

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 85 (VP 1)

Sequente toets met drie beslissingsmogelijkheden
voor het toetsen van een onbekende
waarschijnlijkheid

door

J. de Boer

1952

Statistische Afdeling
van het
Mathematisch Centrum,
Amsterdam

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

Rapport S 85 (VP 1)

Sequente toets met drie beslissingsmogelijkheden
voor het toetsen van een onbekende
waarschijnlijkheid

door

J. de Boer.

Inhoud

	Pag.
1. Summary	1
2. Inleiding	1
3. Het mathematische model	1
4. Beschrijving van de toets	3
5. Eigenschappen van de toets	5
6. Afknotting van de toets	10
7. De analoge klassieke toets	14
8. Conclusies	15

ERRATA

<u>Op pag.:</u>	<u>staat:</u>	<u>moet zijn:</u>
7 r. 4 v.o.	$\beta_2(p)$	$\beta_3(p)$
8 r. 16 v.b.	P'_2	p'_2
8 r. 7 v.o.	want is	want α is
11 r. 7 v.b.	1	2
12 r. 12 v.b.	$\beta_1(p, T)$	$\beta_1(p, T^{(j)})$
16 r. 9 v.b.	P	p

1. Summary.

Sequential test with three possible decisions for testing an unknown probability.

Dividing the real axis by two points in three intervals, we may fix three hypotheses, each corresponding to one of the intervals, according to which the unknown mean of a normal distribution belongs to the related interval.

M.SOBEL and A.WALD [1]¹⁾ developed a sequential test for choosing one of these hypotheses. This test is a combination of two sequential Wald-tests [2]. The same method may be used for deriving a sequential test with three possible decisions for the binomial case, i.e. for testing an unknown probability. This test is described here and some simplifications of the Sobel and Wald-procedure are introduced. By means of some examples this test, truncated or not, is compared with a classical test for the same hypotheses, i.e. a test with fixed sample size.

2. Inleiding.

M.SOBEL en A.WALD [1]¹⁾ hebben een sequente toets ontwikkeld voor het toetsen van het onbekende gemiddelde van een normale verdeling met bekende spreiding, waarbij drie beslissingen mogelijk waren. De door hen gevolgde methode kan, met enkele wijzigingen, ook toegepast worden op het toetsen van de hypothese dat een onbekende waarschijnlijkheid bij een binomiale verdeling in een gegeven interval ligt. De toets, die dan verkregen wordt, wordt in de volgende paragrafen besproken en voor enkele numeriek uitgewerkte gevallen vergeleken met een toets waarbij de uitgebreidheid van de steekproef gegeven is (z.g. klassieke toets) en die dezelfde drie beslissingsmogelijkheden heeft.

3. Het mathematische model.

Wij beschouwen een alternatief met twee mogelijke uitkomsten A en B, die onbekende waarschijnlijkheden p en $q=1-p$ bezitten. De toets berust op een reeks van onderling afhankelijke waarnemingen van dit alternatief; het aantal waarnemingen

1) Numbers within brackets refer to the references at the end of this paper.

is stochastisch.

Twee getallen a_{12} en a_{32} tussen 0 en 1 zijn gegeven (of worden gekozen naar aanleiding van het probleem waarbij de toets wordt toegepast), en de hypothesen H_1 , H_2 en H_3 worden als volgt gedefinieerd:

$$H_1 : p < a_{12}$$

$$H_2 : a_{12} \leq p \leq a_{32}$$

$$H_3 : p > a_{32}$$

Verder is voor de toets vereist dat ter weerszijden van a_{12} twee punten p_1 en p'_2 en ter weerszijden van a_{32} twee punten p''_2 en p_3 gegeven zijn met

$$0 < p_1 < a_{12} < p'_2 < p''_2 < a_{32} < p_3 < 1.$$

De situatie is geschetst in figuur 1.

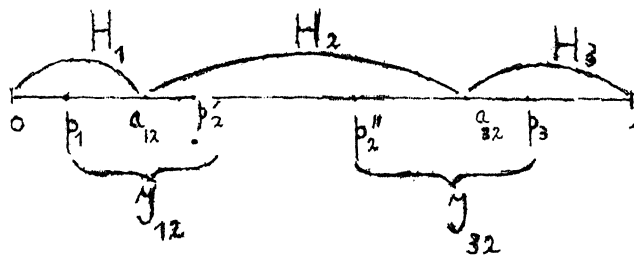


Fig. 1

De verdeling van de parameterruimte.

De volgende definitie voor de begrippen "juiste" en "onjuiste beslissing" wordt dan ingevoerd:

Tabel I
Juistheid van beslissingen

Waarde van p	Juiste beslissing	Onjuiste beslissing
$p \leq p_1$	Accepteer H_1	Accepteer H_2 of H_3
$p_1 < p < p'_2$	" H_1 of H_2	" H_3
$p'_2 \leq p \leq p''_2$	" H_2	" H_1 of H_3
$p''_2 < p < p_3$	" H_2 of H_3	" H_1
$p \geq p_3$	" H_3	" H_1 of H_2

De in figuur 1 met I_{12} en I_{32} aangegeven intervallen worden, in aansluiting aan de inhoud van tabel I, indifferentie-zônes genoemd. De getallen a_{12} en a_{32} zijn verder van weinig belang voor de uitwerking van de toets.

4. Beschrijving van de toets.

Wij zetten in een rechthoekig puntrooster met een rechthoekig assenstelsel de bij de opeenvolgende waarnemingen behorende weg uit door, uitgaande van de O, bij een waarneming met uitslag A één stap in verticale en bij uitslag B één stap in horizontale richting te doen. In dit vlak worden nu aanvaaringsgebieden G_1 , G_2 en G_3 voor H_1 , H_2 resp. H_3 geconstrueerd aan de hand van de door A.WALD [2], pag. 90 e.v., beschreven sequente toets met twee beslissingsmogelijkheden (vgl. fig. 2).

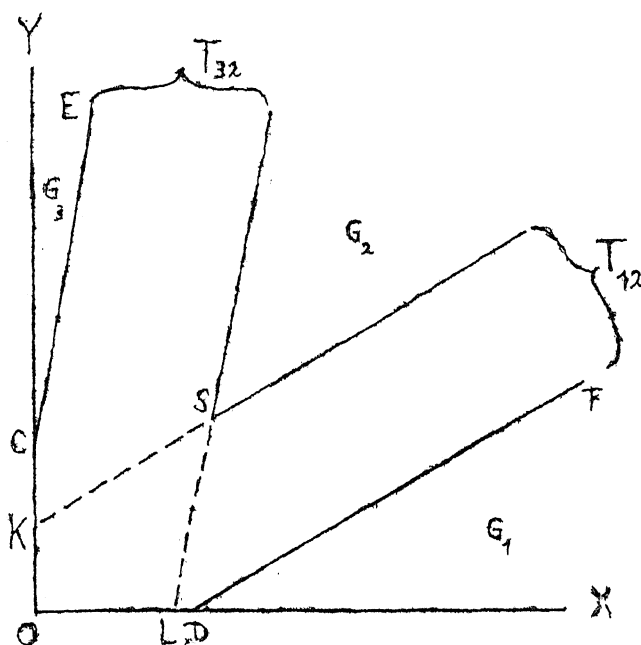


Fig. 2

Schema voor de sequente toets met drie beslissingsmogelijkheden.

