrijheid. De overschrijdingskans (die hier gelijk is aan het oppervlak dat rechts van de gevonden waarde van $\underline{\chi}^2$ ligt) kan in een tabel van de χ^2-verdeling worden opgezocht.

Opmerking.

Het combineren van een aantal dubbele dichotomieën kan nog op vele andere wijzen geschieden, b.v. als volgt:

1. Als men weet dat voor iedere i geldt $p_i = p$ en $p_i' = p'$ (dus dat p en p' niet van proef tot proef verschillen) dan vormt men uit de gegeven ℓ dubbele dichotomieën één nieuwe, die er dan als volgt uitziet:

	A	$\bar{A}$	totaal
eerste reeks	$\Sigma \underline{a}_i$	$\Sigma \underline{c}_i$	Σn_i
tweede reeks	$\Sigma \underline{b}_i$	$\Sigma \underline{d}_i$	Σm_i
	$\Sigma \underline{r}_i$	$\Sigma \underline{s}_i$	ΣN_i

DEEL II

2.1. Het toetsen van de hypothese $p = p_0$.

2.1.1. Inleiding.

We beschouwen een reeks van onafhankelijke experimenten, waarbij ieder experiment één der mogelijke uitkomsten A of $\bar{A}$ heeft. Geven we het aantal malen A aan met $\underline{n}_1$ dan kan $\underline{n}_1$ de waarden $0,1,2,\ldots,n$ aannemen.

Indien bij ieder experiment de kans op A gelijk is aan p dan wordt de waarschijnlijkheidsverdeling van $\underline{n}_1$ gegeven door:

(4) $$P[\underline{n}_1 = n_1] = \binom{n}{n_1} p^{n_1} q^{n-n_1} \qquad (q = 1-p).$$

Deze verdeling wordt de binomiale verdeling of verdeling van BERNOULLI genoemd.

De hypothese die we willen toetsen is:

$$H_0 : \; p = p_0 .$$

2.1.2. De toetsingsmethode.

Als toetsingsgrootheid gebruiken we de grootheid $\underline{n}_1$. De waarschijnlijkheidsverdeling van $\underline{n}_1$ onder de hypothese H_0 wordt gegeven door

(5) $$P[\underline{n}_1 = n_1 \mid H_0] = \binom{n}{n_1} p_0^{n_1} q_0^{n-n_1} \qquad (q_0 = 1-p_0).$$

Tabellen der binomiale verdeling zijn te vinden in [3] ($n = 2(1)49$ en $p_0 = 0{,}01(0{,}01)0{,}50$) en [4] ($n = 2(1)200$ en $p_0 = 0{,}5$).

Voorbeeld 4.

In figuur 2 en tabel VII vinden we een voorbeeld van de binomiale verdeling voor $n = 10$ en $p_0 = 0{,}3$.

Tabel VII

n_1	$\binom{n}{n_1} p_0^{n_1} q_0^{n-n_1}$	n_1	$\binom{n}{n_1} p_0^{n_1} q_0^{n-n_1}$
0	0,0282	6	0,0368
1	0,1211	7	0,0090
2	0,2335	8	0,0014
3	0,2668	9	0,0001
4	0,2001	10	0,0000
5	0,1029		

