

Over bacteriologisch wateronderzoek volgens de verdunningsmethode.

Inhoud

	Pag.
<u>1 Inleiding</u>	1
<u>2 Aannemelijkste schatting van de coli-dichtheid</u>	3
<u>2.1</u> Verdeling van de aantallen coli-bacteriën per monster . . .	3
<u>2.2</u> Aannemelijkste schatting van de coli-dichtheid bij proe- ven met één verdunning	4
<u>2.3</u> Aannemelijkste schatting van de coli-dichtheid bij proe- ven met verschillende verdunningen	4
<u>2.4</u> Enige opmerkingen betreffende de afleiding van vergelij- king (9)	6
<u>2.5</u> Berekening van de aannemelijkste schatting	7
<u>2.6</u> Successieve approximatie der aannemelijkste schatting volgens Finney	9
<u>3 Diverse andere schattingen van de coli-dichtheid</u>	9
<u>3.1</u> Schatting van Mac Crady	9
<u>3.2</u> Tweede schattingsmethode van Fisher	10
<u>3.3</u> Grafische methode van Leeftlang	11
<u>3.4</u> Grafische methode van Haspers	13
<u>4 Nauwkeurigheid van de schattingen</u>	14
<u>4.1</u> Inverse methode	14
<u>4.2</u> Asymptotische formule van Fisher	14
<u>4.3</u> Variatiecoëfficiënten van Halvorson en Ziegler voor één verdunning	16
<u>4.4</u> Onderzoekingen van Halvorson en Ziegler omtrent de nauw- keurigheid van de aannemelijkste schatting bij gebruik van verschillende verdunningen	19
<u>4.5</u> Bepaling van betrouwbaarheidsintervallen met behulp van de tweede schatting van Fisher	21
<u>4.6</u> Bepaling van betrouwbaarheidsintervallen met behulp van de aannemelijkste schatting	24
<u>4.7</u> Enige opmerkingen over de gunstigste opzet van de experi- menten	28
<u>5 Conclusies</u>	30
<u>5.1</u> Proeven met verschillende verdunningen	30
<u>5.2</u> Proeven met één verdunning	30
<u>Litteratuur en Bijlagen.</u>	

Over bacteriologisch wateronderzoek
volgens de verdunningsmethode.

1. Inleiding.

1.1 De verdunningsmethode (Eng.: dilution method) is een vorm van bacteriologisch onderzoek, die vaak gebruikt wordt om het gehalte aan coli-bacteriën van water te bepalen. Men neemt daartoe een aantal monsters van het te onderzoeken water en ent deze op voedingsbodems, die voor de ontwikkeling van coli-bacteriën gunstig zijn. Bij de ontwikkeling van coli-bacteriën in een dergelijke voedingsbodem treedt een gisting op, welke gemakkelijk waar te nemen is. Men veronderstelt, dat er dan en slechts dan gisting optreedt in de voedingsbodem als het geënte monster één of meer coli-bacteriën bevatte. Zulk een monster noemen wij positief en een monster dat geen gisting veroorzaakt (derhalve geen coli-bacteriën bevat), negatief. Het schijnt mogelijk te zijn, dat de coli-bacteriën koloniën vormen in het water. Deze koloniën zijn bij deze methode niet te onderscheiden van een enkel individu. Er bestaat bovendien de mogelijkheid dat de koloniën uiteenvallen. Wij zullen ons met de moeilijkheden, die hieruit voortvloeien, niet bezig houden, omdat zij meer op bacteriologisch terrein liggen.

1.2 De experimentele gegevens, voortvloeiend uit een onderzoek volgens de verdunningsmethode zijn dus:

1e. De volumina van de gebruikte monsters.

Wij verstaan onder het volume van een monster het volume van het onverdunde te onderzoeken water dat in het monster aanwezig is en dat dus, eventueel na verdunning, op een voedingsbodem geënt wordt.

Al naar gelang het doel van het onderzoek en de voorkennis die men omtrent het resultaat heeft werkt men met monsters van één volume v of met monsters van meerdere volumina $v_1, v_2, \dots, v_m$, waarbij m het aantal verdunningen genoemd wordt. In het laatste geval gebruikt men gewoonlijk gelijke aantallen monsters van ieder volume, terwijl de

volumina een meetkundige reeks vormen. De reden a van deze reeks heet de verdunningsfactor.

2e. Voor ieder der m gebruikte verdunningen met monster-volume v_i ($i = 1, 2, \dots, m$).

a) Het totale aantal n_i der gebruikte monsters van deze verdunning.

b) Het aantal daaronder positief bevonden monsters p_i .

Men tracht aan deze experimentele gegevens een interpretatie te geven in de vorm van een schatting van ρ , het aantal coli-bacteriën per cm^3 in het onderzochte water in onverdunde vorm. (Wij zullen dit aantal coli-bacteriën kortweg de coli-dichtheid van het onderzochte water noemen).

1.3 Dit overzichtsrapport is een samenvatting van de literatuur betreffende de verdunningsmethode, voorzien van een kritiek welke onmiddellijk op beschouwingen en resultaten uit de literatuur betrekking heeft.

In paragraaf 2 zullen wij de vergelijking afleiden, waaraan de zogenaamde aannemelijkste schatting van de coli-dichtheid voldoet. De afleiding komt in grote lijnen overeen met de afleiding van H.O. HALVORSON en N.R. ZIEGLER (1933, zie [10]¹⁾). Zij berust op de maximum-likelihood-theorie van R.A. FISHER en is door Fisher in principe reeds gegeven in zijn eerste publicatie over genoemde theorie (1921, zie [5]). De vergelijking zelf komt reeds voor bij M. GREENWOOD en G.U. YULE (1917, zie [3]). Wij zullen aangeven hoe deze vergelijking wordt opgelost en welke tafels er voor de oplossing bestaan.

In paragraaf 3 bespreken wij een aantal andere in de literatuur voorkomende schattingmethoden, waaronder de historisch interessante methode van M.H. Mac CRADY (1915, zie [1]), de veel toegepaste tweede methode van FISHER [5] en de grafische methoden van Ir K.W.H. LEEFLANG [25] en van Dr Ir J.H. HASPERS [28].

