

Stichting
MATHEMATISCH CENTRUM
2e Boerhaavestraat 49
Amsterdam

TW 3

Voorlopig Rapport over Steekproefschema's.

Dr. B.L. van der Waerden

februari 1949

MATHEMATISCH CENTRUM.

Afd. Toegepaste Wiskunde.

Rapport nr.3

Voorlopig rapport over steekproefschema's.

Auteur: Dr. B.L. van der Waerden.

Datum: Februari 1949.

5. 1. Algemeen begrippen.

Notaties

f = aantal fouten in een steekproef van n
 $g = n - f$ = aantal goede exemplaren
 q (qualiteit) = kans dat één stuk goed is
 $p = 1 - q$ = kans dat één stuk fout is

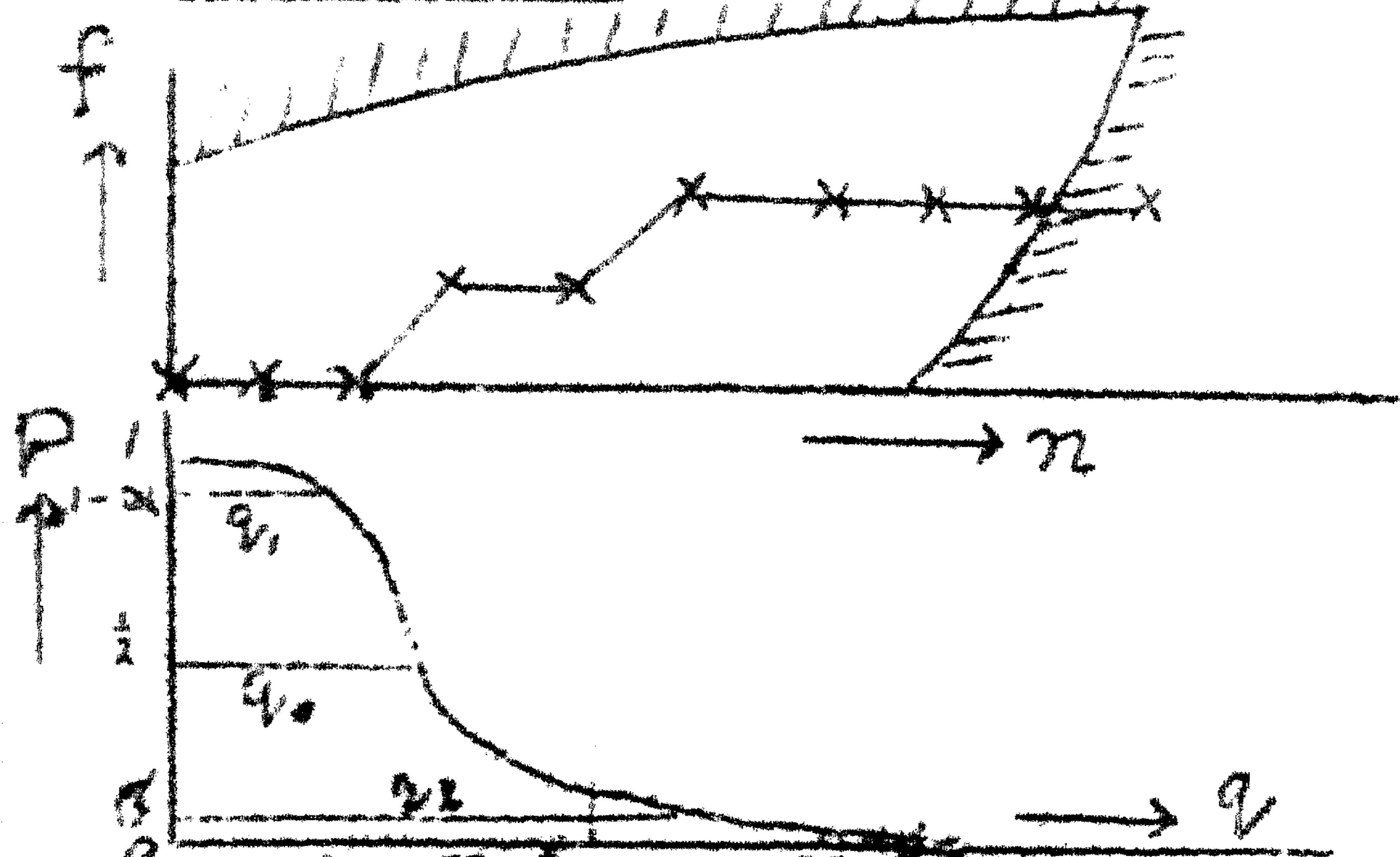
Steekproefschema = regel voor accepteren of verwerpen van de partij op grond van de steekproef.

Toevalspad: Zet f naar boven, n naar rechts uit. Bij elke gevonden

fout gaat het toevalspad één stap schuin naar boven, bij elk goed exemplaar één stap naar rechts, totdat de grens van het schema bereikt is en de partij wordt verworpen of geaccepteerd.

Werkkromme (OC = Operating Characteristics). $P(q)$ = kans dat steekproef van kwaliteit q wordt geaccepteerd.

$Q = 1 - P$ = kans dat steekproef van kwaliteit q wordt verworpen.



Kentallen van de werkkromme:

A. Volgens Wald: $q_{1-\alpha} = q_{\beta} = \theta_0$ (Wald's notatie) en $q_1 = q_2 = \theta_1$, behorende bij gegeven kansen $1 - \alpha$ en β . Kiest men b.v. $1 - \alpha = .95$ en $\beta = .05$, dan betekent dit dat men eist dat een partij van kwaliteit q_1 slechts in 5% van alle gevallen afgekeurd wordt, en dat een partij met kwaliteit q_2 slechts in 5% van alle gevallen goedgekeurd wordt. De getallen q_1 en q_2 bepalen de "strength" van de "test" volgens Wald.