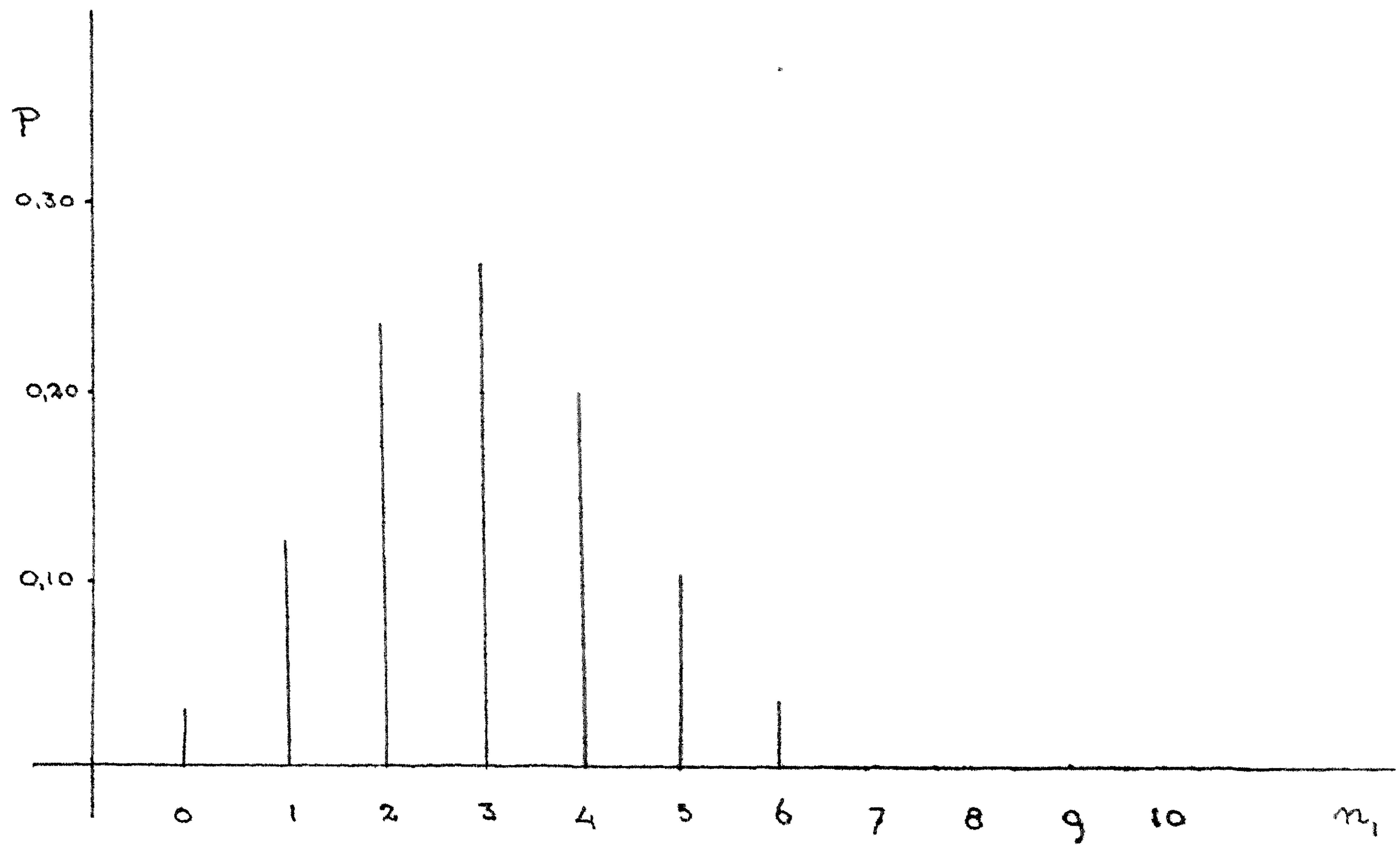

fig. 2 $P[\underline{n}_1 = n_1]$ $n = 10$, $p_0 = 0,3$, $q_0 = 1 - p_0$.

Het gemiddelde en de variantie (spreidingskwadraat) van $\underline{n}_1$ worden gegeven door

(6) $\mathcal{E}(\underline{n}_1 \mid H_0) = n\, p_0$,

(7) $\sigma^2(\underline{n}_1 \mid H_0) = n\, p_0\, q_0$.

De rechtséénzijdige kritieke zone Z_r (welke men gebruikt als men als alternatieve mogelijkheden naast $p = p_0$ beschouwt: $p > p_0$) bestaat uitsluitend uit grote waarden van n_1; de linkséénzijdige Z_ℓ uitsluitend uit kleine waarden van n_1. De rechtséénzijdige overschrijdingskans is de som van de waarschijnlijkheden van die waarden van $\underline{n}_1$, die niet kleiner zijn dan de gevonden waarde van $\underline{n}_1$. De linkséénzijdige overschrijdingskans is de som van de waarschijnlijkheden van die waarden van $\underline{n}_1$, die niet groter zijn dan de gevonden waarde van $\underline{n}_1$.