In paragraaf 4 behandelen wij tenslotte een aantal onderwerpen, betrekking hebbende op de nauwkeurigheid der schattingen, zoals de spreiding van de aannemelijkste schatting, ([5], [9], [18]), variatie-coëfficiënten ([11], [12]) en de constructie van betrouwbaarheidsintervallen ([16], [20], [30]).

1) Nummers tussen vierkante haken verwijzen naar de literatuurlijst (Bijlage I), nummers tussen ronde haken naar formules in de tekst.

Enige bijzondere kwesties, die naar aanleiding van dit literatuuronderzoek door het Mathematisch Centrum nader onderzocht zijn, zullen in een afzonderlijk rapport worden behandeld. Hieronder valt een onderzoek van de "eenvoudige rekenwijze" van Dr Ir J.H.HASPERS, de nauwkeurigheid van de aannemelijkste schatting bij experimenten met monsters van gelijk volume en een verhandeling over normvoorschriften.

2. Aannemelijkste schatting van de coli-dichtheid.

2.1 Verdeling van de aantallen coli-bacteriën per monster.

Stel wij onderzoeken $V \text{ cm}^3$ water, dat $\rho \cdot V$ coli-bacteriën bevat (ρ coli-bacteriën per cm^3), die ieder een volume hebben van $\varepsilon \text{ cm}^3$, kortweg ε -deeltjes, die hetzij coli-bacteriën zijn, hetzij bestaan uit water.

Trekt men nu een ε -deeltje uit $V \text{ cm}^3$, dan is er een bepaalde kans dat dit een coli-bacterie is. Wij veronderstellen nu dat deze kans overal in de onderzochte vloeistof even groot is, dus gelijk is aan $\frac{eV}{\varepsilon} = \rho\varepsilon$ omdat er ρV coli-bacteriën in het water zijn en in totaal $\frac{V}{\varepsilon}$ ε -deeltjes.

Neemt men nu een monster van $v \text{ cm}^3$, dus $\frac{v}{\varepsilon}$ ε -deeltjes, dan is de kans dat hierbij N coli-bacteriën zullen zijn volgens Bernoulli gelijk aan:

$$(1) \quad P[N = N] = \left(\frac{\frac{v}{\varepsilon}}{N} \right) \rho\varepsilon^N (1-\rho\varepsilon)^{\frac{v}{\varepsilon} - N}.$$

Dit geldt strikt genomen alleen, indien men $\frac{v}{\varepsilon}$ ε -deeltjes afzonderlijk trekt met teruglegging. Als v zeer klein is in verhouding tot V , hetgeen in de literatuur (al dan niet stilzwijgend) steeds ondersteld wordt, kan men formule (1) ook gebruiken voor monsters van $v \text{ cm}^3$, die in hun geheel uit het water genomen zijn.

Men heeft hier te doen met een Bernoulli-verdeling, waarvan het aantal termen $\frac{v}{\varepsilon}$ zéér groot is, terwijl het gemiddelde $\frac{v}{\varepsilon} \cdot \rho\varepsilon = \rho v$ klein is. Wij kunnen deze Bernoulli-verdeling dus vervangen door een Poisson-verdeling van de vorm:

$$(2) \quad P[N = N] = \frac{(\rho v)^N}{N!} e^{-\rho v}.$$

Derhalve geldt:

$$(3) \quad P[N=0] = e^{-v\varrho}.$$

Derhalve is de kans, dat een monster positief is, dus één of meer coli-bacteriën bevat:

$$(4) \quad P[N > 0] = 1 - e^{-v\varrho}.$$

2.2 Aannemelijkste schatting van de coli-dichtheid bij proeven met één verdunning.

Stel wij nemen n monsters van $v \text{ cm}^3$ uit water dat ϱ coli-bacteriën per cm^3 bevat, dan is de kans, dat hiervan p monsters positief en $q=n-p$ negatief zullen zijn, volgens de verdeling van Bernoulli:

$$(5) \quad P[p=p] = \binom{n}{p} (1 - e^{-v\varrho})^p (e^{-v\varrho})^q.$$

Stel nu, dat wij een experiment met n monsters van $v \text{ cm}^3$ hebben gedaan en dat p monsters positief en q negatief bleken te zijn. De aannemelijkste schatting (maximum likelihood-estimate) volgens Fisher van ϱ , is nu het getal $\hat{r}$, dat gesubstitueerd voor ϱ in (5), $P[p=p]$ maximaal maakt. We kunnen $\hat{r}$ het gemakkelijkst vinden door de logarithmische afgeleide van $P[p=p]$ naar ϱ gelijk te stellen aan 0, en uit de verkregen vergelijking ϱ op te lossen. Men vindt zodoende

$$(6) \quad \hat{r} = -\frac{1}{v} \ln \frac{q}{n}.$$

N.B.: Als $q=0$ is, levert deze formule $\hat{r}=\infty$, als $q=n$ is, $\hat{r}=0$. Deze resultaten zijn zeer onbevredigend, in feite kan men in de gevallen $q=0$ (alle monsters positief) en $q=n$ (geen enkel monster positief) geen bevredigende schatting van ϱ geven, omdat men dan kennelijk een te groot resp. te klein monstervolume gekozen heeft. Wij spreken in deze gevallen van een ongedetermineerd resultaat.

2.3 Aannemelijkste schatting van de coli-dichtheid bij proeven met verschillende verdunningen.

Stel wij nemen nu monsters van m verschillende volumina en wel n_1 monsters van volume v_1 , n_2 monsters van volume v_2 etc. t/m n_m monsters van volume v_m , alle uit water met coli-dichtheid ϱ . De kans, dat het volgende resultaat optreedt:

$$(7) \left\{ \begin{array}{llll} \text{volume } v_1 & p_1 & \text{monsters positief, } q_1 = n_1 - p_1 & \text{negatief} \\ " & v_2 & p_2 & " \\ " & \vdots & \vdots & \\ " & v_i & p_i & " \\ " & \vdots & \vdots & \\ " & v_m & p_m & " \end{array} \right. , \begin{array}{l} q_2 = n_2 - p_2 \\ q_i = n_i - p_i \\ q_m = n_m - p_m \end{array}$$

is dan:

$$(8) \quad P[p_1 = p_1, \dots, p_m = p_m] = \prod_{i=1}^m P[p_i = p_i] = \\ = \prod_{i=1}^m \binom{n_i}{p_i} (1 - e^{-v_i \rho})^{p_i} (e^{-v_i \rho})^{q_i}$$

We kunnen nu, gegeven het proefresultaat (7), wederom de aannemelijkste schatting van ρ bepalen. Daartoe zal men de waarde van ρ moeten zoeken waarvoor de waarschijnlijkheid van het proefresultaat (7), derhalve de uitdrukking (8) maximaal wordt.