B. Volgens Hamaker zijn de kentallen van een werkkromme: $q_0 = q_{\frac{1}{2}}$ behorende bij $P = .5$ en

$$(1) \quad s = - \left(\frac{dP}{dq} \right)_0 \quad \text{in hetzelfde punt.}$$

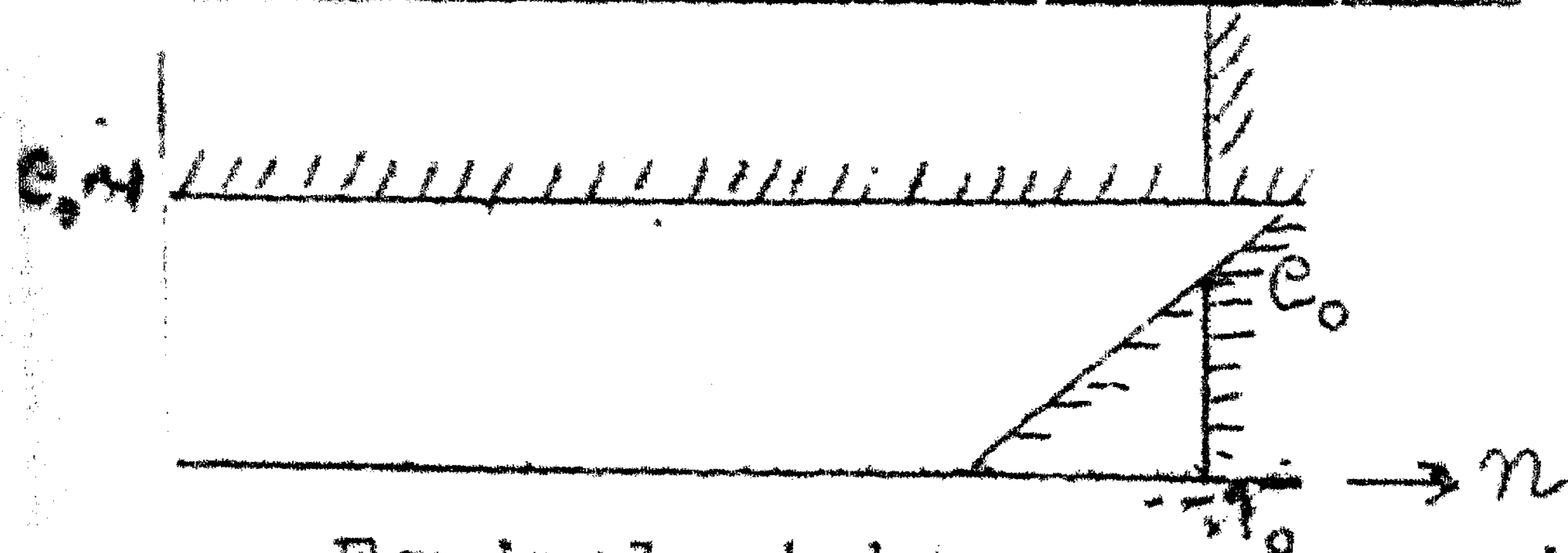
Naderen de $1 - \alpha$ en β van Wald beide tot $\frac{1}{2}$, dan naderen q_1 en q_2 beide tot q_0 en het quotient

$$\frac{q_{\beta} - q_{1-\alpha}}{(1-\alpha) - \beta}$$

nadert tot s . De Ha-kentallen kunnen dus als grensgevallen van de Wa-kentallen worden beschouwd. De Ha-kentallen hebben het voordeel dat zij tot eenvoudiger formules leiden.

Twee schema's heten Ha-equivalent als ze dezelfde q_0 en s hebben. Ze heten Wa-equivalent (of bij Wald: of equal strength) als ze dezelfde q_1 en q_2 hebben (bij gegeven α en β). Ze heten streng equivalent, als de werkkrommen elkaar dekken.

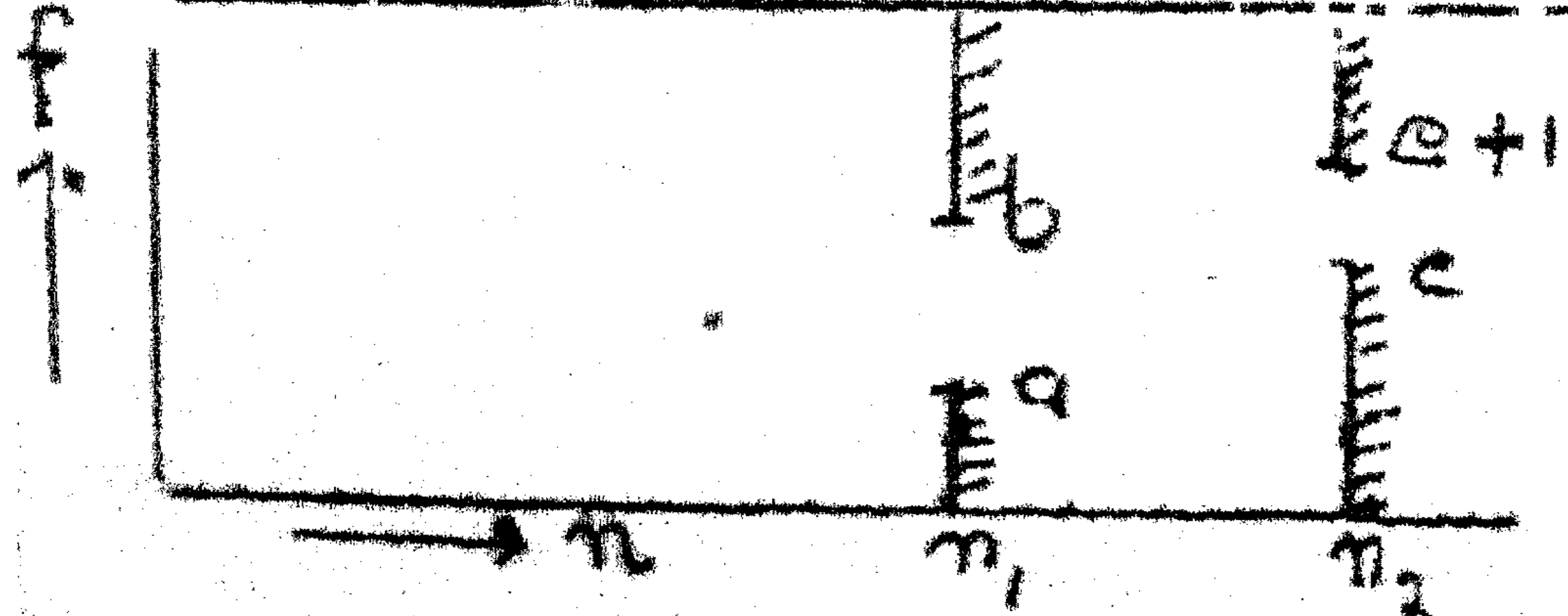
Enkelvoudig steekproefschema:



De omvang van de steekproef is constant: $n = n_0$. De gebiedsgrens is dus een verticale rechte lijn, waarop een punt (n_0, c_0) gegeven is. Een steekproef van n_0 met meer dan c_0 fouten wordt verworpen, met hoogstens c_0 fouten aangenomen.

Equivalent hiermee, maar iets efficiënter is een schema, begrensd door een horizontale lijn $f = c_0 + 1$ en een lijn onder $\frac{1}{2}$ door het punt (n_0, c_0) . Men kan immer met de keuring ophouden zodra men $c_0 + 1$ fouten of $n_0 - c_0$ goede exemplaren heeft.