In voorbeeld 4 bestaat de rechtséénzijdige kritieke zone Z_r (met onbetrouwbaarheidsdrempel 0,05) uit de waarden $n_1 = 6;7;8;9$ en 10. De kans α' dat $\underline{n}_1$ in Z_r valt is, onder de hypothese H_0 ($p = 0,3$) (zie tabel VII) gelijk aan:

$\alpha' = 0,0368 + 0,0090 + 0,0014 + 0,0001 = 0,0483$.

De linkséénzijdige kritieke zone Z_ℓ (met onbetrouwbaarheidsdrempel 0,05) bestaat uit de waarde $n_1 = 0$. Hier wordt $\alpha' = 0,0282$.

De tweezijdige kritieke zone Z' (met onbetrouwbaarheidsdrempel α) kan men nu vormen uit een rechtséénzijdige Z_r en een linkséénzijdige Z_ℓ, ieder met onbetrouwbaarheidsdrempel $\frac{1}{2}\alpha$. De overschrijdingskans wordt nu gedefinieerd als tweemaal de éénzijdige overschrijdingskans (en wel moet hierbij de kleinste der twee éénzijdige overschrijdingskansen genomen worden). Men kan de tweezijdige kritieke zone echter ook vormen als is aangegeven in 1.1.2. (Deze kritieke zone zullen we aangeven met Z.) We zoeken dan de waarden van n_1 met de kleinste waarschijnlijkheden bij elkaar totdat de gekozen onbetrouwbaarheidsdrempel het toevoegen van een nieuwe waarde verhindert. De tweezijdige overschrijdingkans is de som van alle waarschijnlijkheden die niet groter zijn dan de waarschijnlijkheid van de gevonden waarde van n_1 [4]).

In <u>voorbeeld 4</u> bestaat de tweezijdige kritieke zone Z' (gevormd uit een links- en rechtséénzijdige, ieder met onbetrouwbaarheidsdrempel 0,025) uit de waarden $n_1 = 7;8;9$ en 10. De kans α' dat n_1 in Z' valt, onder de hypothese H_0, is gelijk aan:

$$\alpha' = 0{,}0090 + 0{,}0014 + 0{,}0001 = 0{,}0105.$$

De tweezijdige kritieke zone Z, gevormd door het bij elkaar zoeken van de waarden van n_1 met de kleinste waarschijnlijkheden, bestaat (bij een onbetrouwbaarheidsdrempel $\alpha = 0{,}05$) uit de waarden $n_1 = 0;\ 7;\ 8;\ 9$ en 10.
De kans α', dat n_1 in Z valt, onder hypothese H_0, is nu:

$$\alpha' = 0{,}0001 + 0{,}0014 + 0{,}0090 + 0{,}0282 = 0{,}0387.$$

De twee bovenbeschreven tweezijdige kritieke zones vallen samen als $p_0 = \frac{1}{2}$. De waarschijnlijkheidsverdeling van n_1 is dan nl. symmetrisch. Als $p_0 \neq \frac{1}{2}$ is zal, in het algemeen, de werkelijke onbetrouwbaarheid α' van Z dichterbij α liggen dan die van Z'.

----

4) Een kritieke zone van het type van Z', dus met hoogstens $\frac{1}{2}\alpha$ aan beide kanten, kan ook gebruikt worden voor de tweezijdige toets van de hypothese $p = p'$ (zie 1.1.2). De reden, dat deze methode daar niet en hier wel beschreven is, is, dat wij de kritieke zone Z in het algemeen prefereren. Deze leidt echter bij het bepalen van een tweezijdig betrouwbaarheidsinterval (zie 2.2) van een onbekende kans tot complicaties, die met behulp van Z' vermeden kunnen worden. Dit deed zich bij deel I niet voor.

2.1.3. Benadering voor de verdeling van n_1 voor grote waarden van n.

Voor grote waarden van n en als p_0 niet te ver van $\frac{1}{2}$ verwijderd is, is de grootheid n_1 bij benadering normaal verdeeld met gemiddelde np_0 en variantie np_0q_0 (zie formules (6) en (7)).