Zij p de (onbekende) kans op A bij iedere waarneming en zij T_{12} de gewone sequente toets voor het toetsen van de hypothese $p \leq p_1$ tegen $p \geq p'_2$ en T_{32} de analoge toets voor $p \geq p_3$ tegen $p \leq p''_2$. Dan worden de grenzen van de twee bijbehorende "kanalen" (zie fig. 2) gegeven door

$$(p_i + q_i = 1 \text{ voor } i = 1, 3; p'_2 + q'_2 = p''_2 + q''_2 = 1):$$

$$T_{12}: \quad x \log \frac{q'_2}{q_1} + y \log \frac{p'_2}{p_1} = \log A_{12} \quad (\text{lijn KS}) \quad (1)$$

$$\text{en} \quad x \log \frac{q'_2}{q_1} + y \log \frac{p'_2}{p_1} = \log B_{12} \quad (\text{lijn DF}), \quad (2)$$

resp.

$$T_{32}: \quad x \log \frac{q_2''}{q_3} + y \log \frac{p_2''}{p_3} = \log A_{32} \quad (\text{lijn LS}) \quad (3)$$

$$x \log \frac{q_2''}{q_3} + y \log \frac{p_2''}{p_3} = \log B_{32} \quad (\text{lijn CE}). \quad (4)$$

Hierbij gelden volgens Wald (1.c.) de volgende relaties voor de constanten A_{12} , A_{32} , B_{12} en B_{32} : zij

$$\begin{aligned} \alpha_{12} &= \text{kans om met } T_{12} \text{ tot } p \geq p_2' \text{ te besluiten als } p = p_1 \text{ is} \\ \alpha_{21} &= \text{" " " } T_{12} \text{ " } p \leq p_1 \text{ " " " } p = p_2' \text{ "} \\ \alpha_{23} &= \text{" " " } T_{32} \text{ " } p \geq p_3 \text{ " " " } p = p_2'' \text{ "} \\ \alpha_{32} &= \text{" " " } T_{32} \text{ " } p \leq p_3 \text{ " " " } p = p_3 \text{ "}, \end{aligned}$$

dan is

$$A_{ij} \approx \frac{1 - \alpha_{ji}}{\alpha_{ij}} \quad \text{en} \quad B_{ij} \approx \frac{\alpha_{ji}}{1 - \alpha_{ij}} \quad (i, j = 1, 2, 3). \quad (5)$$

Uit de relaties (1) t/m (5) volgt gemakkelijk dat de helling van LS groter is dan die van KS en dat x_L en y_K positief zijn - als $\alpha_{ij} < \frac{1}{2}$ ($i, j = 1, 2, 3$), zoals we steeds zullen onderstellen - zodat LS en KS elkaar inderdaad in het eerste kwadrant snijden, zoals in figuur 2 aangegeven is. In verband met het volgende is het echter van belang op te merken dat zonder verdere voorwaarden niet noodzakelijk voldaan is aan de relaties

$$x_D \geq x_L \quad \text{en} \quad y_C \geq y_K. \quad (6)$$

Wij zullen echter steeds eisen dat aan deze relaties wel voldaan is en verderop de hiervoor nodige extra voorwaarden onderzoeken. De reden van deze eis blijkt bij de definitie van de sequente toets met drie beslissingsmogelijkheden.

Er zijn nu nl. drie mogelijkheden voor de weg die door de waarnemingen wordt aangegeven:

1. De weg passeert ¹⁾ eerst LS en dan DF. In dat geval wordt volgens T_{12} besloten tot $p \leq p_1$ en volgens T_{32} tot $p \leq p_2'$.

2. De weg passeert eerst KS en dan CE. Dan wordt volgens T_{12} besloten tot $p \geq p_2'$ en volgens T_{32} tot $p \geq p_3$.

3. De lijn passeert eerst KS en dan LS of het verlengde daarvan, of eerst LS en dan KS of het verlengde ervan. In deze gevallen wordt volgens T_{12} besloten tot $p \geq p_2'$ en volgens T_{32} tot $p \leq p_2''$.

1) Onder "passeren" wordt steeds verstaan: bereiken of overschrijden.

Wij definiëren nu de toets T met drie beslissingsmogelijkheden naar aanleiding van deze drie mogelijkheden als volgt:

Zodra de weg één der grenslijnen van T_{12} en één der grenslijnen van T_{32} gepasseerd is, is de toets afgelopen. Eindigt de weg in G_1 (dit is de eerste der drie bovenstaande mogelijkheden), dan wordt H_1 geaccepteerd. Eindigt de weg in G_3 (de tweede mogelijkheid), dan wordt H_3 geaccepteerd en eindigt de weg in G_2 , of zonder G_1 , G_2 of G_3 bereikt te hebben (derde mogelijkheid), dan wordt H_2 geaccepteerd.

Wij zien nu ook dat er moeilijkheden zouden kunnen ontstaan als aan de relaties (6) niet was voldaan. Immers dan zouden ook de volgende mogelijkheden (of althans minstens één ervan) hebben bestaan:

a. De weg passeert eerst DF en vervolgens CE (na weer over DF teruggegaan te zijn). Volgens T_{12} zou dan tot $p \leq p_1$ en volgens T_{32} tot $p \geq p_3$ besloten worden.

b. De weg passeert eerst CE en vervolgens DF (na weer over CE teruggegaan te zijn). Volgens T_{12} zou dan weer tot $p \leq p_1$ en volgens T_{32} tot $p \geq p_3$ besloten worden.

Daar met deze wegen bij de definitie van T niets te beginnen is, zijn wij genoodzaakt ze uit te sluiten.

Uit (1) t/m (4) volgt dat (6) geschreven kan worden als

$$\frac{\log B_{12}}{\log q'_2/q_1} \geq \frac{\log A_{32}}{\log q''_2/q_3} \quad \text{en} \quad \frac{\log B_{32}}{\log p''_2/p_3} \geq \frac{\log A_{12}}{\log p'_2/p_1} \quad (7),$$

waaruit, tezamen met (5), beperkingen volgen voor de keuze der waarden van de p 's en de α 's.

5. Eigenschappen van de toets.

De eigenschappen van de toets worden beschreven door zijn karakteristiek en door de waarschijnlijkheidsverdeling van het aantal waarnemingen dat nodig is om een beslissing te verkrijgen. Deze worden afzonderlijk behandeld.

5.1. De karakteristiek.

Onder deze naam vatten wij de drie functies van p samen (p is de kans op uitkomst A), die de kansen op de drie mogelijke beslissingen voorstellen. Wij gebruiken de volgende notatie voor de kans dat hypothese H_i zal worden aanvaard:

$$\beta_i(p) = P[H_i | p] \quad (i = 1, 2, 3). \quad (8)$$

Indien nodig voegen wij tussen de haken nog de letter T_{12} , T_{32} of T toe om aan te geven op welke toets β_1 betrekking heeft. Indien dit niet gebeurd is, bedoelen wij steeds T .

Daar de toetsen T_{12} en T_{32} (zie WALD [2]) beide met waarschijnlijkheid 1 tot een beslissing leiden, is dit ook voor T het geval. Derhalve geldt:

$$\sum_{i=1}^3 \beta_i(p) = 1 \quad \text{voor iedere } p. \quad (9)$$

Ook is bijv.:

$$\beta_1(p_1) = 1 - \alpha_{12} \quad \text{en} \quad \beta_3(p_3) = 1 - \alpha_{32}.$$

Verder volgen uit de in de vorige paragraaf gegeven definitie van T de vergelijkingen:

$$\beta_1(p, T) = \beta_1(p, T_{12}) \quad \text{en} \quad \beta_3(p, T) = \beta_3(p, T_{32}). \quad (10)$$

De rechterleden zijn echter door Wald (bij benadering; deze toevoeging laten wij verder weg) berekend, zodat met behulp van (9) de karakteristiek direct kan worden getekend. Deze bezit ongeveer een gedaante als de in figuur 3 geschetste.

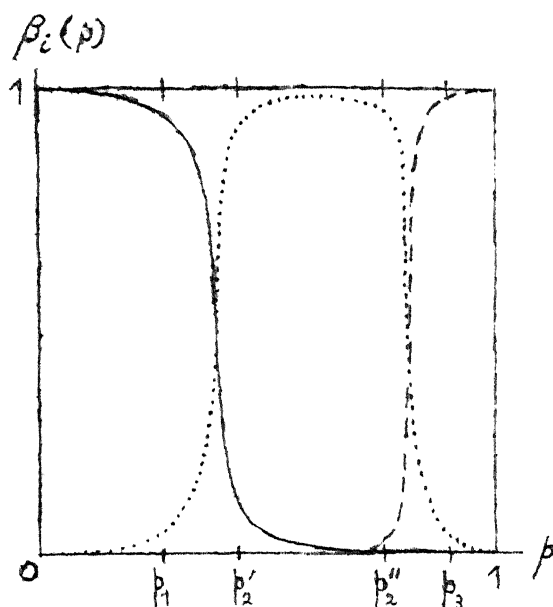


Fig. 3
Karakteristiek van T .