Dubbel steekproefschema:



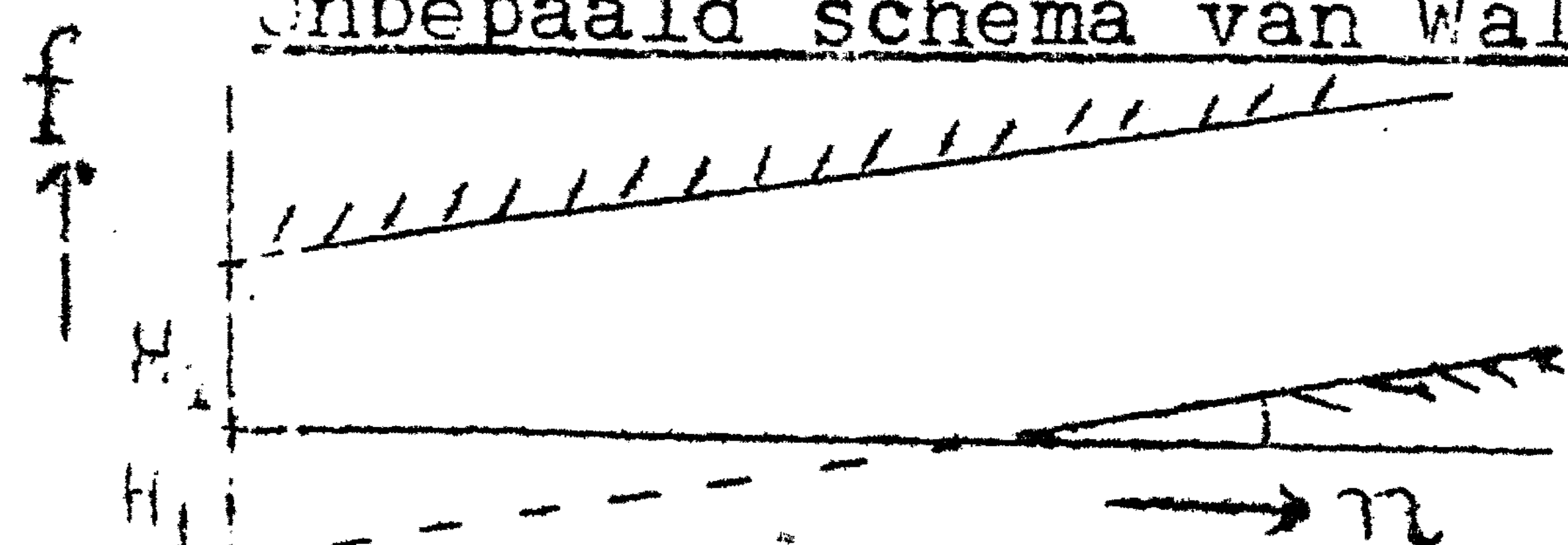
Men neemt eerst een steekproef van n_1 stuks. Zijn er hoogstens a fouten in, dan wordt de partij goedgekeurd; zijn er b of meer fouten in, dan afgekeurd. Ligt het aantal fouten tussen a en b

dan wordt de steekproef tot n_2 stuks uitgebreid, en bij meer dan c fouten de partij afgekeurd, anders goedgekeurd.

Bij gelijke scherpte is dit schema meestal efficiënter dan het enkelvoudige: men komt gemiddeld met een kleinere steekproef uit.

Nog efficiënter is het

Onbepaald schema van Wald:



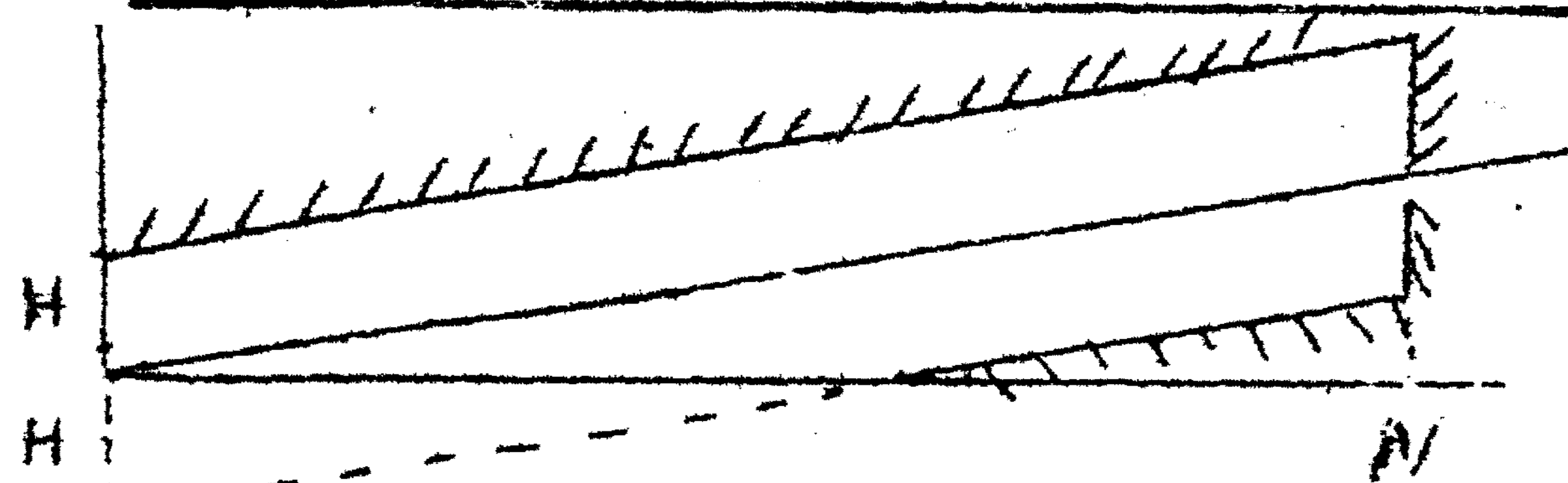
Het gebied wordt begrensd door 2 evenwijdige lijnen ter hoogte $-H$ en H (of algemener $-H_1$ en $+H_2$) met richtingscoëfficiënt q_0 , dus met vergelijkingen

$$(2) \quad f = -H + q_0 n \quad (\text{goedkeurlijn})$$

$$(3) \quad f = H + q_0 n \quad (\text{afkeurlijn})$$

De proef wordt onbepaald voortgezet tot men één der lijnen treft. De kans dat dit nooit gebeurt is nul.

Afgebroken onbepaald schema:



Als voren, maar de test wordt bij maximaal N stuks afgebroken, en de partij wordt verworpen of geaccepteerd alnaar-

mate $f > Nq_0$ of $f \leq Nq_0$ uitvalt.

Bij elk schema, waarvan de werkkromme bekend is, hoort een Ha-equivalent enkelvoudig schema, dat dezelfde kentallen q_0 en s heeft (tenminste als men ook niet-gehele waarden van n en c_0 in het enkelvoudige schema toelaat en de bijbehorende kentallen q_0 en s interpoleert).

Op dezelfde wijze behoort ook bij elk schema een Wa-equivalent enkelvoudig schema, dat dezelfde kentallen q_1, α en q_2, β heeft, waarbij α en β nog willekeurig kunnen worden gekozen.