Door de logarithmische afgeleide van deze uitdrukking gelijk te stellen aan nul, vinden wij, dat de aannemelijkste schatting $\hat{r}$ van ρ moet voldoen aan de vergelijking:

$$(9) \quad \sum_{i=1}^m p_i v_i \frac{e^{-v_i \hat{r}}}{1 - e^{-v_i \hat{r}}} = \sum_{i=1}^m q_i v_i$$

Hiermee is equivalent de vergelijking

$$(9a) \quad \sum_{i=1}^m p_i v_i \frac{1}{1 - e^{-v_i \hat{r}}} = \sum_{i=1}^m n_i v_i$$

Ook hier treedt de mogelijkheid van een ongedetermineerd resultaat op. Als alle monsters positief zijn ($p_i = n_i$ voor $i=1, 2, \dots, m$) blijkt uit (9a) dat slechts $\hat{r} = \infty$ voldoet, als alle monsters negatief zijn ($p_i = 0$ voor $i=1, 2, \dots, m$) voldoet slechts de waarde $\hat{r} = 0$.

Als slechts bij één verdunning positieve monsters optreden (b.v. $p_i = 0$ voor $i > 1$, $p_1 \neq 0$), kan men $\hat{r}$ expliciet uit vergelijking (9) oplossen:

$$(10) \quad \hat{r} = \frac{1}{v_1} \ln \frac{\sum_{i=1}^m n_i v_i}{\sum_{i=1}^m q_i v_i} \quad (\text{Zie [23]})$$

Het is duidelijk dat formule (6) een bijzonder geval is van formule (10).

Indien bij meerdere verdunningen positieve monsters optreden, zal men voor de berekening van $\hat{r}$ uit vergelijking (9) zijn toevlucht moeten nemen tot een iteratieproces. Hierop komen wij terug in 2.5.

2.4 Enige opmerkingen betreffende de afleiding van vergelijking (9).

De in 2.3 weergegeven afleiding van vergelijking (9) komt, behoudens de notatie, in grote lijnen overeen met de afleiding van H.O. HALVORSON en N.R. ZIEGLER (1933, zie [10]) R.A. FISHER leidt in zijn eerste publicatie over de theorie van "maximum likelihood" (1921, zie [5]) met behulp van deze theorie reeds een vergelijking af, die, zoals wij zullen aantonen, overeenstemt met vergelijking (9).

In de notatie van FISHER wordt de coli-dichtheid voorgesteld door de letter n , terwijl de monster-volumina een meetkundige reeks $\frac{1}{a^x}$ vormen ($x=1,2,\dots,N$); $p=e^{-\frac{n}{a^x}}$ is de kans dat een monster van volume $\frac{1}{a^x}$ negatief is (vergelijk (3)) en $q=1-p$.

Fisher komt tot de conclusie dat de aannemelijkste schatting $\hat{n}$ van n , de oplossing naar n is van de vergelijking:

$$(11) \quad S_1(\ln p) - S_2\left(\frac{p}{q} \ln p\right) = 0,$$

waarin S_1 sommatie over de negatieve, S_2 sommatie over de positieve monsters voorstelt.

Indien wij overgaan op onze notatie, door te substitueren: $\frac{1}{a^x} = v_x$, $\hat{n} = \hat{r}$ en aannemen, dat er p_x positieve en q_x negatieve monsters zijn met volume v_x , dan gaat de vergelijking (11) over in:

$$-\hat{r} \sum_{x=1}^m q_x v_x + \hat{r} \sum_{x=1}^m p_x v_x \frac{e^{-\hat{r} v_x}}{1 - e^{-\hat{r} v_x}} = 0,$$

welke vergelijking equivalent is met vergelijking (9).

De schatting, die uit formule (9) volgt wordt in de litteratuur vaak het "waarschijnlijkste aantal coli-bacteriën" (Most probable number of B.coli) genoemd. Deze benaming is een overblijfsel uit de tijd, dat men dit probleem behandelde met de methode der inverse waarschijnlijkheden (gebaseerd op het theorema van BAYES). Men gaat daarbij uit van de veronderstelling, dat, a priori, alle waarden van ρ tussen 0 en een zeker maximum even waarschijnlijk zijn.

Daarmee bepaalt men de voorwaardelijke waarschijnlijkheidsverdeling van ρ bij gegeven resultaat van een proef. De waarschijnlijkste dichtheid is dan de modus van deze verdeling. Tegen deze methode bestaat het bekende bezwaar, dat iedere veronderstelling over een a priori-verdeling van ρ zinloos is, omdat ρ geen stochastische grootte, doch een onbekende parameter is.

M.GREENWOOD en G.U.YULE hebben reeds in 1917 volgens de inverse methode een vergelijking voor de "waarschijnlijkste dichtheid" van de coli-bacteriën afgeleid (zie [3]). Deze vergelijking is dezelfde als vergelijking (9). Mede door deze toevallige overeenstemming is de verouderde benaming waarschijnlijkste dichtheid i.p.v. aannemelijkste schatting van de dichtheid, in zwang gebleven. Ook na het verschijnen van de publicatie [5] van FISHER blijven vele schrijvers voor de afleiding van vergelijking (9) naar het artikel [3] van M.GREENWOOD en G.U.YULE verwijzen (o.a. J.K.HOSKINS in [13]).