— β_1
..... β_2
----- β_3

Uit deze karakteristiek kan met behulp van tabel I een grafiek worden afgeleid voor de functie $\gamma(p)$ die de kans op het treffen van een onjuiste beslissing aangeeft. Deze func-

tie voldoet nl. volgens tabel I en volgens de definitie van α_{ij} (zie § 4) aan de volgende relaties:

Tabel II
Relaties tussen γ en β_i ($i = 1, 2, 3$)

Voor	Relatie
$p \leq p_1$	$\gamma = 1 - \beta_1$
$p = p_1$	$\gamma = \alpha_{12}$
$p_1 < p < p'_2$	$\gamma = \beta_3$
$p'_2 \leq p \leq p''_2$	$\gamma = 1 - \beta_2$
$p''_2 < p < p_3$	$\gamma = \beta_1$
$p = p_3$	$\gamma = \alpha_{32}$
$p \geq p_3$	$\gamma = 1 - \beta_3$

In figuur 4 vindt men de functie γ uitgezet, afgeleid uit de karakteristiek van figuur 3.

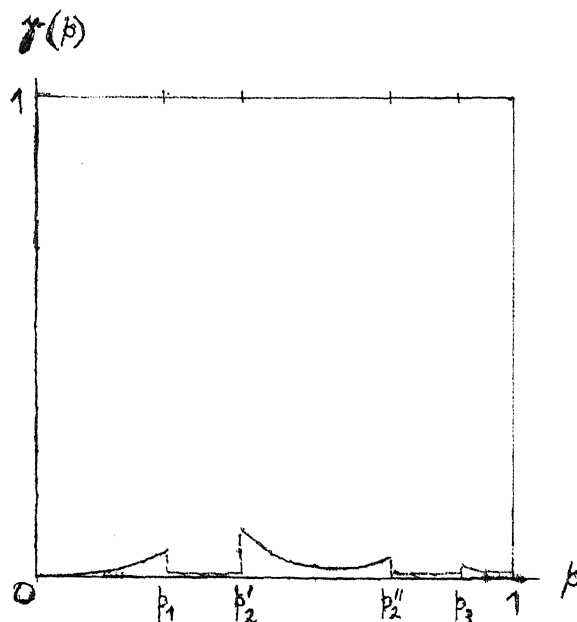


Fig. 4

De kans $\gamma(p)$ op een onjuiste beslissing.

Ten slotte kunnen wij de grootheid

$$\alpha = \text{Max. } \gamma(p) \quad 0 \leq p \leq 1 \quad (11)$$

invoeren en deze de onbetrouwbaarheidsdrempel van T noemen.

De krommen $\beta_1(p)$ en $\beta_2(p)$ bezitten ieder één buigpunt, dat volgens de theorie van Wald gelegen is in het interval (p_1, p'_2) resp. (p''_2, p_3) . Derhalve is $\gamma(p)$ in de intervallen $(0, p_1)$ en (p_1, p'_2) monotoon toenemend, in (p'_2, p_3) en $(p_3, 1)$

monotoon afnemend en in (p_2', p_2'') concaaf. Verder is

$$\beta_1(p_2') + \beta_2(p_2') + \beta_3(p_2') = 1,$$

dus

$$\beta_3(p_2') < 1 - \beta_2(p_2').$$

Evenzo geldt:

$$\beta_1(p_2'') < 1 - \beta_2(p_2''),$$

zodat wij krijgen:

$$\alpha = \text{Max.} \{ \alpha_{12}, 1 - \beta_2(p_2'), 1 - \beta_2(p_2''), \alpha_{32} \}, \quad (12)$$

terwijl bovendien de ongelijkheid

$$\text{Max.} \{ 1 - \beta_2(p_2'), 1 - \beta_2(p_2'') \} < \alpha_{21} + \alpha_{23} \quad (13)$$

geldt, maar deze bovengrens zal in het algemeen te groot zijn.

Het ligt echter voor de hand om bij voorgeschreven α te nemen:

$$\alpha_{12} = \alpha_{32} = \alpha_{21} + \alpha_{23} = \alpha, \quad (14)$$

daar men er dan zeker van is dat α bereikt, maar niet overschreden wordt.

Voor het symmetrische geval ($p_1 + p_3 = p_2' + p_2'' = 1$), dat wij later nog zullen gebruiken ter vergelijking van T met een klassieke toets, nemen wij dan bovendien

$$\alpha_{21} = \alpha_{23} = \frac{1}{2} \alpha \quad (15)$$

Volgens (5) wordt dan

$$A_{12} = A_{32} = \frac{2-\alpha}{2\alpha} \quad \text{en} \quad B_{12} = B_{32} = \frac{\alpha}{2(1-\alpha)}. \quad (16)$$

In dit geval is ook aan (7) voldaan, want

$$\frac{\log \frac{\alpha}{2(1-\alpha)}}{\log \frac{q_2'}{q_1}} \geq \frac{\log \frac{2-\alpha}{2\alpha}}{\log \frac{q_2''}{q_3}}, \quad \text{omdat, als we het linkerlid schrijven}$$

ven als

$$\frac{\log \frac{2(1-\alpha)}{\alpha}}{\log \frac{q_1}{q_2'}},$$

geldt (want is in de practijk altijd zeer klein):

$$\frac{2(1-\alpha)}{\alpha} > \frac{2-\alpha}{2\alpha}$$

en

$$\frac{q_1}{q_2'} = \frac{1-q_3}{1-q_2''} < \frac{q_2''}{q_3},$$

want de functie $y = x(1-x)$ neemt monotoon toe van $x=0$ tot $x=\frac{1}{2}$, zoals door differentiatie te zien is en omdat

$$q_3 < q_2'' \leq \frac{1}{2}$$

is, geldt:

$$q_3(1-q_3) < q_2''(1-q_2'').$$

Evenzo is de tweede ongelijkheid van (7) te bewijzen.

5.2. De mathematische verwachting van het aantal waarnemingen.

De grootheid, die wij nu zullen beschouwen, is $E(\underline{n}|p)$ ¹⁾, d.i. de mathematische verwachting van het aantal waarnemingen dat nodig is om met toets T tot een beslissing te komen als de kans op uitkomst A bij iedere stap p is. Voor deze grootheid geldt de volgende gelijkheid:

$$E(\underline{n}|p) = E(\underline{n}_{12}|p) + E(\underline{n}_{32}|p) - E(\underline{n}^*|p), \quad (17)$$

waarin $\underline{n}_{12}$ en $\underline{n}_{32}$ het aantal waarnemingen voorstellen dat nodig is om met T_{12} , resp. T_{32} tot een beslissing te komen, terwijl $\underline{n}^*$ het aantal stappen is dat nodig is om de grens KSL (zie figuur 2) voor het eerst te bereiken.

De juistheid van (17) kan als volgt worden ingezien: We kunnen twee soorten wegen onderscheiden:

1^e. Wegen eindigend aan één der grenzen van T_{12} . Deze geven we aan met $w_{12}^{(j)}$, waarin $j = 1, 2, \dots$, omdat er aftelbaar veel wegen zijn. Deze passeren alle de lijn LS. Het totale aantal stappen van de weg $w_{12}^{(j)}$ geven wij aan met $n_{12}^{(j)}$ en het aantal stappen tot en met het eerste snijpunt met LS noemen wij $m_{12}^{(j)}$.

2^e. Wegen eindigend aan één der grenzen van T_{32} . Deze passeren de lijn KS en worden aangegeven met $w_{32}^{(j)}$ ($j = 1, 2, \dots$). Het totale aantal stappen zijn $n_{32}^{(j)}$ en het aantal tot en met het eerste snijpunt met KS: $m_{32}^{(j)}$.