In de limiet voor $\alpha = \beta = \frac{1}{2}$ gaat het Wa-equivalente in het Ha-equivalente enkelvoudige schema over.

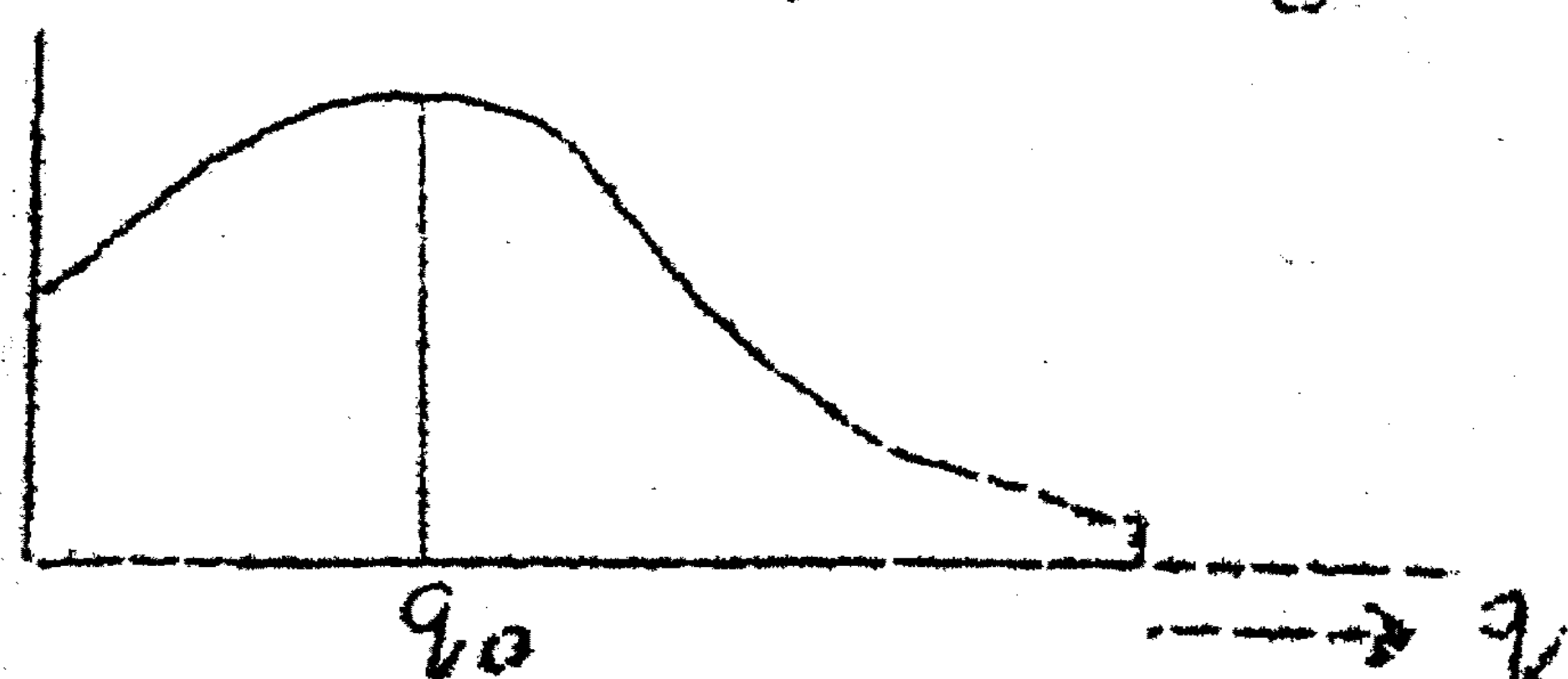
Wanneer de werkkromme van een schema niet al te erg van die van een enkelvoudig schema afwijkt, zullen wij ter wille van de eenvoud ons tot het equivalentiebeprip van Hamaker beperken. Hamaker heeft empirisch gevonden dat sterke afwijkingen alleen dan optreden als het steekproefschema sterk asymmetrisch is t.o.v. de lijn $f = q_0 n$. Zolang men b.v. in het schema van Wald $H_1 = H_2$ kiest, valt de werkkromme vrij nauwkeurig met die van een enkelvoudig schema samen. Kiest men echter $H_1 \neq H_2$, dan krijgt men afwijkende werkkrommen. Maar deze sterk asymmetrische schema's blijken zeer inefficiënt te zijn.

De gemiddelde omvang $\hat{n}$ van de steekproef is

$$(4) \quad \hat{n} = E(n) = \sum n P_n$$

Hierin is P_n de kans dat na n stappen de beslissing valt.

$\hat{n}$ is een functie van q , die gewoonlijk in de buurt van q_0 een maximum heeft (zie figuur). Dit is ook te verwachten: Als de partij juist de twijfelqualiteit q_0 heeft, duurt het gewoonlijk langer voor de beslissing valt dan wanneer de partij uitgesproken goed of slecht is. - Hoe lager de kromme en vooral het maximum ligt, des te efficiënter is het steekproefschema.

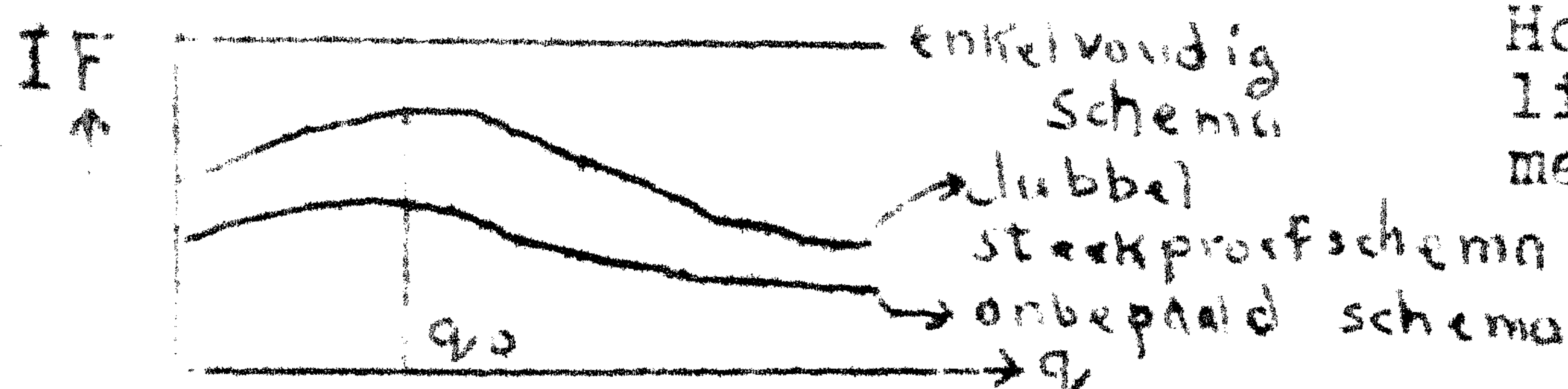


Efficiency-quotienten. Om de efficiency door een factor tussen 0 en 1 uit te drukken, kan men 2 wegen inslaan: men kan volgens Wald voor bepaalde waarden van q (b.v. q_1 en q_2) de $\hat{n}$ vergelijken met de kleinste mogelijke $\hat{n}$, of men kan volgens Hamaker de $\hat{n}$ vergelijken met de n_0 van de equivalente enkelvoudige steekproef. We bespreken eerst de tweede mogelijkheid.