2.5 Berekening van de aannemelijkste schatting.

In de litteratuur vindt men bij H.O.HALVORSON en N.R.ZIEGLER [10], J.K.HOSKINS [14], S.SWAROOP [18] en R.POMEROY [23] berekeningsmethoden, die in principe alle neerkomen op het oplossen van de vergelijking (9) door middel van "trial and error". Dit wil zeggen: men substitueert voor de onbekende een aantal in aanmerking komende waarden en kiest hieruit als benaderende oplossing die waarde, waarvoor het verschil tussen linker- en rechterlid in absolute zin het kleinst is. Indien men bijvoorbeeld $\hat{r}$ in 3 decimalen nauwkeurig uit (9) wil bepalen, zal men een reeks waarden substitueren, die 10^{-3} verschillen. Daarbij gaat men van tevoren na in welke buurt $\hat{r}$ moet liggen, bijvoorbeeld door ρ eerst met behulp van formule (6) te schatten uit het resultaat van een verdunning, waarbij zowel positieve als negatieve monsters zijn opgetreden (zoals in § 4 zal blijken, kunnen wij hiervoor het beste een verdunning kiezen waarbij 60 à 80 procent van monsters positief is).

Men treft in de litteratuur uitvoerige tabellen aan, om het aan deze methode verbonden rekenwerk te verlichten. HALVORSON en ZIEGLER geven in [10] tabellen van de functie e^{-x} , te gebruiken voor berekeningen met formule (6), en tabellen van de functie $\frac{1}{1-e^{-x}}$ voor de oplossing van vergelijking (9a). HOSKINS geeft in [14] tabellen van $k \frac{1}{e^{\lambda}-1}$

als functie van k en λ en SWAROOP in [18] tabellen van

$$E_x = \frac{1}{e^{n/10^x} - 1} \text{ voor } x=1,2,3; n=1,2,\dots,1000.$$

Men heeft ook getracht het oplossen van vergelijking (9) te vereenvoudigen, door deze vergelijking in een andere vorm te brengen. POMEROY (zie [23]), voert in de formule het totale volume van de positieve monsters $A = \sum_{i=1}^m p_i v_i$ en het totale volume van de negatieve monsters $B = \sum_{i=1}^m q_i v_i$ in. Formule (10) gaat dan over in $\hat{r} = \ln \frac{A}{B}$ en indien men slechts bij twee verdunningen positieve monsters aantreft is de vergelijking van POMEROY wellicht iets eenvoudiger dan de gebruikelijke. Voor de oplossing blijft hij echter aangewezen op de methode van "trial and error".

Men kan de aannemelijkste schatting $\hat{r}$ van ϱ ook bepalen door in de zogenaamde "likelihood-function":

$$\begin{aligned} L &= \ln P[p_1 = p_1, \dots, p_m = p_m] = \\ &= \text{const.} \times \left\{ \sum_{i=1}^m p_i \log(1 - e^{-v_i \varrho}) + \sum_{i=1}^m q_i \log e^{-v_i \varrho} \right\} \end{aligned}$$

een aantal in aanmerking komende waarden voor ϱ te substitueren en als benadering voor $\hat{r}$ die waarde te nemen waarvoor L het grootst is. Dit is dus ook een methode van "trial and error", doch nu toegepast op de "likelihood-function" in plaats van op vergelijking (9). HOSKINS heeft hiertoe tabellen vervaardigd van $\log(1 - e^{-10^{-x} \varrho})$ en $\log e^{-10^{-x} \varrho}$ als functies van x en ϱ (zie [13]).

Pasklare tabellen van $\hat{r}$ vindt men in de artikelen [10] en [14]. De tabellen van HALVORSON en ZIEGLER [10] hebben betrekking op experimenten, waarbij $v_1=10$, $v_2=1$, $v_3=0,1$ en $n_1=n_2=n_3=10$. De tabellen van HOSKINS [14], die veel uitvoeriger zijn, zijn bruikbaar bij dezelfde reeks van monster volumina (10; 1; 0,1) doch nu met $n_1 \leq 5$, $n_2 \leq 5$, $n_3 \leq 5$, terwijl tevens nog minder uitgebreide tabellen voor de reeksen (50, 10, 1) en (100, 50, 10) zijn opgenomen.

Wij wijzen er op, dat een tabel voor $\hat{r}$ bij gegeven v_1 , v_2 en v_3 even goed gebruikt kan worden voor de bepaling van $k\hat{r}$ bij gegeven kv_1 , kv_2 en kv_3 , omdat $\hat{r}$ in vergelijking (9) uitsluitend in de combinaties $\hat{r}v_i$ ($i=1,2,\dots,m$) voorkomt. Men kan dus met behulp van de tabellen uit [10] $\hat{r}$ eveneens bepalen indien b.v. $v_1=100$, $v_2=10$ en $v_3=1$ is. Hiertoe heeft men de in de tabel gevonden waarde slechts door 10 te delen.

2.6 Successieve approximatie der aannemelijkste schatting volgens Finney.

D.J.FINNEY heeft getracht een methode aan te geven, waarmee men de aannemelijkste schatting van de coli-dichtheid, op principieel andere wijze dan hierboven, door successieve approximatie zou kunnen bepalen (zie [27]). Zijn methode berust op de overeenstemming die er zou bestaan tussen de vergelijking voor de aannemelijkste schatting en de minimum voorwaarde van de methode der kleinste quadraten ter bepaling van een rechte regressielijn. Voor deze laatste bepaling is door FINNEY in [27] een successieve approximatie-methode ontwikkeld, waarbij echter verondersteld wordt, dat de waargenomen punten een constante spreiding ten opzichte van de regressielijn hebben. De overeenkomstige grootte uit de vergelijking voor de aannemelijkste schatting is niet constant. Hierdoor vervalt de theoretische fundering van Finney's methode. Wij zullen deze methode, die bovendien weinig minder omslachtig schijnt te zijn, dan de berekening van $\hat{r}$ uit vergelijking (9) met "trial and error", verder buiten beschouwing laten.

3. Diverse andere schattingen van de coli-dichtheid.

3.1 Schatting van Mac Crady.

M.H.Mac Crady gaat voor de afleiding van zijn schatting uit van de volgende redenering (zie [1]). Stel men neemt een steekproef, groot $V \text{ cm}^3$ uit het te onderzoeken water en kent een monster van $v \text{ cm}^3$ uit deze steekproef op een voedingsbodem. Bij het nemen van het monster bestaat een kans $1 - \frac{v}{V}$, dat een in de steekproef aanwezige coli-bacterie, niet in het monster terechtkomt. Indien er x coli-bacteriën in het monster zijn, is de kans dus $(1 - \frac{v}{V})^x$, dat het monster negatief is. Deze uitdrukking gaat voor $V \rightarrow \infty$ over in $e^{-\rho V}$ (aangezien $x = \rho V$), hetgeen overeenkomt met vergelijking (3). Mac Crady maakt deze overgang echter niet en werkt verder met de kans $(1 - \frac{v}{V})^x$, waarbij hij een gedragslijn volgt, die grote overeenstemming vertoont met de methode van de aannemelijkste schattingen. Dit is merkwaardig in zoverre de publicatie [1] stamt uit 1915, derhalve 6 jaren voor het verschijnen van Fisher's eerste artikel over de theorie van "maximum likelihood".