Noemen wij de kans op een bepaalde weg w, bij gegeven p, nu $P(w|p)$, dan gelden de volgende relaties:

$$\left. \begin{aligned} E(\underline{n}|p) &= \sum_i P(w_{12}^{(i)}|p) \cdot n_{12}^{(i)} + \sum_j P(w_{32}^{(j)}|p) \cdot n_{32}^{(j)} \\ E(\underline{n}_{12}|p) &= \sum_i P(w_{12}^{(i)}|p) \cdot n_{12}^{(i)} + \sum_j P(w_{32}^{(j)}|p) \cdot m_{32}^{(j)} \\ E(\underline{n}_{32}|p) &= \sum_i P(w_{12}^{(i)}|p) \cdot m_{12}^{(i)} + \sum_j P(w_{32}^{(j)}|p) \cdot n_{32}^{(j)} \\ E(\underline{n}^*|p) &= \sum_i P(w_{12}^{(i)}|p) \cdot m_{12}^{(i)} + \sum_j P(w_{32}^{(j)}|p) \cdot m_{32}^{(j)} \end{aligned} \right\} \quad (18)$$

Hieruit volgt (17) direct. Voor de eerste twee termen van (17) zijn door WALD [2] benaderingen zowel als bovengrenzen afgeleid en $E(\underline{n}^*|p)$ kan voor het hier behandelde probleem bij

1) Stochastische grootheden worden door een onderstreepte letter aangegeven.

ieder schema gemakkelijk exact worden berekend. Om een ruwe indruk van $E(\underline{n}|p)$ te verkrijgen, kunnen we de volgende benadering gebruiken:

$$\begin{aligned} \text{a) Voor } p < p'_2: E(\underline{n}|p) &\approx E(\underline{n}_{12}|p) \\ \text{b) Voor } p > p''_2: E(\underline{n}|p) &\approx E(\underline{n}_{32}|p). \end{aligned} \quad (19)$$

Uit (18) volgt dat deze schattingen te laag zijn, maar als we nu voor de rechterleden de door Wald afgeleide bovengrenzen gebruiken, krijgen we toch een behoorlijke indruk van de functie $E(\underline{n}|p)$. Voor $p'_2 \leq p \leq p''_2$ vertoont deze functie, zoals intuïtief gemakkelijk in te zien is, een "zak". Men vergelijke figuur 5.

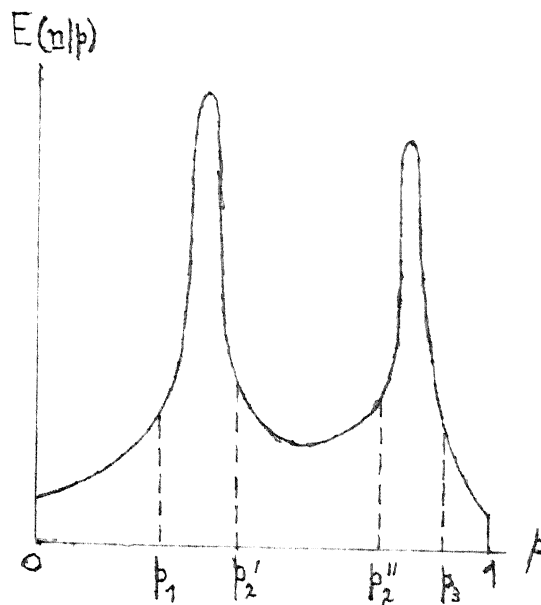


Fig. 5
Schets van $E(\underline{n}|p)$

Opmerking. SOBEL en WALD [1] gebruiken een andere methode voor het door hen behandelde probleem (vgl. § 2). Deze methode is echter omslachtiger en bezit geen voordelen boven de hier aangegevene, althans niet voor het hier behandelde probleem.

6. Afknotting van de toets.

Daar $E(\underline{n}|p)$ in (17) zeer groot kan zijn als p in I_{12} of in I_{32} ligt, zoals uit figuur 5 blijkt en het vaak gewenst is een bovengrens N voor het aantal waarnemingen vast te stellen, moeten wij een regel opstellen met behulp waarvan de toets T bij $\underline{n} = N$ afgeknot kan worden.

Deze regel luidt als volgt (analoog aan de regel voor afknotting bij een enkelvoudige Wald-toets), als er na N waarnemingen nog geen beslissing is:

Positie na N waarnemingen	Regel
$x \log \frac{q_2'}{q_1} + y \log \frac{p_2'}{p_1} < 0$	Accepteer H_1
$x \log \frac{q_2'}{q_1} + y \log \frac{p_2'}{p_1} \geq 0$ $x \log \frac{q_2''}{q_3} + y \log \frac{p_2''}{p_3} \geq 0$	Accepteer H_2
$x \log \frac{q_2''}{q_3} + y \log \frac{p_2''}{p_3} < 0$	Accepteer H_3

De aldus uit T verkregen toets geven wij aan met N^T . In figuur 1 komt deze regel hierop neer dat we door O twee lijnen trekken, resp. evenwijdig aan DF en CE (vgl. de formules (1) t/m (4)), die drie gebieden begrenzen. Het gebied dat grenst aan de x-as (resp. y-as) leidt tot acceptatie van H_1 (resp. H_3), terwijl in het "middelste" gebied H_2 geaccepteerd wordt.

Het is duidelijk dat de onbetrouwbaarheidsdrempel van de toets door afknotting hoger komt te liggen.

Als we voor een toets T de bijbehorende lijnen met vergelijkingen (1) t/m (4) op millimeterpapier uitzetten, dan kunnen we met behulp van een tabel voor de binomiale verdeling [3] bij elke waarneming de kans, dat H_1 of H_3 geaccepteerd wordt, berekenen; dus ook $\beta_1(p, N^T_{12})$ en $\beta_3(p, N^T_{32})$ als T afgeknot wordt tot N^T . Uit (9) en (10) volgt dan dat de karakteristiek van N^T bekend is. Deze formules leveren ook de karakteristiek van de niet-afgeknotte toets, als $\beta_1(p, T_{12})$ en $\beta_3(p, T_{32})$ (zie WALD [2], blz. 51) berekend zijn.

Uit de voorgaande alinea volgt direct dat wij ook gemakkelijk $E(\underline{n}|p)$ voor de toets N^T kunnen berekenen, nl. door te bedenken dat de definitie luidt:

$$E(\underline{n}|p) = \sum_{n=0}^N P(\underline{n} = n|p) \cdot n, \quad (20)$$

waarin $P(\underline{n} = n|p)$ de kans voorstelt dat N^T bij de n^e waarneming eindigt, als p de kans op A is. Voor de toets T maken wij gebruik van (17), waarin we voor $E(\underline{n}^*|p)$ een soortgelijke formule invullen als de bovenstaande en $E(\underline{n}_{j2}|p)$ ($j = 1, 3$) berekenen als bij WALD [2], blz. 53.

Als voorbeelden van de invloed van afknotting op de karakteristiek en de waarde van $E(\underline{n}|p)$ beschouwen wij drie symmetrische toetsen $T^{(1)}$, $T^{(2)}$ en $T^{(3)}$ met afknotting bij $N = 25$ of $N = 35$, gedefinieerd door:

Tabel III
Numeriek berekende toetsen

Toets	p_1	p_2	α	N
$T^{(1)}$	0,20	0,40	0,05	25 resp. 35
$T^{(2)}$	0,10	0,40	0,05	25
$T^{(3)}$	0,15	0,45	0,05	25

Omdat wij hier met symmetrische toetsen te maken hebben, is het voor het tekenen van de karakteristiek van een toets $T^{(j)}$ voldoende om $\beta_1(p, T_{12}^{(j)})$ te kennen. Er geldt dan immers:

$$\beta_3(p, T_{32}^{(j)}) = \beta_1(1 - p, T_{12}^{(j)}) \quad (j = 1, 2, 3) \quad (21)$$

en uit (9) en (10) volgt dan dat $\beta_i(p, T)$ ($i = 1, 2, 3$) bekend zijn. Ook is het voor $E(\underline{n}|p)$ en $\gamma(p)$ voldoende de waarden van deze grootheden voor $p \leq \frac{1}{2}$ te kennen, want

$$\text{en} \quad \left. \begin{aligned} E(\underline{n}|\frac{1}{2} + d) &= E(\underline{n}|\frac{1}{2} - d) \\ \gamma(\frac{1}{2} + d) &= \gamma(\frac{1}{2} - d) \end{aligned} \right\} (0 \leq d \leq \frac{1}{2}). \quad (22)$$

We bepalen nu bij elke toets voor enkele waarden van p de bijbehorende $\beta_1(p)$, d.i. de kans om hypothese H_1 te accepteren en krijgen de volgende uitkomsten, die in tabel IV (pag. 12 a) vermeld staan.