We passen weer (evenals in 1.1.3) een continuïteitscorrectie toe.

De tweezijdige overschrijdingskans (die we gebruiken als zowel $p > p_0$ als $p < p_0$ als alterternatieve mogelijkheden worden beschouwd) is bij benadering gelijk aan twee maal het oppervlak van de normale verdeling met gemiddelde 0 en spreiding 1 dat rechts van $\frac{|n_1 - np_0| - \frac{1}{2}}{\sqrt{np_0q_0}}$ ligt.

Beschouwt men als alternatieve mogelijkheden naast $p = p_0$ alleen $p > p_0$ dan is de rechtséénzijdige overschrijdingskans bij benadering het oppervlak rechts van

$$\frac{n_1 - np_0 - \frac{1}{2}}{\sqrt{np_0q_0}} .$$

De linkséénzijdige overschrijdingskans is bij benadering gelijk aan het oppervlak links van

$$\frac{n_1 - np_0 + \frac{1}{2}}{\sqrt{np_0q_0}} .$$

Voorbeeld 5.

In 1930 werden in Amsterdam 6848 jongens en 6374 meisjes geboren [5]).

Met behulp van de bovenbeschreven methode kunnen we de hypothese toetsen dat de kans op een jongensgeboorte gelijk is aan de kans op een meisjesgeboorte, d.w.z. we kunnen de hypothese H_0 : $p = \frac{1}{2}$ toetsen. Als alternatieve mogelijkheden beschouwen wij, naast $p = \frac{1}{2}$: $p \neq \frac{1}{2}$. We toetsen dus tweezijdig. We krijgen: $n = 6848 + 6374 = 13.222$; $n_1 = 6848$ en $p_0 = \frac{1}{2}$. Dus $n_1 - np_0 = 237$ en

$$\frac{|n_1 - np_0| - \frac{1}{2}}{\sqrt{np_0q_0}} = \frac{2 \times 236,5}{\sqrt{13.222}} = 4,11.$$

De tweezijdige overschrijdingskans is 0,0002. De hypothese H_0 wordt dus verworpen ten gunste van de hypothese dat de kans op een jongensgeboorte groter is dan die op een meisjesgeboorte.

----

5) Zie: "De bevolking van Amsterdam", deel I, Amsterdam 1933, pag. 47.

Voor grote waarden van n en kleine waarden van p_0 (of $1-p_0$) benadert men de verdeling van $\underline{n}_1$ met de Poissonverdeling met gemiddelde np_0. Voor kleine p_0 geldt bij benadering:

$$P[\underline{n}_1 = n_1 \mid H_0] = \frac{e^{-np_0}(np_0)^{n_1}}{n_1!} .$$

De rechtséénzijdige kritieke zone Z_r bestaat weer uitsluitend uit grote waarden van $\underline{n}_1$, de linkséénzijdige Z_ℓ uitsluitend uit kleine waarden van $\underline{n}_1$.

De tweezijdige kritieke zone Z' (met onbetrouwbaarheidsdrempel α) bestaat uit een rechtséénzijdige en een linkséénzijdige ieder met onbetrouwbaarheidsdrempel $\frac{1}{2}\alpha$ [6]).

Tabellen der Poissonverdeling vindt men in [7] en [8]. Een nomogram der Poissonverdeling vindt men in [9].

Opmerking.

Bij kleine waarden van $1-p_0 = q_0$ vervangt men n_1 door $n_2 = n - n_1$. Dan geldt bij benadering:

$$P[\underline{n}_2 = n_2 \mid H_0] = \frac{e^{-nq_0}(nq_0)^{n_2}}{n_2!} .$$

2.2. Het bepalen van een betrouwbaarheidsinterval voor een kans p.

2.2.1. Het bepalen van een betrouwbaarheidsinterval in het algemeen.

Een betrouwbaarheidsinterval $\underline{\mathfrak{I}}$ voor een onbekende parameter θ (b.v. het gemiddelde van een waarschijnlijkheidsverdeling of een onbekende kans) is een interval waarvan de grenzen stochastisch zijn en dat de eigenschap bezit, behoudens een zekere onbetrouwbaarheid α, de ware waarde van θ te bevatten.