Mac Crady geeft alleen getallen voorbeelden van zijn schattingsmethode. A.WOLMAN en W.WEAVER hebben aan de hand van zijn theorie formules opgesteld voor het geval van 3

verdunningen met verdunningsfactor 10 (zie [2]). De verkregen schattingsvergelijking wordt zodanig vereenvoudigd, dat zij de overeenkomstige vergelijking, vervat in (9), dichter benadert. J.H.LINSCHOTEN heeft een nomogram vervaardigd, waarmee men de vergelijking van WOLMAN en WEAVER in bepaalde gevallen gemakkelijk kan oplossen (zie [8]). Verder bestaan er tabellen van de schatting van Mac Crady¹⁾.

Bij de methode van Mac Crady wordt ondersteld, dat de steekproef van volume V representatief is voor het reservoir, waaruit deze steekproef genomen is. Eigenlijk wordt alleen het aantal coli-bacteriën in V geschat. Men zal dus V groot moeten nemen, maar er is geen enkel nader criterium voor deze grootte aan te geven, terwijl de uitkomsten nog van V afhangen en wel meer naar mate V kleiner is. De limietovergang voor $V \rightarrow \infty$, die door alle latere schrijvers is toegepast kan dus beschouwd worden als een perfectionering van de methode van Mac Crady en er is geen reden deze limietovergang niet uit te voeren.

3.2 Tweede schattingsmethode van Fisher.

In artikel [5] bespreekt R.A.FISHER nog een tweede schattingsmethode. Hierbij wordt de verwachting van het totaal aantal negatieve monsters $\sum_{i=1}^m q_i = \sum_{i=1}^m n_i e^{-e v_i}$ gelijk gesteld aan het werkelijk gevonden aantal $\sum_{i=1}^m q_i$. Als schatting r van ρ neemt men nu de oplossing van deze vergelijking naar ρ .

Deze methode bezit het voordeel, dat een zeer beknopte tabel voor de bepaling van r voldoende is, indien de monstervolumina een meetkundige reeks met verdunningsfactor a vormen en van iedere verdunning evenveel (n) monsters genomen worden.

In dat geval blijkt namelijk te gelden

$$\ln r = x \ln a - K$$

waarin $x = \frac{mn - \sum_{i=1}^m q_i}{n}$ is en K behalve voor de grootste en kleinste monstervolumina bij de experimenten gebruikt, nageenoe constant is. Een tabel van K , met behulp waarvan men r kan bepalen behoeft dus niet uitgebreid te zijn; men vindt een dergelijke tabel in de "Statistical tables" van

1) Zie Bijlage III No. 5

FISHER en YATES (zie [26]).

Men gebruikt bij deze methode $\frac{1}{n} \sum_{i=1}^m q_i$ als schatting van $\frac{1}{n} \sum_{i=1}^m q_i$; dit is een zuivere ¹⁾ en, omdat de spreiding van $\frac{1}{n} \sum_{i=1}^m q_i$ voor $n \rightarrow \infty$ tot nul nadert, tevens een bruikbare ¹⁾ schatting; r is een functie van deze schatting, ρ de overeenkomstige functie van $\frac{1}{n} \sum_{i=1}^m q_i$; hieruit volgt dat r een bruikbare schatting is van ρ .

Op analoge wijze kan men aantonen, dat $\hat{r}$ eveneens een bruikbare schatting van ρ is.

Het zal in paragraaf 4 echter blijken, dat de aannemelijkste schatting veel nauwkeuriger is dan de tweede schatting van Fisher. Indien de nauwkeurigheid van de schatting mede in acht genomen wordt, zal het gebruik van de aannemelijkste schatting dus de voorkeur verdienen; de tweede schatting is echter zeer geschikt indien men snel een antwoord geeft, zij het iets minder nauwkeurige, interpretatie aan proefresultaten wil geven.

3.3 Grafische methode van Leeftlang.

Indien men n monsters trekt met een volume v , is de verwachting van de frequentie der positieve monsters (de verwachting van $\frac{p}{n}$) gelijk aan:

$$(12) \quad \sum \frac{p}{n} = 1 - e^{-v\rho}$$

hetgeen onmiddellijk volgt uit (5).

Uit (12) volgt:

$$(13) \quad \ln v = \ln \left\{ -\ln \left(1 - \sum \frac{p}{n} \right) \right\} - \ln \rho$$

Indien men nu een grafische voorstelling vervaardigt van $\ln v$ als functie van $E = \sum \frac{p}{n}$ bij gegeven ρ , dan ontstaat een kromme, die de gedaante heeft van k in figuur 1, welke wij de S-kromme van LEEFTLANG zullen noemen.

1) De definities van "zuivere" en "bruikbare" schattingen vindt men in memorandum S 53 (M 21):

"Enige definiteis uit de schattingstheorie" (Bijlage IV).