Hierbij dient opgemerkt te worden, dat uit $\alpha = 0,05$ en Wald [2], blz. 95 volgt, dat

$$\beta_1(s) = \frac{\log A_{12}}{\log A_{12} - \log B_{12}} = 0,45 \text{ voor } T^{(j)}$$

$$(j = 1, 2, 3)$$

waarbij s gedefinieerd is door:

$$s = \frac{\log \frac{1-p_1}{1-p_2}}{\log \frac{p_2}{p_1} + \log \frac{1-p_1}{1-p_2}}.$$

We vinden voor s bij $T^{(1)}$, $T^{(2)}$ en $T^{(3)}$ resp.: 0,29, 0,23, 0,28.

Tabel IV
 $\beta_1(p)$ voor de toetsen van tabel III

Toets	p=0	0,10	0,15	0,20	0,23	0,28	0,29	0,30	0,35	0,40	0,45	0,50	0,55	0,60	
$T^{(1)}$	1			0,95			0,45			0,025					1
$_{25}T^{(1)}$	1	1,00		0,89				0,51		0,15		0,02		0,00	0
$_{35}T^{(1)}$	1	0,99		0,90				0,50		0,11		0,01		0,00	0
$T^{(2)}$	1	0,95			0,45					0,025					0
$_{25}T^{(2)}$	1	0,96		0,61				0,20		0,05		0,00			0
$T^{(3)}$	1		0,95			0,45					0,025				0
$_{25}T^{(3)}$	1	1,00	0,98		0,80				0,29		0,06		0,01		0

Uit tabel IV, tabel II en de definitie van α krijgen wij nu:

Tabel V
 $\beta(p)$ en α voor de toetsen van tabel III

Toets	$p=0$	0,10	0,15	0,20	0,40	0,45	α
$T^{(1)}$	0			0,05	0,025		0,05
$_{25}T^{(1)}$	0			0,11	0,15		0,15
$_{35}T^{(1)}$	0			0,10	0,11		0,11
$T^{(2)}$	0	0,05			0,025		0,05
$_{25}T^{(2)}$	0	0,04			0,05		0,05
$T^{(3)}$	0		0,05			0,025	0,05
$_{25}T^{(3)}$	0		0,02			0,07	0,07

Zoals uit deze tabellen en de met behulp daarvan ver-
vaardigde grafieken¹⁾ 1 t/m 4 blijkt, is $T^{(2)}$ de enige der drie
toetsen die ook na afknotting bij $N = 25$ voldoet aan: $\alpha = 0,05$.
De karakteristieken van $T^{(2)}$ en $_{25}T^{(2)}$ stemmen veel beter over-
een dan bij de andere toetsen het geval is. Bij $T^{(1)}$ is zelfs
 $N = 35$ niet voldoende om een behoorlijke overeenstemming te
krijgen. Zoals te verwachten was, is α kleiner naarmate de
indifferentiezônes groter zijn en bij 2 toetsen met even grote
indifferentiezônes is α het kleinste voor die toets, waarbij
deze zônes het verst van elkaar liggen.

Voor $E(\underline{n}|p)$ krijgen wij tabel VI (pag. 13 a).

De noemer van de formule voor $E(\underline{n}_{12}|p)$ (zie Wald [2], blz. 53)
wordt bij $T^{(1)}$, resp. $T^{(2)}$ en $T^{(3)}$ nul voor $p = 0,29$, resp.
0,23 en 0,28. Voor deze waarden berekenen wij $E(\underline{n}_{12}|p)$ zoals
aangegeven is bij [2], blz. 176.

We zien in de grafieken 5 t/m 7 dat, evenals bij de
karakteristiek, de overeenstemming tussen afgeknotte en niet-
afgeknotte toets het grootst is bij $T^{(2)}$, minder bij $T^{(3)}$ en
het slechtst bij $T^{(1)}$.

1) De grafieken staan op blz. 16a-16d.

Tabel VI
E(n|p) voor de toetsen van tabel III

Toets	p=0	0,10	0,15	0,20	0,23	0,25	0,28	0,29	0,30	0,40	0,45	0,50
$T^{(1)}$	13	20,48		35,36		56,88		53,93		27,50		23,09
$T_{25}^{(1)}$	13	18,44		22,77					23,49	21,87		20,64
$T_{35}^{(1)}$	13	19,22		27,58					30,22	25,81		22,09
$T^{(2)}$	9	15,78		18,40	18,36	18,03			16,57	12,82		11,04
$T_{25}^{(2)}$	9	15,21		17,82		17,13			15,53	12,16		10,83
$T^{(3)}$	9	13,91		18,76		23,80	24,53		24,03	18,08		15,02
$T_{25}^{(3)}$	9		15,95	17,82					18,64	16,11	15,01	14,41

7. De analoge klassieke toets

Wij beschouwen nu een toets T^* voor H_1 , H_2 en H_3 (zie figuur 1), gebaseerd op een constant aantal (N) waarnemingen. Zij m het aantal A's onder deze n waarnemingen, dan wordt T^* als volgt gedefinieerd:

$$T^* \begin{cases} \text{Accepteer } H_1, & \text{als } m < a_{12}N \\ & " \quad H_2, \quad " \quad a_{12}N \leq m \leq a_{32}N \\ & " \quad H_3, \quad " \quad m > a_{32}N. \end{cases}$$

De juistheid en onjuistheid van beslissingen wordt weer gedefinieerd volgens tabel I van § 3. De karakteristiek β^* wordt in dit geval gegeven door

$$\left. \begin{aligned} \beta_i^*(p) &= \sum_{v \in I_i} \binom{N}{v} p^v q^{N-v} \quad (i = 1, 2, 3) \\ &\quad \text{met} \\ I_1 &\equiv v < a_{12}N; \quad I_2 \equiv a_{12}N \leq v \leq a_{32}N \\ I_3 &\equiv v > a_{32}N \end{aligned} \right\} \quad (23)$$

De karakteristiek heeft een dergelijke vorm als de in figuur 3 geschetste; hetzelfde geldt t.o.v. figuur 4 voor de functie $\gamma^*(p)$ die de kans op het treffen van een onjuiste beslissing aangeeft. Dit alles is gemakkelijk na te gaan. De door (11), met $\gamma^*(p)$ in plaats van $\gamma(p)$ gedefinieerde onbetrouwbaarheidsdrempel geven wij met α^* aan. Deze wordt nu:

$$\alpha^* = \text{Max.} \{1 - \beta_1^*(p_1), 1 - \beta_2^*(p_1'), 1 - \beta_2^*(p_2''), 1 - \beta_3^*(p_3)\}. \quad (24)$$

In het symmetrische geval ($a_{12} + a_{32} = 1$, $p_1 + p_3 = p_2' + p_2'' = 1$) wordt (24)

$$\alpha^* = \text{Max.} \left\{ \sum_{v \geq a_{12}N} \binom{N}{v} p_1^v q_1^{N-v}, 1 - \sum_{a_{12}N \leq v \leq a_{32}N} \binom{N}{v} p_2^v q_2^{N-v} \right\}. \quad (25)$$

Al deze grootheden zijn functies van N . Voor $N \rightarrow \infty$ naderen β_i^* ($i = 1, 2, 3$) tot 1 en α^* tot 0.

Als voorbeelden nemen wij drie toetsen die overeenkomen met de in tabel III genoemde.

Hierbij is n met behulp van een tabel voor de binomiale verdeling of de normale benadering van de binomiale verdeling [3] uit (25) opgelost.

Voorts kunnen wij nu met (23) de karakteristiek van elke toets verkrijgen. We zien dan uit de grafieken 5 t/m 7 dat voor elke p geldt:

$$E(\underline{n}_{T(j)}|p) < N_{T(j)}^* \quad (j = 2, 3)$$

en dat dit voor $j=1$ voor bijna elke p geldt, slechts met uitzondering van waarden van p in de omgeving van 0,25 en 0,75.