Het algemene principe ter bepaling van een betrouwbaarheidsinterval is het volgende:

Zij gegeven een aantal waarnemingen $x_1, x_2, \ldots, x_n$ van een stochastische grootheid $\underline{x}$; θ is een onbekende parameter van de waarschijnlijkheidsverdeling van $\underline{x}$. Zij nu T een toets voor de hypothese

$$H_0 : \theta = \theta_0$$

dan is het betrouwbaarheidsinterval $\underline{\mathfrak{I}}$ de verzameling van al die waarden θ_0 die bij de toepassing van T op grond van de waarne-

---

6) Het in 1.1.2 gebruikte principe van het bijeenzoeken der kleinste waarschijnlijkheden leidt in dit geval zeer vaak tot éénzijdigheid en wel rechtséénzijdigheid der kritieke zone, zodat wij het gebruik hier niet kunnen aanraden.

mingen $x_1, x_2, \ldots, x_n$ niet voor verwerping in aanmerking komen.

Is de toets T toegepast met een onbetrouwbaarheidsdrempel α dan is dit ook de onbetrouwbaarheidsdrempel van het betrouwbaarheidsinterval.

Want de kans om de juiste hypothese te verwerpen is $\leqq \alpha$. Daar het betrouwbaarheidsinterval $\underline{\mathfrak{I}}$ bestaat uit alle niet verworpen waarden θ_0, is de kans dat $\underline{\mathfrak{I}}$ de ware waarde van θ niet bevat $\leqq \alpha$.

2.2.2. Exacte bepaling van één- en tweezijdige betrouwbaarheidsintervallen voor een kans.

Onder een éénzijdig naar boven begrensd betrouwbaarheidsinterval voor een kans p verstaat men een betrouwbaarheidsinterval, waarvan de bovengrens stochastisch is en de ondergrens 0. Om dit betrouwbaarheidsinterval te bepalen, maken wij gebruik van de toets voor de hypothese

$$H_0: \quad p = p_0.$$

Dan geldt:

$$P[\underline{n}_1 = n_1 \mid H_0] = \binom{n}{n_1} p_0^{n_1} q_0^{n-n_1}.$$

We kiezen nu een linkséénzijdige kritieke zone en verwerpen H_0 dus als

$$P[\underline{n}_1 \leqq n_1 \mid H_0] = \sum_{i=0}^{n_1} \binom{n}{i} p_0^i q_0^{n-i} \leqq \alpha$$

is.

Deze som neemt af als p_0 toeneemt; verwerpen we dus een bepaalde waarde p_0 dan is dit ook het geval met alle p waarvoor geldt $p > p_0$. De bovengrens van het betrouwbaarheidsinterval is dus de kleinste p_0 waarvoor

$$\sum_{i=0}^{n_1} \binom{n}{i} p_0^i q_0^{n-i} \leqq \alpha$$

is, of: de kleinste p_0, waarvoor

$$\sum_{i=n_1+1}^{n} \binom{n}{i} p_0^i q_0^{n-i} \geqq 1-\alpha$$

is.

Onder een éénzijdig naar beneden begrensd betrouwbaarheidsinterval voor een kans p verstaan we een betrouwbaarheidsinterval waarvan de ondergrens stochastisch is en de bovengrens 1. De bepaling geschiedt geheel analoog als boven, maar nu met een rechtséénzijdige kritieke zone. We verwerpen H_0 nu dus als:

$$P[\underline{n}_1 \geq n_1 \mid H_0] = \sum_{i=n_1}^{n} \binom{n}{i} p_0^i q_0^{n-i} \leq \alpha$$

is.

Deze som neemt af als p_0 afneemt; dus als een bepaalde waarde p_0 verworpen wordt is dit ook het geval met alle p waarvoor geldt $p < p_0$. De ondergrens van het betrouwbaarheids-interval is de grootste p_0, waarvoor

$$\sum_{i=n_1}^{n} \binom{n}{i} p_0^i q_0^{n-i} \leq \alpha$$

is.

Voorbeeld 6.

Stel $n = 20$ en $n_1 = 6$.