De Inspectiefactor IF wordt als functie van q gedefinieerd door

$$IF = \frac{\hat{n}}{n_0}$$

Hierin is n de omvang van de Ha-equivalente enkelvoudige steekproef.



Hoe beter het schema, des te lager ligt de IF-kromme. Het maximum ligt meestal nabij $q = q_0$.

De efficiency volgens Wald:

Wald heeft bewezen dat $\hat{n}(q)$ aan de volgende ongelijkheid voldoet)

$$(5) \quad \hat{n}(q) \geq \frac{p(q) \ln \frac{\beta}{1-\alpha} + q(q) \ln \frac{1-\beta}{\alpha}}{p \ln p_2/p_1 + q \ln q_2/q_1}$$

In 't bijzonder geldt dus voor $q = q_1$

$$(6) \quad \hat{n}(q_1) \geq \frac{(1-\alpha) \ln \frac{\beta}{1-\alpha} + \alpha \ln \frac{1-\beta}{\alpha}}{p_1 \ln p_2/p_1 + q_1 \ln q_2/q_1}$$

en voor $q = q_2$

$$(7) \quad \hat{n}(q_2) \geq \frac{\beta \ln \frac{\beta}{1-\alpha} + (1-\beta) \ln \frac{1-\beta}{\alpha}}{p_2 \ln p_2/p_1 + q_2 \ln q_2/q_1}$$

Hierin is $p_1 = 1 - q_1$ en $p_2 = 1 - q_2$ gesteld.

Er zijn dus bij gegeven α, β, q_1 en q_2 onderste grenzen van $\hat{n}(q_1)$ en $\hat{n}(q_2)$, waar men nooit beneden kan komen. Laat deze onderste grenzen n_1 en n_2 zijn. De efficiency van elk schema kan men dus beoordeelen naar de quotiënten

$$(8) \quad \frac{\hat{n}(q_1)}{n_1} \quad \text{en} \quad \frac{\hat{n}(q_2)}{n_2}$$

Hoe lager deze zijn, des te beter is het steekproefschema.

Laat men α en β tot 0,5 naderen, dan naderen q_1 en q_2 beide tot q_0 , en men vindt in de limiet

$$\hat{n}(q) = -2 \left(\frac{dP}{dq} \right)_0 \frac{p-q}{p_0 - q_0}$$

Laat men q ook nog tot q_0 naderen, dan vindt men

$$(9) \quad \hat{n}(q_0) \geq 4 p_0 q_0 \left(\frac{dP}{dq} \right)_0^2 = 4 p_0 q_0 s^2$$

De verkwistingsfactor. Tot een zeer eenvoudige maat voor de efficiency van een steekproefschema komt men, door te stellen

$$(10) \quad V = \frac{\hat{n}(q_0)}{4 p_0 q_0 s^2}$$

Volgens (9) is V altijd ≥ 1 . V is de gemiddelde omvang van de steekproef gedeeld door de kleinst mogelijke gemiddelde omvang van een Ha-equivalent schema in het kritieke geval $q = q_0$, dus voor partijen die juist op de grens tussen goed- en afkeuren staan.

We noemen V de verkwistingsfactor. We zullen zien dat voor het enkelvoudige steekproefschema V nagenoeg gelijk is aan

$$\frac{\pi}{2} = 1.57$$

en dat voor het schema van Wald de factor V vlak bij 1 ligt.

§2. Het enkelvoudige schema.

Voor een enkelvoudige steekproef van n stuks met criterium c (waarbij de partij dus afgekeurd wordt als er meer dan c fouten zijn) geldt:

$$(11) \quad P = p^n + \binom{n}{1} p^{n-1} q + \dots + \binom{n}{c} p^{n-c} q^c$$

De afgeleide is

$$(12) \quad - \frac{dP}{dq} = \frac{dP}{dp} = \binom{n}{c} (n-c) p^{n-c-1} q^c \\ = n \frac{(n-1)!}{c! (n-c-1)!} p^{n-c-1} q^c$$

De som P kan bij benadering door een integraal van een Gausskromme worden voorgesteld. Neemt men nog de correctietermen van de orde $n^{-1/2}$ voor de scheefheid van de Gausskromme in aanmerking, en zoekt men de waarde $q = q_0$, waarvoor P juist $\frac{1}{2}$ wordt, dan vindt men met goede benadering

$$(13) \quad (n + \frac{1}{3}) q_0 \sim c + \frac{2}{3}$$

Voor grote n en niet te grote c gaat (13) in de door Dr. Hamaker gebruikte formule

$$n q_0 = c + 0,67$$

over.

Deze waarde van q_0 ligt vlak bij het maximum van de functie (12). In de buurt van het maximum verandert de functie niet sterk; wij kunnen dus als goede benadering voor (12) het maximum zelf nemen, dat bij $q = \frac{c}{n-1}$ ligt, en vinden

$$(14) \quad s = - \left(\frac{dP}{dq} \right)_{q_0} \sim \sqrt{\frac{n}{2\pi p_0 q_0}}$$

Als men p_0 door 1 vervangt en 2π door 6,25, dan komt men op de formule van Hamaker

$$n \sim 6,25 q_0 s^2$$

Formules (13) en (14) kunnen worden gebruikt om bij gegeven s en q_0 de constanten $n = n_0$ en $c = c_0$ van de equivalente enkelvoudige steekproef te bepalen:

$$(15) \quad n_0 \sim 2\pi p_0 q_0 s^2$$

$$(16) \quad c_0 \sim (n_0 + \frac{1}{3}) q_0^{-1/2}$$

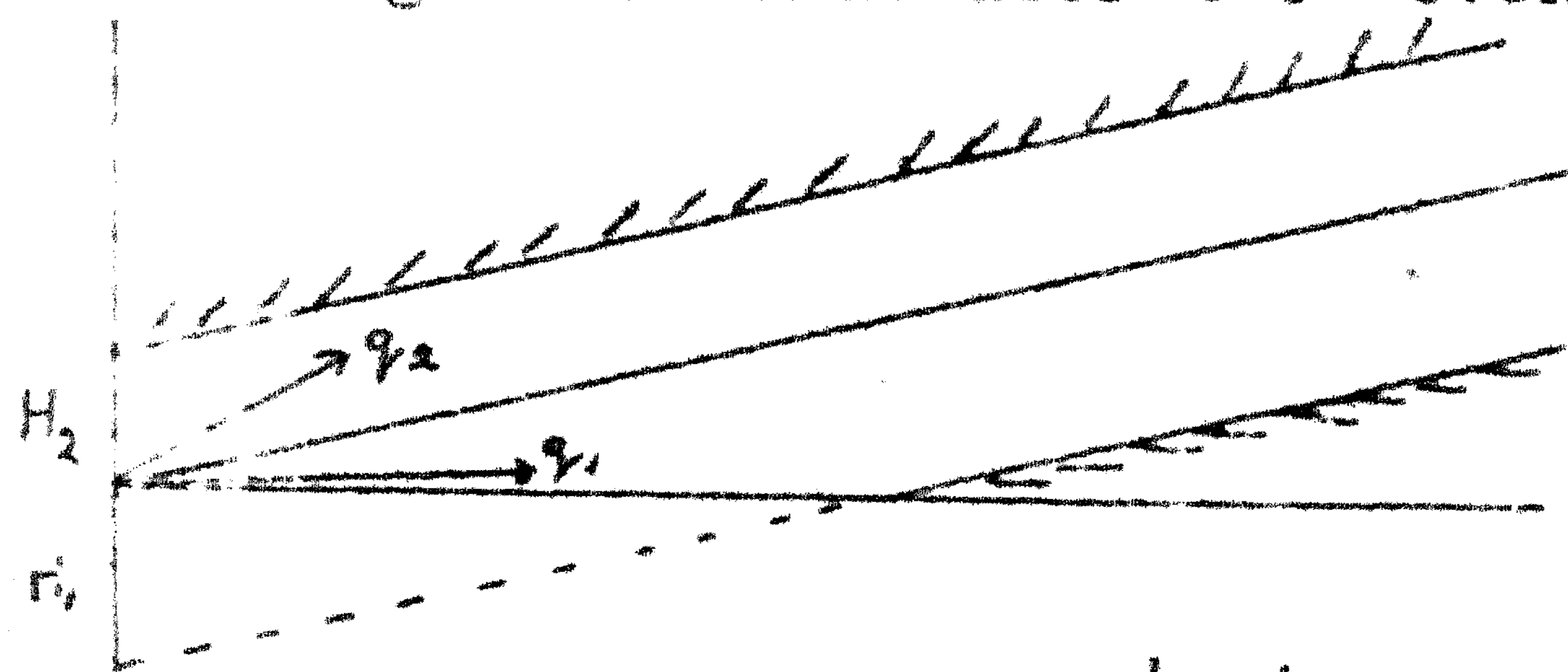
Formule (15) betekent dat de verkwistingsfactor V van het enkelvoudige steekproefschema ongeveer

$$\frac{2\pi p_0 q_0 s^2}{4 p_0 q_0 s^2} = \frac{\pi}{2} = 1,57$$

bedraagt.

§3. Het onbepaald schema van Wald.

In een fundamentele verhandeling¹⁾ heeft A.Wald, uitgaande van algemene waarschijnlijkheidsprincipes, een steekproefschema ontworpen, dat begrensd wordt door twee evenwijdige rechte lijnen.



$f = -H_1 + nq_0$ (goedkeurlijn)
 $f = H_2 + nq_0$ (afkeurlijn)
 H_1, H_2 en q_0 worden zo bepaald dat de kans op afkeuring van een partij met kwaliteit q_1 (of beter) hoogstens α is, en de kans op goedkeuring van een partij met kwaliteit q_2 (of erger) hoogstens β . Dit bereikt men door te stellen

$$(17) \quad H_1 = \frac{\log \frac{1-\alpha}{\beta}}{\log \frac{q_2}{q_1} - \log \frac{p_2}{p_1}}$$

$$(18) \quad H_2 = \frac{\log \frac{1-\beta}{\alpha}}{\log \frac{q_2}{q_1} - \log \frac{p_2}{p_1}}$$

$$(19) \quad q_0 = \frac{\log \frac{1-\beta}{\alpha}}{\log \frac{q_2}{q_1} - \log \frac{p_2}{p_1}}$$

Nemen we $\alpha = \beta$ (b.v. = 0,05), dan wordt $H_1 = H_2$.

De werkkromme wordt gedefinieerd door de parameter voorstelling

$$(20) \quad q = \frac{1 - \left(\frac{1-q_2}{1-q_1} \right)^h}{\left(\frac{q_2}{q_1} \right)^h - \left(\frac{1-q_2}{1-q_1} \right)^h} \quad (h \text{ is de parameter})$$

$$(21) \quad P(q) = \frac{\left(\frac{1-\beta}{\alpha} \right)^h - 1}{\left(\frac{1-\beta}{\alpha} \right)^h - \left(\frac{\beta}{1-\alpha} \right)^h}$$

Neem nu weer $\alpha = \beta$ aan, dus $H_1 = H_2 = H$. Als h tot nul nadert, nadert $P(q)$ tot $\frac{1}{2}$ en q tot q_0 . Het eerste Ha-kental van het schema van Wald is dus juist q_0 . Het tweede Ha-kental vindt men door differentiatie:

$$(22) \quad s = - \left\{ \frac{dP}{dq} \right\}_0 = - \frac{H}{2 q_0 p_0}$$

De gemiddelde steekproefgrootte $\hat{n}$ is in goede benadering

$$(23) \quad \hat{n}(q) \sim \frac{P(q) \log \frac{1-\beta}{1-\alpha}}{q \log \frac{q_2}{q_1}} + \frac{Q(q) \log \frac{1-\beta}{\alpha}}{(1-q) \log \frac{p_2}{p_1}}$$

Voor $q = q_1$ en $q = q_2$ gaat het rechterlid van (23) juist in dat van (6) en (7) over. Het schema van Wald heeft dus van alle Wa-equivalente schema's bijna de laagst mogelijke $\hat{n}(q_1)$ en $\hat{n}(q_2)$. Wald vermoedt, dat zijn schema zelfs absoluut de laagst mogelijke $\hat{n}(q_1)$ en $\hat{n}(q_2)$ heeft.

Neem nu weer $\alpha = \beta$ aan, het belangrijkste geval. Laat men, bij onveranderde H , α en β beide tot 0,5 naderen, en stelt ook $q = q_0$, dan wordt de ingewikkelde uitdrukking (23) zeer veel eenvoudiger. De limiet van (23) voor $q = q_0$ is namelijk, zoals we onder (9) al berekend hebben:

$$(24) \quad \hat{n}(q_0) \sim 4 p_0 q_0 s^2$$

of volgens onze definitie van de verkwistingsfactor V .

$$V = \frac{\hat{n}(q_0)}{4 p_0 q_0 s^2} \sim 1$$

1) A.Wald, Sequential tests of statistical hypotheses, Ann.math.Statistics 16 (1945) p.117-186. Zie ook J.Amer, Statist.Assoc.40 (1945) p.277 en het boek van A.Wald, Sequential Analysis.