Tabel VII

Klassieke toetsen, corresponderende met de toetsen van tabel III

Toets	p_1	a_{12}	p_2'	α^*	N
$T^{(1)*}$	0,20	0,30	0,40	0,05	55
$T^{(2)*}$	0,10	0,25	0,40	0,05	23
$T^{(3)*}$	0,15	0,30	0,45	0,05	26

Ten slotte bepalen wij N voor vier toetsen die overeenkomen met de afgeknotte sequente toetsen.

Tabel VIII

Klassieke toetsen, corresponderende met de afgeknotte sequente toetsen

Toets	p_1	a_{12}	p_2'	α^*	N
$_{25}T^{(1)*}$	0,20	0,30	0,40	0,15	25
$_{35}T^{(1)*}$	0,20	0,30	0,40	0,11	33
$_{25}T^{(2)*}$	0,10	0,25	0,40	0,05	23
$_{25}T^{(3)*}$	0,15	0,30	0,45	0,07	20

Hieruit blijkt dat geldt:

$$E(\underline{n}_{_{25}T(j)}|p) < N_{_{25}T(j)}^* = 25 \quad (j = 1, 2, 3)$$

en analoog voor $_{35}T^{(1)}$ en $_{35}T^{(1)*}$. Voor elke waarde van p vergt de afgeknotte sequente toets dus gemiddeld minder waarnemingen bij gelijke α dan de corresponderende klassieke toets.

8. Conclusies

Zowel bij de beschreven al of niet afgeknotte sequente toets als bij de klassieke toets geldt dat bij gegeven schema's de onbetrouwbaarheidsdrempel hoger ligt naarmate de indifferenzênes kleiner zijn en meer naar het midden van de p -as lig-

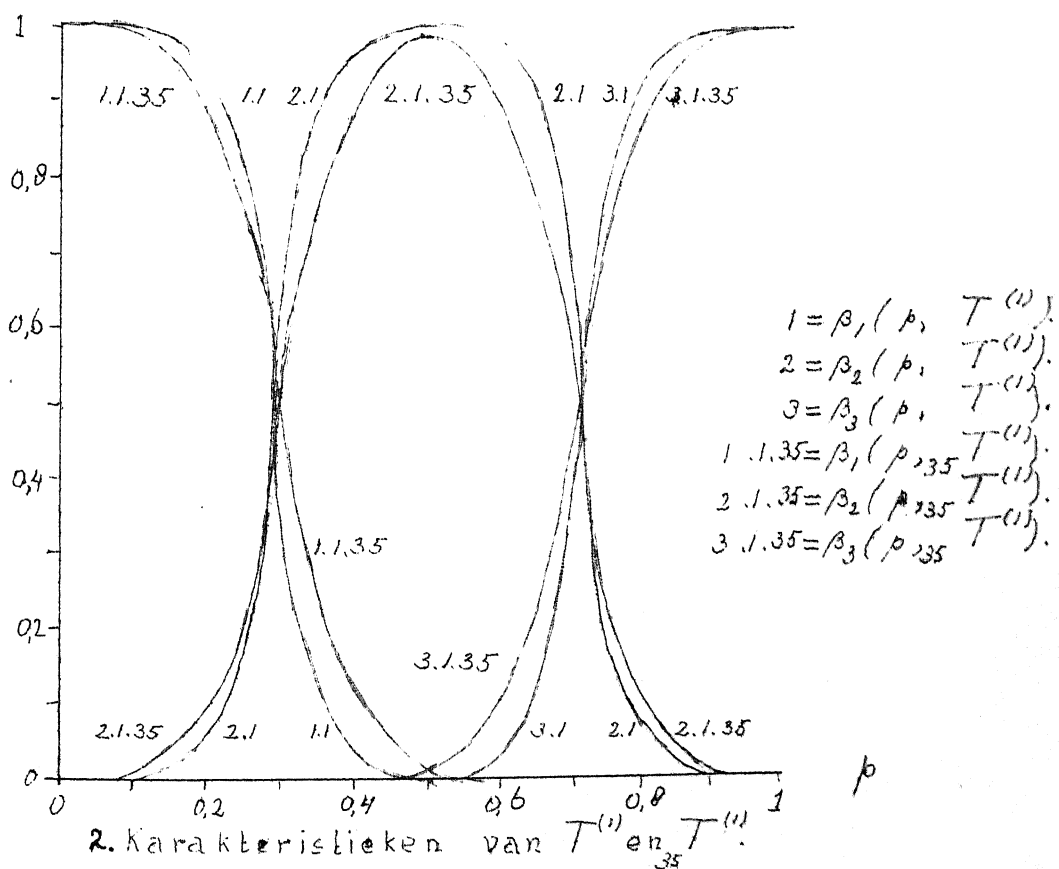
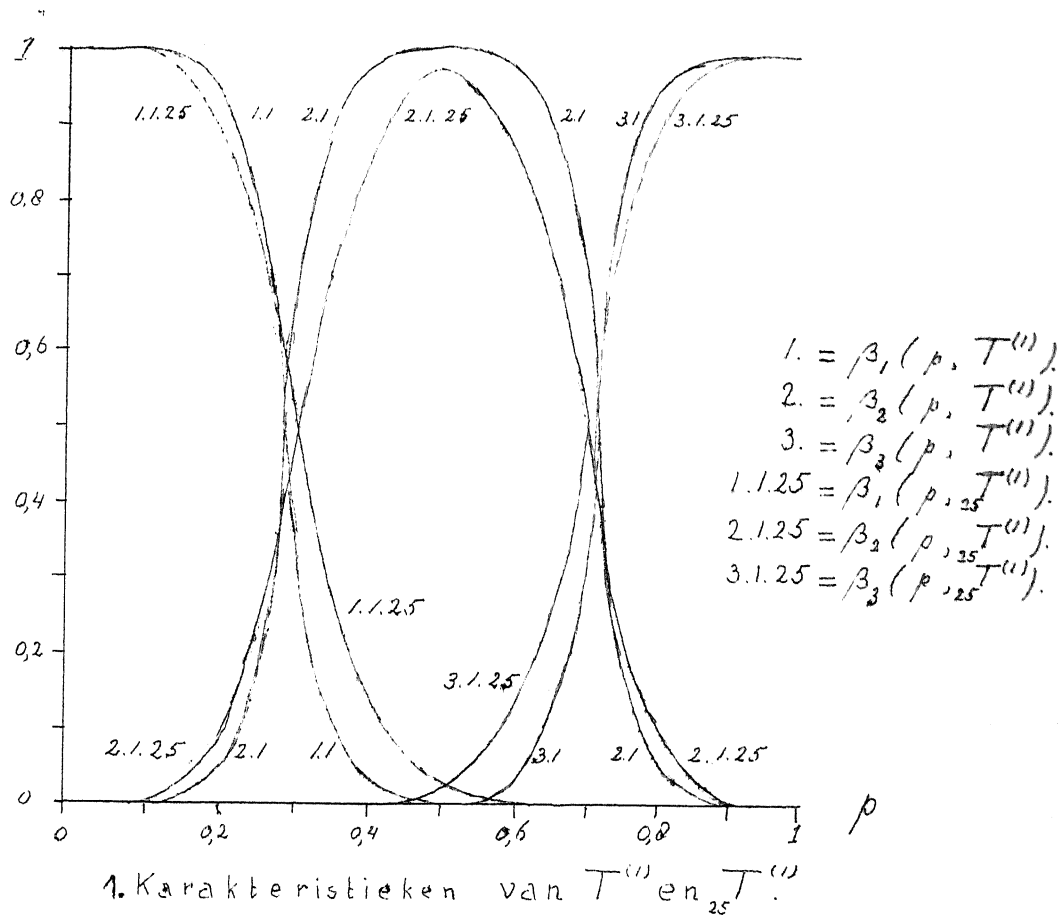
gen wat intuïtief direct duidelijk is (p is de kans op één der twee mogelijke uitkomsten).