We willen nu een naar beneden begrensd betrouwbaarheids-interval voor p bepalen met onbetrouwbaarheidsdrempel 0,05

In tabel [3] vinden we

$$\sum_{i=6}^{20} \binom{20}{i} p_0^i q_0^{20-i}$$

voor $p_0 = 0{,}01(0{,}01)0{,}50$.

We moeten nu de grootste p_0 bepalen waarvoor

$$\sum_{i=6}^{20} \binom{20}{i} p_0^i q_0^{20-i} \leq 0{,}05$$

is.

In de tabel vinden we:

$$\sum_{i=6}^{20} \binom{20}{i} p_0^i q_0^{20-i} \begin{cases} = 0{,}0369738 \text{ voor } p_0 = 0{,}13 \\ = 0{,}0506727 \text{ voor } p_0 = 0{,}14, \end{cases}$$

dus de ondergrens van het betrouwbaarheidsinterval is 0,13.

De ware p is dus behoudens een onbetrouwbaarheid $\leq 0{,}05$ $> 0{,}13$.

Voor een naar boven begrensd betrouwbaarheidsinterval moeten we de kleinste p_0 zoeken waarvoor

$$\sum_{i=7}^{20} \binom{20}{i} p_0^i q_0^{20-i} \geq 0{,}95$$

is.

Uit de tabel zien we dat deze p_0 groter dan 0,50 is. Daarom vervangen we p_0 door q en q_0 door p ($p+q=1$) We krijgen dan:

$$\sum_{i=7}^{20} \binom{20}{i} p_0^i q_0^{20-i} = \sum_{i=7}^{20} \binom{20}{i} p^{20-i} q^i = \sum_{i=0}^{13} \binom{20}{i} p^i q^{20-i} =$$

$$= 1 - \sum_{i=14}^{20} \binom{20}{i} p^i q^{20-i}.$$

Gezocht wordt nu dus de grootste p waarvoor

$$\sum_{i=14}^{20} \binom{20}{i} p^i q^{20-i} \leq 0{,}05$$

is.

De tabel geeft:

$$\sum_{i=14}^{20} \binom{20}{i} p^i q^{20-i} \begin{cases} = 0{,}0480128 \text{ voor } p = 0{,}49 \\ = 0{,}0576591 \text{ voor } p = 0{,}50, \end{cases}$$

dus $p = 1 - p_0 = 0{,}49$ of $p_0 = 0{,}51$.

Het naar boven begrensd betrouwbaarheidsinterval met onbetrouwbaarheid 0,05 bestaat dus uit alle p waarvoor geldt: $p < 0{,}51$.

Een tweezijdig betrouwbaarheidsinterval voor p (met onbetrouwbaarheidsdrempel α) kunnen we nu vinden door een naar boven en een naar beneden begrensd betrouwbaarheidsinterval te bepalen, ieder met onbetrouwbaarheidsdrempel $\frac{1}{2}\alpha$.

Dit komt er dus op neer dat we de hypothese

$$H_0 : p = p_0$$

toetsen en de in 2.1.2 gedefinieerde kritieke zone Z' gebruiken.

Voor de bepaling van een tweezijdig betrouwbaarheidsinterval kunnen we ook de kritieke zone Z gebruiken (zie 2.1.2). De bepaling van het betrouwbaarheidsinterval geschiedt dan weer volgens het in 2.2.1 beschreven principe en is vrij bewerkelijk. Men krijgt op deze wijze echter in het algemeen een nauwer betrouwbaarheidsinterval.

Voorbeeld 7.

Stel $n = 20$ en $n_1 = 6$ (dus dezelfde gegeven als in voorbeeld 6).

We willen nu een tweezijdig betrouwbaarheidsinterval voor p bepalen uit twee éénzijdige, ieder met onbetrouwbaarheidsdrempel 0,025.

De ondergrens van dit interval is de grootste p_0, waarvoor

$$\sum_{i=6}^{20} \binom{20}{i} p_0^i q_0^{20-i} \leq 0{,}025.$$

In tabel [3] vinden we:

$$\sum_{i=6}^{20} \binom{20}{i} p_0^i q_0^{20-i} \begin{cases} = 0{,}0175482 \text{ voor } p_0 = 0{,}11 \\ = 0{,}0260185 \text{ voor } p_0 = 0{,}12. \end{cases}$$

De ondergrens van het interval is dus 0,11.