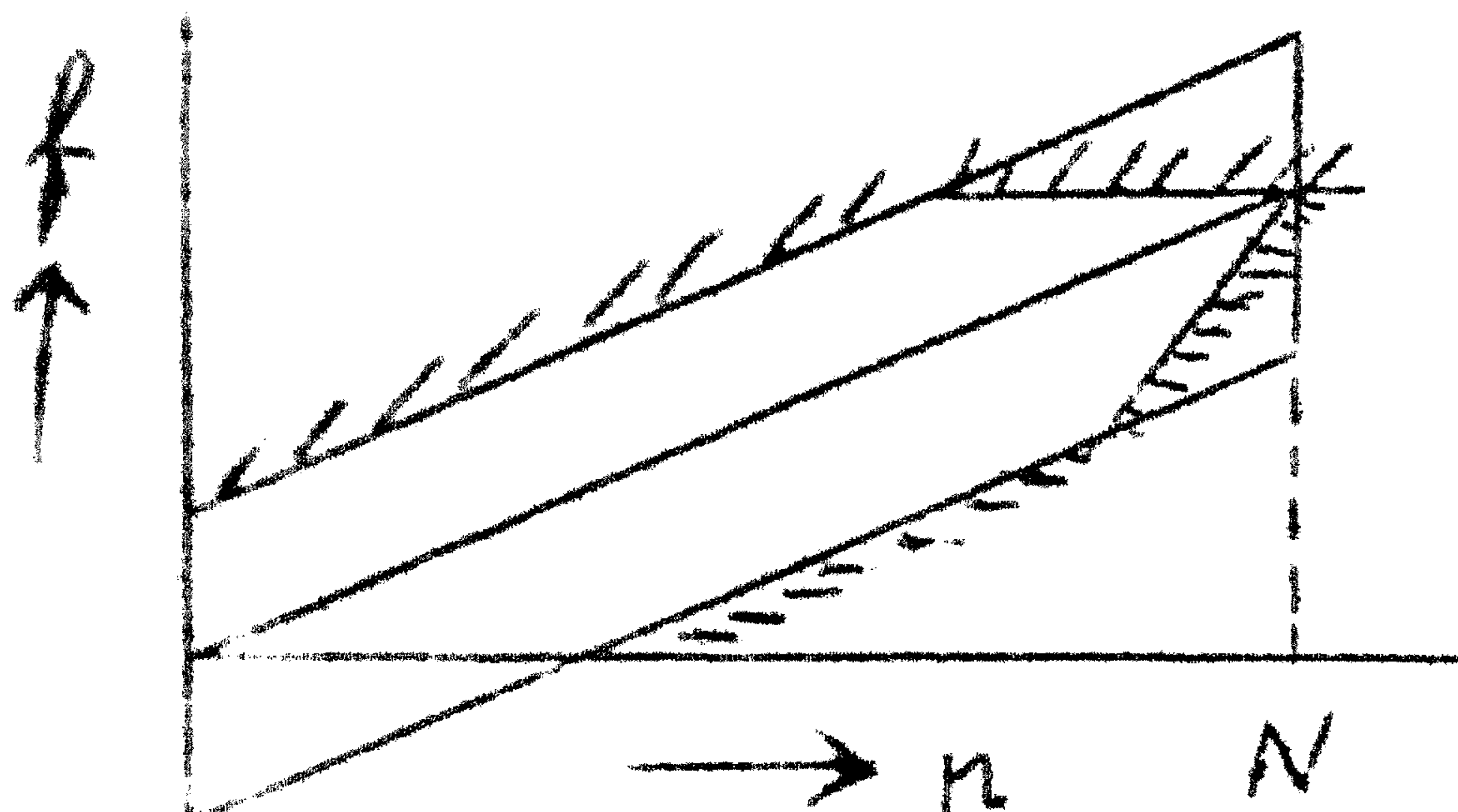
§ 4. Program voor verder onderzoek.

In de tekeningen bij § 4 zullen wij naar boven het aantal fouten f , naar rechts het aantal onderzochte exemplaren $n=f+g$ afzetten.

De bedoeling is voor verschillende vormen van grenslijnen te berekenen:

1. de werkkromme $P=P(q)$
 2. de kentallen c_0 en n_0 van de equivalente enkelvoudige steekproef (in de zin van Hamaker).
 3. de inspectiefactor $\hat{1}F = \hat{n}/n_0$ als functie van q
 4. de verdelingsfunctie van n in het ongunstigste geval $q=q_0$
- en door vergelijking de gunstigste vorm van de grenslijnen te bepalen. Het schijnt praktisch dit program in twee delen te splitsen al naar de vorm van de aangenomen grenslijnen.

A. Potlood en parallellogram



Wanneer men de onbepaalde steekproef bij een steekproefgrootte N afbreekt en dan afkeurt of goedkeurt alnaarmate f groter of kleiner dan $q_0 N$ is, dan krijgt men een gebied in de vorm van een parallellogram.

Men kan natuurlijk ook al ophouden zodra men meer dan $q_0 N$ fouten of meer dan $(1-q_0)N$ goede exemplaren heeft gevonden. Het gebied krijgt dan de vorm van een potlood. Dit schema is equivalent met het vorige.

De inspectiefactor is voor het potlood iets kleiner dan voor het parallellogram. Het verschil zal wel gemakkelijk te schatten zijn. Voor de berekening is het parallellogram eenvoudiger. Laten we ons dus voorlopig tot het parallellogram beperken.