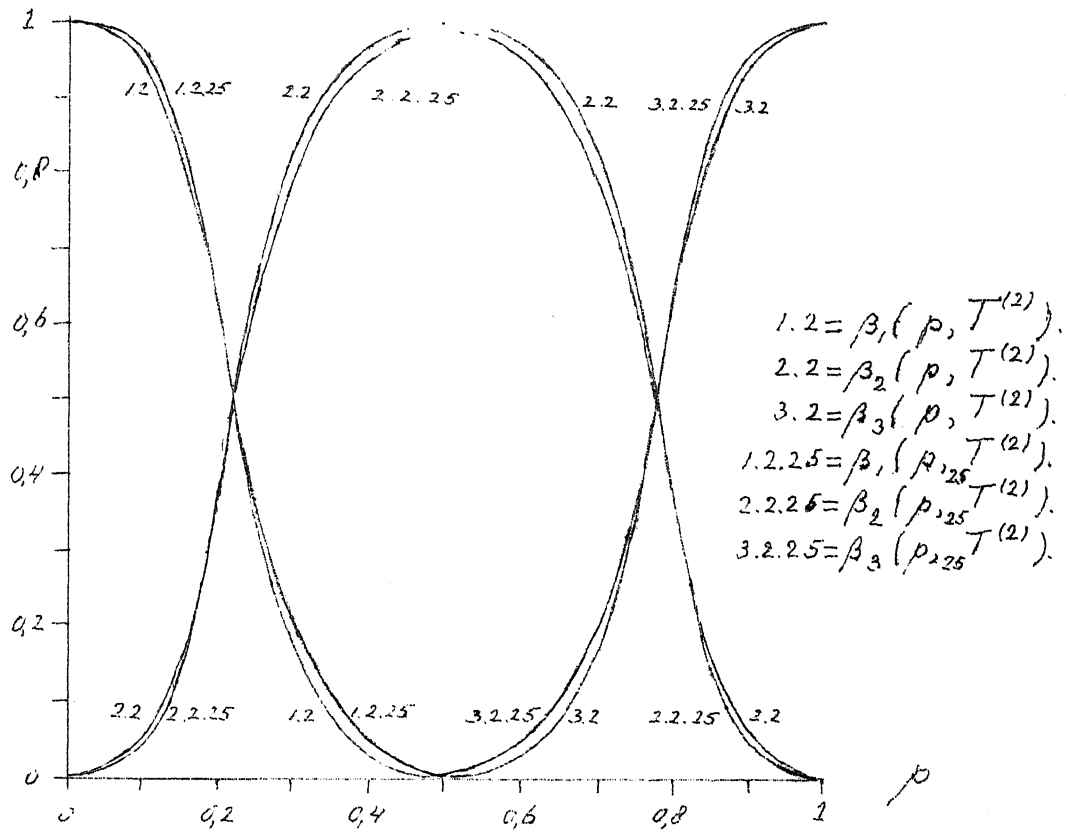
Bij vergelijking van de sequente en klassieke toetsen aan de hand van enige voorbeelden blijkt dat als regel de niet afgeknotte toets bij gelijke onbetrouwbaarheidsdrempel gemiddeld minder waarnemingen vergt dan de corresponderende klassieke toets, behalve bij toetsen waarbij de indifferentiezônes betrekkelijk klein zijn en bovendien de werkelijke waarde van P ongeveer midden in een indifferentiezône ligt. Voor de afgeknotte sequente toets geldt t.o.v. de klassieke toets dezelfde regel, nu zonder bovengenoemde uitzondering.

Dit betekent dus dat afgeknotte sequente toetsen uniform doeltreffender zijn dan klassieke toetsen met dezelfde onbetrouwbaarheidsdrempels.

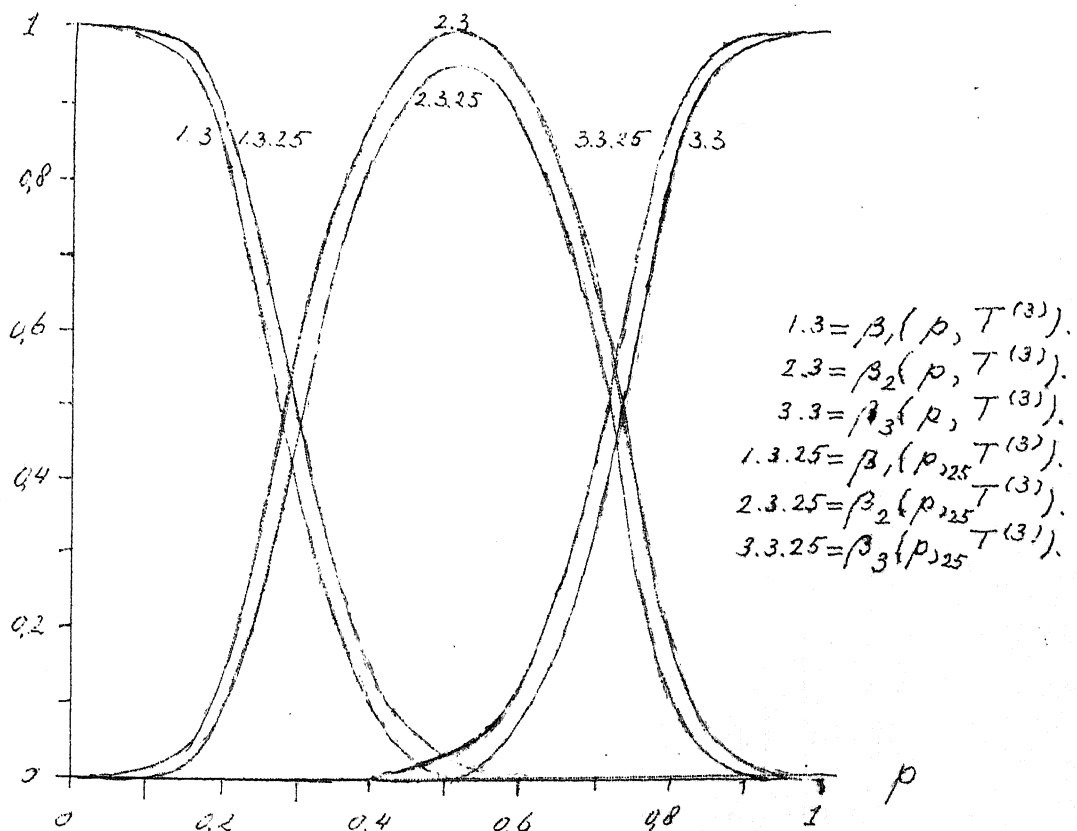
Litteratuur:

- [1] M.Sobel en A.Wald, A sequential decision procedure for choosing one of three hypotheses concerning the unknown mean of a normal distribution.
The annals of mathematical statistics, 20 (1949), 502.
- [2] A.Wald, Sequential analysis, John Wiley & Sons, Inc.
N.Y. (1947).
- [3] Tables of the binomial probability distribution, National Bureau of Standards, Applied Mathematics Series, 6 (1949).