De bovengrens is de kleinste p_0 waarvoor

$$\sum_{i=7}^{20} \binom{20}{i} p_0^i q_0^{n-i} \geq 0{,}975$$

is.

Stellen we weer $p = 1 - p_0$ dan zoeken we de grootste p waarvoor

$$\sum_{i=14}^{20} \binom{20}{i} p^i q^{20-i} \leq 0{,}025$$

is.

De tabel geeft:

$$\sum_{i=14}^{20} \binom{20}{i} p^i q^{20-i} \begin{cases} = 0{,}0214144 \text{ voor } p = 0{,}45 \\ = 0{,}0265129 \text{ voor } p = 0{,}46 \end{cases}$$

De bovengrens van het betrouwbaarheidsinterval is dus $1 - 0{,}45 = 0{,}55$.

Bepalen wij het betrouwbaarheidsinterval met behulp van de kritieke zone Z dan vinden we voor de ondergrens 0,13 en voor de bovengrens 0,53.

2.2.3. <u>Benaderingen voor een betrouwbaarheidsinterval voor een kans p</u>.

Voor grote waarden van n en als p niet te ver van $\frac{1}{2}$ verwijderd is, maken wij gebruik van de normale benadering van de verdeling van n_1.

Een betrouwbaarheidsinterval bestaat uit alle niet verworpen waarden van p_0; dus uit alle p_0, waarvoor (bij gegeven n en n_1) de overschrijdingskans $> \alpha$ is.

Voor de bepaling van een naar boven begrensd betrouwbaarheidsinterval kiezen we een linker kritieke zone. De linkséénzijdige overschrijdingskans wordt bepaald door

$$\frac{n_1 - n p_0 + \frac{1}{2}}{\sqrt{n p_0 q_0}}.$$

Dus de bovengrens van het naar boven begrensd betrouwbaarheidsinterval (met onbetrouwbaarheid α) is die p_0 waarvoor

$$\frac{n_1 - n p_0 + \frac{1}{2}}{\sqrt{n p_0 q_0}} = -\xi_\alpha ,$$

waarin ξ_α gevonden wordt uit de normale verdeling:

$$\frac{1}{\sqrt{2\pi}} \int_{\xi_\alpha}^{\infty} e^{-\frac{1}{2}x^2} dx = \alpha .$$

De ondergrens van het naar beneden begrensd betrouwbaarheidsinterval is die p_0, waarvoor

$$\frac{n_1 - n p_0 - \frac{1}{2}}{\sqrt{n p_0 q_0}} = +\xi_\alpha .$$

Literatuur

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

[2] Fry, T.C., Probability and its engineering uses, D.van Nostrand Company, New York 1928, p. 439-452.

[3] Tables of the binomial probability distribution, U.S. Dept. of Commerce, Nat. Bureau of Standards, Washington 1949.

[4] Wijngaarden, A.van, Table of the cumulative symmetric binomial distribution, Proc. Kon. Ned. Akad. van Wet., 53 (1950).

[5] Fisher, R.A., Statistical methods for research workers, London 1948, p. 96.

[6] Hemelrijk, J., Waarschijnlijkheidsrekening en statistiek, Vacantiecursus Mathematisch Centrum, Amsterdam 1950, §2, 3 en 4.

[7] Molina, E.C., Poisson's exponential binomial limit, D.van Nostrand Company, New York, 1945.

[8] Hartley, H.O. and E.S.Pearson, Tables of the χ^2-integral and of the cumulative Poisson distribution, Biometrika 37 (1950), p. 313-325.

[9] Dodge, H.F. and H.G.Romig, Sampling inspection tables, J.Wiley and Sons, 1944, p. 44.

[10] Dixon, W.J. and F.J.Massey, Introduction to statistical analysis, Mc Graw-Hill Book Company, Inc., New York 1951.

[11] Deming, W.E., Some theoryof sampling, John Wiley and Sons, Inc., Chapman and Hall, London 1950.