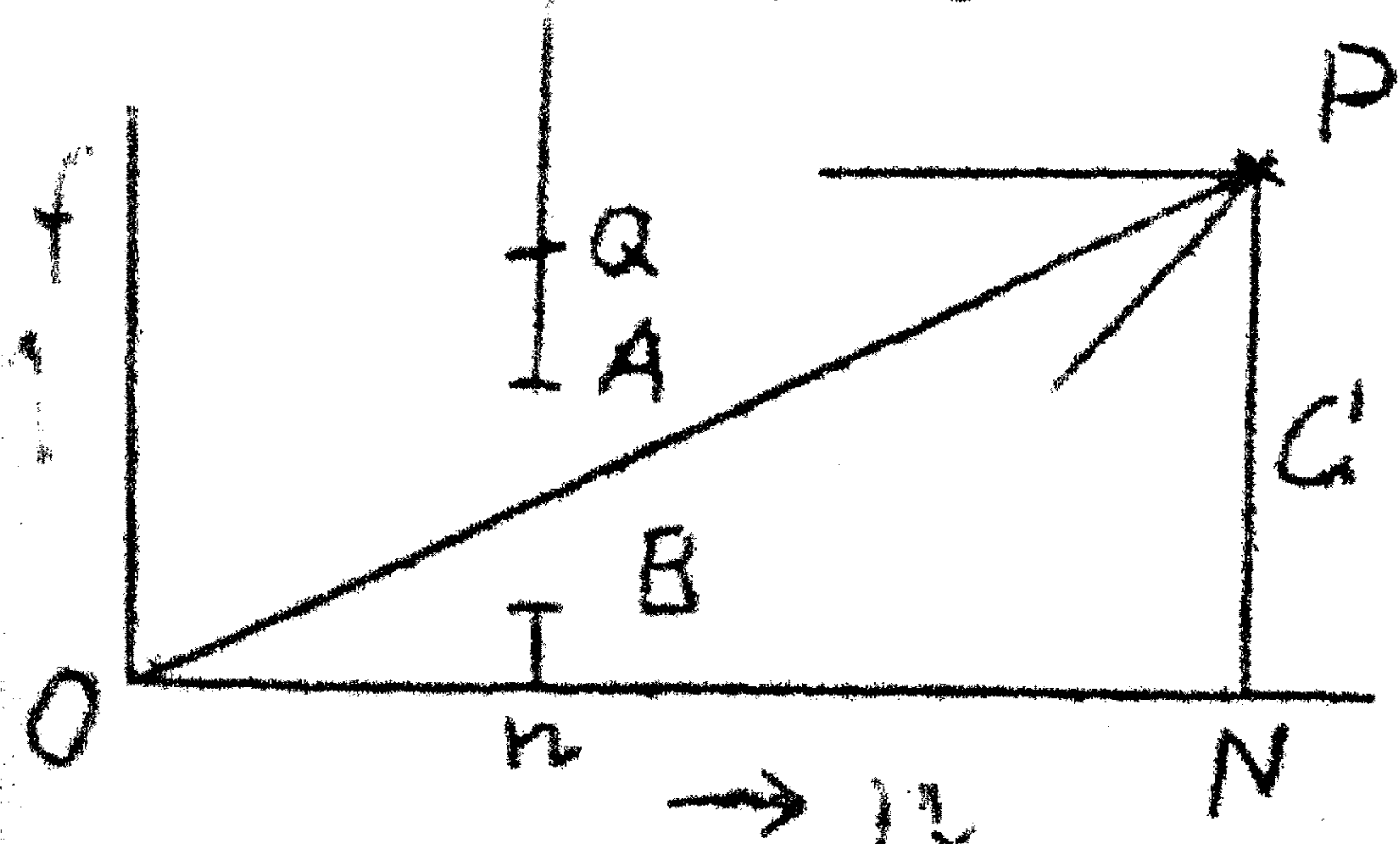
Voor de bepaling van $P(q)$ en $\hat{1}F$ bij gegeven q_0, H en N zijn er 4 methoden, waar 4 leden van ons team mee bezig zijn:

1. van Dantzig: Exakte oplossing voor eenvoudige waarden van de helling $q_0 = 1/(b+1)$.
2. Kemperman: Exakte oplossing voor kleine N en niet te kleine q_0 . Het eenvoudigste geval is $q_0 = 1/2$.
3. Van der Waerden: Benaderde oplossing met Gauszkromme en warmtegeleidingsvergelijking.
4. Rekenafdeling: Exakte oplossing voor kleine q_0 en niet te kleine N .

De vier methoden vullen elkaar aan. De resultaten van het onderzoek volgens 1, 2 en 3 zullen zonder twijfel belangrijke aanwijzingen voor de doelmatige inrichting van de berekening 4 geven. We zullen dus volgens 1, 2 en 3 dienen te beginnen.

B. Andere vormen van grenslijnen.

Wanneer de gunstigste vorm van potloden en parallellogrammen gevonden is, dan kunnen we eens nagaan of de $\hat{1}F$ door keuze van andere grenslijnen nog belangrijk kleiner gemaakt kan worden.



Op grond van heuristische beschouwingen meen ik, dat het gunstig zal zijn, de grenslijnen als volgt aan te nemen:

Laat P het punt met coördinaten N, C zijn, waar de goedkeur- en de afkeurlijn bij elkaar komen. Denk nu op een horizontale afstand n van O een diafragma AB geplaatst. Wanneer nu het grootste deel van de toevalspaden die tenslotte onder het punt P terechtkomen, ook onder langs het punt A gaan, en

wanneer eveneens het grootste deel van de toevalspaden die tenslotte boven P terecht komen ook boven langs B gaan, dan zal het diafragma AB de werkkromme van het enkelvoudige steekproefschema (N, C) niet roemenswaard beïnvloeden.

Kies dus een getal ϵ , b.v. $0,01$. Kies A zo dat van alle toevalspaden die van O naar P gaan slechts een gedeelte ϵ boven langs A gaan. Dan zal

$$n + 2$$

hetzelfde nog sterker gelden voor de toevalspaden die tenslotte onder P terecht komen. Analooog wordt B gekozen.

Op deze wijze krijgen we voor elke waarde van n een dubbel steekproefschema, dat in steilheid van de werkkromme niet veel scheelt met het enkelvoudige, maar dat zeker een kleinere $\hat{n}$ heeft.

Wanneer nu alle zo gevonden diafragma's (voor alle tussen 0 en N gelegen waarden van n) achter elkaar worden geschakeld, dan kan men hopen een gunstig steekproefschema te krijgen.

Laat P gehele coördinaten hebben. Het aantal toevalspaden van 0 naar P is $\binom{N}{c}$. Het aantal dat van 0 via een punt Q met coördinaten (n, f) naar P loopt, gedeeld door het totale aantal, is

$$(1) \quad \frac{\binom{n}{f} \binom{N-n}{c-f}}{\binom{N}{c}}$$

Dit sommeren we over alle $f > F$ en stellen de som op ε . Zo vinden we de hoogte F van het punt A, en analooog voor B.

Volgens Czuber Nr. 86 (Beispiel XLVIII) kunnen de uitdrukkingen (1) bij benadering door een Gausskromme worden voorgesteld, met middelwaarde nq_0 en spreiding $s^2 = n \frac{N-n}{N} q_0(1-q_0)$

Wordt dus γ zo bepaald dat de staart integraal van de Gausskromme juist ε is:

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}t^2} dt = \varepsilon$$

dan is de hoogte van A

$$F = nq_0 + \gamma s$$

en die van B

$$f = nq_0 - \gamma s$$

Beide punten A en B liggen dus op de kromme

$$(f - nq_0)^2 = \gamma^2 s^2 = \gamma^2 n \frac{N-n}{N} q_0(1-q_0)$$

Stel nu

$$\gamma^2 \frac{q_0(1-q_0)}{N} = \beta$$

dan vindt men als grenslijn de ellips

$$(f - nq_0)^2 = \beta n(N-n)$$

met OP als middellijn en met verticale raaklijnen in O en P.

Het "potlood" is geen slechte benadering van deze ellips (zie figuur) !