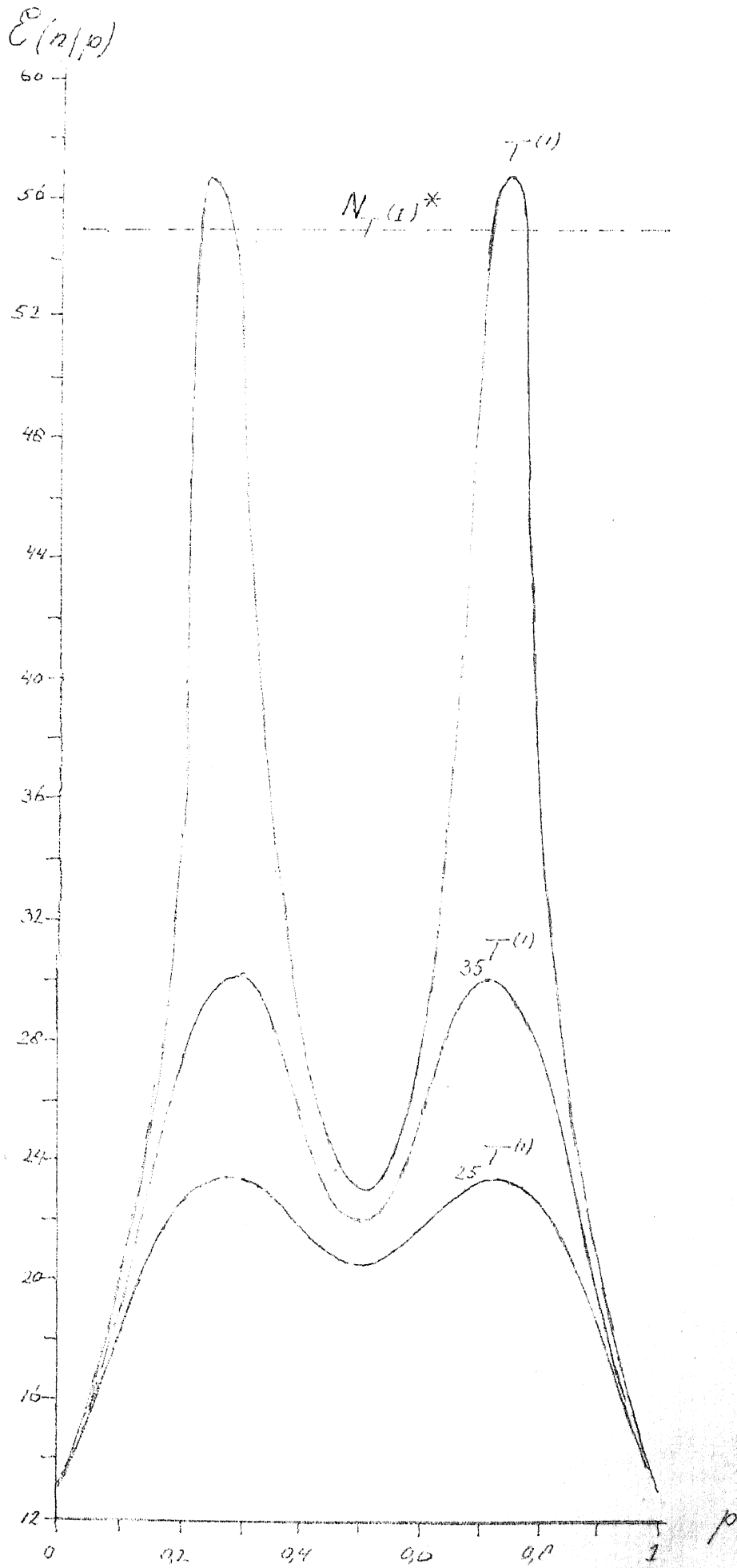




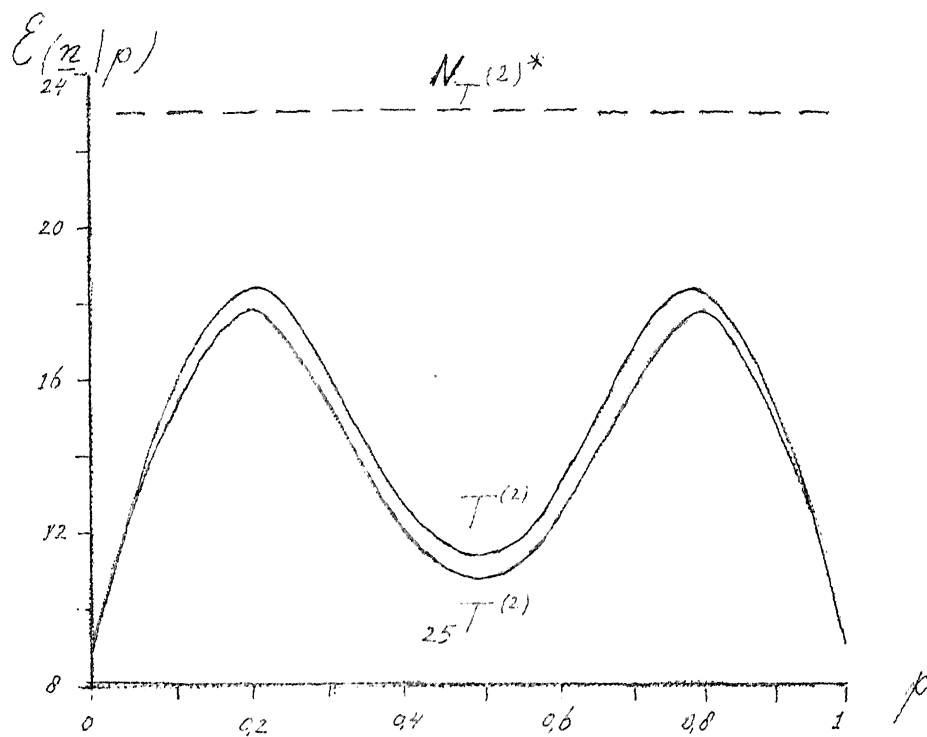
3. Karakteristieken van $T^{(2)}$ en $T_{25}^{(2)}$.



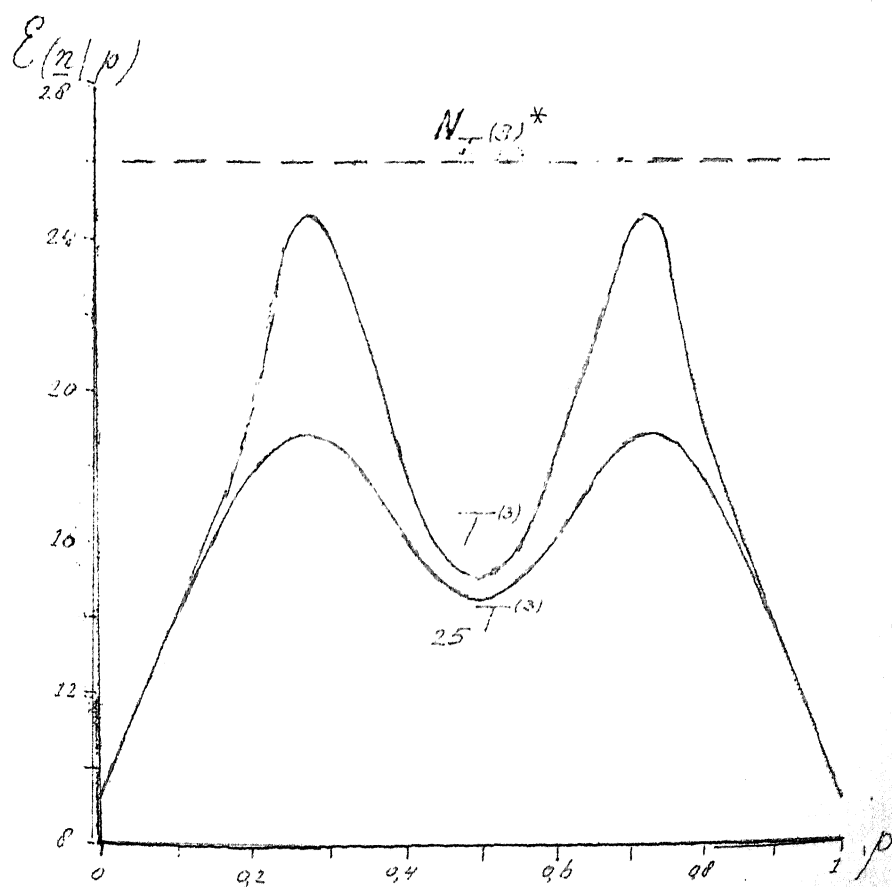
4. Karakteristieken van $T^{(3)}$ en $T_{25}^{(3)}$.



5. $E(n/p)$ van $T_{25}^{(1)}$, $T_{35}^{(1)}$ en $T^{(1)}$ in vergelijking met $N_{T(1)}^*$



6. $E(\frac{n}{p})$ van $T^{(2)}$ en $25T^{(2)}$ in vergelijking met $N_{T(2)}^*$



7. $E(\frac{n}{p})$ van $T^{(3)}$ en $25T^{(3)}$ in vergelijking met $N_{T(3)}^*$