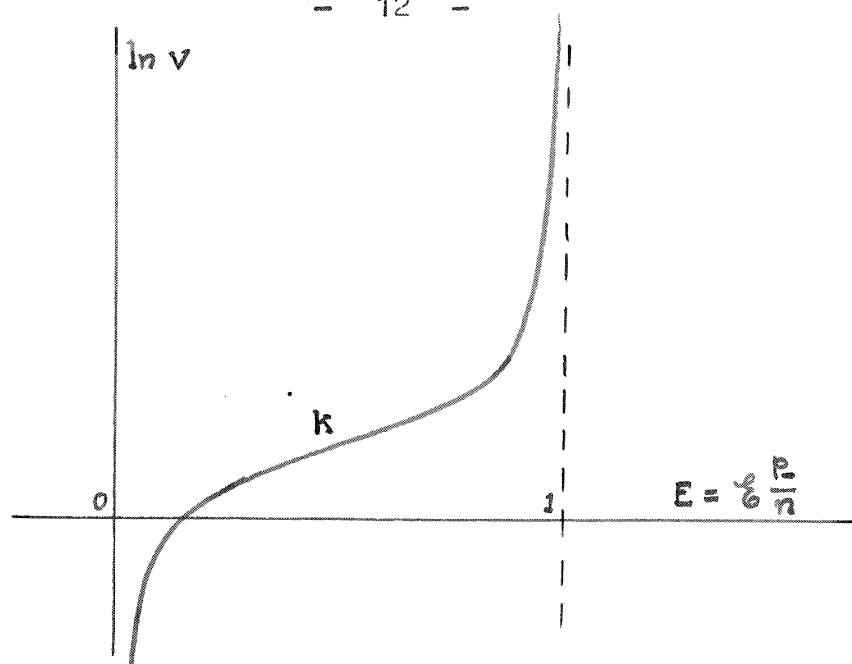


fig. 1 (S-kromme van LEEFLANG).

Een verandering van ρ betekent slechts een verschuiving van de S-kromme, zoals onmiddellijk volgt uit (13).

Een verdunningsexperiment levert ons bij iedere gebruikte monster-volume v_i de frequentie van de positieve monsters $\frac{p_i}{n_i}$. De punten $(\frac{p_i}{n_i}, \ln v_i)$ die hiermee in de grafiek corresponderen, kunnen beschouwd worden als waarnemingen van punten van een kromme, gegeven door vergelijking (13). Indien men bij deze punten op een verantwoorde wijze een S-kromme van Leeftlang k_0 aanpast, zal de met k_0 corresponderende ρ , die wij met r_0 zullen aanduiden een schatting zijn van de ware coli-dichtheid. Indien k_0 eenmaal aangepast is, kan men r_0 gemakkelijk bepalen uit het snijpunt $(1 - e^{-1}, \ln v_0)$ van k_0 met de rechte: $E = 1 - e^{-1} = 0,63...$; immers dan geldt:

$$\ln v_0 = \ln \{- \ln e^{-1}\} - \ln r_0 = - \ln r_0$$

$$\text{dus: } r_0 = \frac{1}{v_0}.$$

De methode van Ir K.W.H. LEEFLANG (zie [25]) berust op het boven beschreven principe. Ir Leeftlang past bij de waargenomen punten echter geen S-kromme aan, doch trekt door deze punten uit de hand een vloeiende kromme k' . Indien k' in gedaante sterk van een S-kromme afwijkt, concludeert Ir Leeftlang, dat het waarnemingsmateriaal niet aan de eisen voldoet op grond waarvan vergelijking (13) is opgesteld. Ir Leeftlang veronderstelt, dat de monsters dan niet gelijkwaardig zijn, hetgeen in de voorbeelden, die hij geeft zeker gerechtvaardigd is, omdat het daar monsters van rivierwater betreft, die op verschillende tijdstippen genomen zijn.

Wijkt echter de "waargenomen kromme" k' niet sterk van een S-kromme af, dan bepaalt Ir Leeftang zijn schatting r_2 uit het snijpunt van k' met de rechte $E = 1 - e^{-1}$, analoog aan de hierboven beschreven bepaling van r_0 .

Door het gebruik van een aangepaste S-kromme in plaats van de kromme k' , kan men de methode van Ir Leeftang uit theoretisch oogpunt meer aanvaardbaar maken. De moeilijkheid is hier het vinden van een "verantwoorde" aanpassing. Verder zal men moeten aangeven, wanneer geen verantwoorde aanpassing mogelijk is (wanneer de monsters als niet gelijkwaardig kunnen worden beschouwd) terwijl het ook gewenst is, de nauwkeurigheid van het resultaat te onderzoeken. Wij zullen hierop in het reeds genoemde aanvullende rapport terugkomen.

3.4 Grafische methode van Haspers.

Indien

$$(14) \quad z = \log \ln \frac{1}{1 - \frac{p_i}{n}}$$

is, volgt uit (13) onmiddellijk:

$$(15) \quad \log \rho = z - \log v$$

De grafische voorstellingen van $\log \rho$ als functie van z bij verschillende waarden van v zullen rechten worden, alle evenwijdig met de rechte $\log \rho = z$.

De lijnen, overeenkomend met de S-kromme van Leeftang, zijn in het $(z, \log \rho)$ -diagram dus recht en evenwijdig met de z -as. Indien men waarnemingsresultaten $(\frac{p_i}{n_i}, v_i)$ nu tot $(z_i = \log \ln \frac{1}{1 - \frac{p_i}{n_i}}, v_i)$ transformeert, dan kan men de punten

gekaracteriseerd door (z_i, v_i) met behulp van de rechten $\log \rho = z - \log v$ gemakkelijk in het $(z, \log \rho)$ -diagram overbrengen. Past men bij de verkregen punten een rechte l_0 aan, parallel met de z -as, dan geeft de hoogte van deze rechte boven de z -as een schatting van $\log \rho$.

Op dit principe berust de methode van Dr Ir J.H. HASPERS (zie [28]). Deze verbindt de waargenomen punten echter door een gebroken rechte l' en bepaalt zijn schatting r_h uit het snijpunt van l' met de rechte

$$z = \log \ln \frac{1}{e^{-1}} = 0$$

De schatting r_h is daarom zeer grof, want in feite wordt slechts gebruikt gemaakt van punten (z_i, v_i) , die ter weerszijden het dichtst bij de ρ -as ($z=0$) liggen.

De methode kan echter wellicht verbeterd worden door gebruik van een aangepaste rechte l_0 in plaats van l' . Men zal dan moeilijkheden van dezelfde aard moeten overwinnen als beschreven in de laatste alinea van 3.3

4. Nauwkeurigheid van de schattingen.

4.1 Inverse methode.

Een deel van de publicaties betreffende de nauwkeurigheid van de schatting der coli-dichtheid uit verdunnings-experimenten is gebaseerd op de veronderstelling dat ρ a priori een homogene verdeling heeft. Zoals reeds in 2.4 is opgemerkt, is deze veronderstelling ongemotiveerd. Wij zullen dan ook alle conclusies, die hierop berusten, verder buiten beschouwing laten.

4.2 Asymptotische formule van Fisher.

Vele schrijvers gebruiken voor hun beschouwingen betreffende de nauwkeurigheid der schatting $\hat{r}$ een stelling uit de theorie van de aannemelijkste schattingen van R.A.FISHER. Wij zullen deze stelling hieronder eerst in het kort uiteenzetten.

Men gaat bij FISHER's theorie uit van n waarnemingen $x_1, \dots, x_n$ van een stochastische grootheid x met verdelingsdichtheid $f(x|\theta)$, die bekend is behoudens de parameter θ . Fisher definieert nu als "likelihood function" ("aannemelijkheid")

$$(16) \quad L = \ln \prod_{i=1}^n f(x_i|\theta) = \sum_{i=1}^n \ln f(x_i|\theta)$$

De aannemelijkste schatting $\hat{t}_n$ van θ , behorende bij bovengenoemde n waarnemingen is die waarde van θ , waarvoor L maximaal is. Volgens de definitie van t_n geldt dus:

$$(17) \quad \left(\frac{\partial L}{\partial \theta} \right)_{\theta = \hat{t}_n} = 0$$

Men kan nu bewijzen (zie b.v. Prof. Dr D.van Dantzig, Kadercursus Mathematische Statistiek¹⁾, pg. 203-210; Whr. 292-299):

1) In gestencilde vorm uitgegeven door het "Mathematisch Centrum".

$$(18) \quad \lim_{n \rightarrow \infty} \sigma_{\hat{t}_n | \theta}^2 = - \frac{1}{\theta \frac{\partial^2 L}{\partial \theta^2}}$$

waarin $\sigma_{\hat{t}_n | \theta}^2$ de spreiding van $\hat{t}_n$ is, bij gegeven parameter θ .

Deze formule komt overeen met de asymptotische formule van FISHER. Het door Prof. VAN DANTZIG gegeven bewijs geldt echter slechts onder bepaalde voorwaarden. Deze voorwaarden zijn:

- 1) $\ln f(x|\theta)$ is voor iedere reële x en iedere toegelaten θ gedefinieerd en tweemaal partieel continu naar θ differentieerbaar.
- 2) Van de stochastische grootheden $\frac{\partial \ln f(x|\theta)}{\partial \theta}$ en $\frac{\partial^2 \ln f(x|\theta)}{\partial \theta^2}$ bestaan de eerste en tweede momenten; van $\frac{\partial \ln f(x|\theta)}{\partial \theta}$ bovendien het derde moment.
- 3) De aannemelijkste schatting $\hat{t}_n(x_1, \dots, x_n)$ bestaat voor iedere n .
- 4) Het eerste en tweede moment van $\hat{t}_n$ bestaan en blijven begrensd voor iedere toegelaten θ , als $n \rightarrow \infty$.

Bij de toepassing van deze stelling op de resultaten van een verdunningsexperiment laat men de p_i corresponderen met de x_i en de natuurlijke logaritmie van $\prod_{i=1}^m P[p_i = p_i]$ (zie (8)), met de likelihood-function L . De p_i bezitten echter een discrete verdeling, die behalve van θ ook nog afhangt van v_i . Daarbij komt nog dat de verdeling van $\hat{r}$ geen eerste en geen tweede moment bezit. Er bestaat namelijk steeds een positieve kans dat $\hat{r} = \infty$ (zie de opmerkingen betreffende een ongedetermineerd resultaat in 2.2 onder formule (6) en in 2.3 onder formule (9a)). Aan de voorwaarden 1) en 4) is dus in ieder geval niet voldaan. Hoewel genoemde stelling van Fisher wellicht te bewijzen is onder iets algemenere voorwaarden, dan de bovenvermelde is het zeker niet mogelijk haar af te leiden als niet voldaan is aan voorwaarde 4) (hetgeen hier het geval is), aangezien dan de spreiding van de aannemelijkste schatting niet bestaat. Wij mogen dus wel concluderen dat toepassing van de stelling in dit geval niet gerechtvaardigd is.

Onder degenen, die de stelling desondanks hebben toegepast moet allereerst FISHER zelf genoemd worden (zie [5]),

die de stelling onder andere gebruikt om de doeltreffendheid¹ van zijn tweede schatting te bepalen (zie 3.2).

Verder vinden wij de stelling toegepast voor het berekenen van de "spreiding van r " bij S.SWAROOP (zie [18]) en voor de berekening van de "spreiding van $\ln \hat{r}$ " bij E.S.ALLEN (zie [9]). Behalve de reeds genoemde bezwaren kan men tegen hun methode nog aanvoeren, dat zij de uiteraard onbekende ρ uit de asymptotische formule van Fisher, vervangen door de aannemelijkste schatting $\hat{r}$ van ρ .

W.G.COCHRAN (zie [30]) heeft een grafiek vervaardigd, waarin de verhouding van de "spreiding van de aannemelijkste schatting" tot deze schatting wordt uitgezet tegen de ware coli-dichtheid. De "spreiding" is hier eveneens berekend met behulp van de asymptotische formule van Fisher.

Genoemde grafische voorstelling blijkt voor een geval met verdunningsfactor 10 een aantal schommelingen te vertonen, dat gelijk is aan het aantal gebruikte verdunningen; bij een geval met verdunningsfactor 2 zijn geen schommelingen te onderkennen. Dit effect is in overeenstemming met hetgeen wij in 4.3 zullen zien: er is bij proeven met één verdunning één ρ waarbij de nauwkeurigheid maximaal is. Indien men een grote verdunningsfactor gebruikt, verschillen de gebruikte monster-volumina zo sterk, dat in een "nauwkeurighedsdiagram" de bij de verschillende verdunningen behorende nauwkeurighedsmaxima gescheiden naar voren komen; dit zal niet meer het geval zijn als de verdunningsfactor klein is, waardoor de maxima dichter bij elkaar liggen. Op grond van deze overweging verdient het aanbeveling met een kleine verdunningsfactor te werken.

De nauwkeurighedsmaat, die Cochran gebruikt heeft, is echter, zoals boven opgemerkt is, onbetrouwbaar, zodat men aan de hand van zijn grafische voorstelling geen conclusies kan trekken.

4.3 Variatiecoëfficiënt van Halvorson en Ziegler voor één verdunning.

Het aantal negatieve monsters bij gebruik van één verdunning is evenals het aantal positieve monsters binomiaal verdeeld volgens (5):

$$P[\underline{q} = q] = \binom{n}{q} (1 - e^{-ev})^{n-q} e^{-evq},$$

1) Zie het memorandum S 53 (M 21):

"Enige definities uit de schattingstheorie" (Bijlage IV).

waaruit volgt, dat geldt:

$$(19) \quad \bar{b}q = n e^{-\rho v}$$

en

$$(20) \quad \sigma_q = \sqrt{\bar{b}(q - \bar{b}q)^2} = \sqrt{n e^{-\rho v} (1 - e^{-\rho v})}.$$

De binomiale verdeling van q wordt geapproximeerd door een normale verdeling met gemiddelde $\mu = \bar{b}q$ en spreiding $\sigma = \sigma_q$. Indien deze normale verdeling exact zou gelden, kon men schrijven:

$$P[\mu - \sigma \leq q \leq \mu + \sigma] = 0,68.$$

Hier zal echter slechts bij benadering gelden:

$$(21) \quad P[\bar{b}q - \sigma_q \leq q \leq \bar{b}q + \sigma_q] \approx 0,68,$$

welke benadering beter wordt naarmate n groter is.

Men kan nu van de verdeling van q overgaan op de verdeling van $\hat{r}$ door de transformatie:

$$(6) \quad \hat{r} = -\frac{1}{v} \ln \frac{q}{n}.$$

Indien men nu definieert:

$$(22) \quad \begin{cases} r_1 = -\frac{1}{v} \ln \frac{1}{n} (\bar{b}q + \sigma_q) = \frac{1}{v} \ln \frac{1}{e^{-\rho v} + \frac{\sigma_q}{n}} \\ r_2 = -\frac{1}{v} \ln \frac{1}{n} (\bar{b}q - \sigma_q) = \frac{1}{v} \ln \frac{1}{e^{-\rho v} - \frac{\sigma_q}{n}} \end{cases}$$

zal gelden:

$$(23) \quad P[r_1 \leq \hat{r} \leq r_2] \approx 0,68$$

We zien hieruit, dat de kans, dat de aannemelijkste schatting binnen het interval (r_1, r_2) valt bij benadering constant is. Op grond hiervan kan men de wijidte $r_2 - r_1$ van het interval (r_1, r_2) beschouwen als een maat voor de nauwkeurigheid van de schatting $\hat{r}$. De grootheid

$$(24) \quad V = \frac{r_2 - r_1}{\rho} = \frac{1}{\rho v} \ln \frac{n e^{-\rho v} + \sigma_q}{n e^{-\rho v} - \sigma_q}$$

is dus een maat voor de relatieve nauwkeurigheid van de aannemelijkste schatting $\hat{r}$ in betrekking tot de ware coli-dichtheid ρ .

De variatiecoëfficiënt van HALVORSON en ZIEGLER (zie [1]) kan uit de boven gedefinieerde grootheid V verkregen worden door in het rechterlid van (24) ρ door $\hat{r}$ te vervangen.

Het is een algemene neiging van de toepassers van de theorie van Fisher om de aannemelijkste schatting te ge-

bruiken alsof zij gelijk ware aan de geschatte parameter. Dit wordt heel dikwijls in de hand gewerkt doordat men voor parameter en schatting hetzelfde symbool gebruikt. Ook Halvorson en Ziegler doen dit. Wil men formule (24) toepassen bij een concreet experiment door schatting van ρ dan zal men moeilijk anders te werk kunnen gaan dan door substitutie van de gevonden waarde van $\hat{r}$ voor ρ , daar deze laatste onbekend is. Men dient hierbij echter wel te bedenken, dat men zodoende slechts een zeer globale indruk verkrijgt van de nauwkeurigheid, omdat $\hat{r}$ zeer sterk van ρ kan afwijken, hetgeen uit de hieronder te bespreken tabellen van Halvorson en Ziegler duidelijk blijkt.

Men kan V echter ook beschouwen voor verschillende bekende waarden van ρ . Wij zien uit de vergelijkingen (20) en (24) dat V uitsluitend een functie is van n en ρv . Als functie van n is V monotoon dalend; V nadert tot 0 voor $n \rightarrow \infty$ en is dan $O(\sqrt{n})$. Halvorson en Ziegler hebben vooral aandacht besteed aan het verband tussen V en ρv bij gegeven n . Zij hebben tabellen vervaardigd van V voor $n = 10$ en $n = 100$ en voor ρv lopende van 0,4 tot 5,0. Zij constateren dat het minimum van V voor $n=10$ ligt bij $\rho v=1.2$, d.w.z. als de kans op een positief resultaat $1-e^{-\rho v}$ gelijk is aan 0,70 en voor $n=100$ bij $\rho v=1,5$ dus bij $1-e^{-\rho v}=0,78$.

Dit resultaat is in zoverre merkwaardig, dat men wellicht maximale nauwkeurigheid zou verwachten bij $e^{-\rho v}=\frac{1}{2}$, dus bij gelijke kansen op een positief dan wel een negatief monster. Voor $e^{-\rho v}=\frac{1}{2}$ geldt, dat de kans op een ongedetermineerd resultaat (zie opmerking onder formule (6)) minimaal is. De waarde $v=\frac{1}{\rho} \ln 2$ die voldoet aan $e^{-\rho v}=\frac{1}{2}$ wordt daarom door J.K.HOSKINS en C.T.BUTTERFIELD (zie [15]) het gunstigste monster-volume genoemd. Verwijzend naar de resultaten van Halvorson en Ziegler adviseren deze schrijvers echter om het monster-volume steeds iets groter dan $\frac{1}{\rho} \ln 2$ te kiezen.

HALVORSON en ZIEGLER constateren verder dat de variatiecoëfficiënt V zeer sterk van de coli-dichtheid ρ afhangt. Aangezien men in de praktijk voor de aanvang der experimenten zeer weinig van de grootte van ρ weet, zal men v niet zo kunnen kiezen, dat de nauwkeurigheid zo groot mogelijk is en zal men met een nauwkeurigheid genoegen moeten nemen, die veel geringer is dan de maximale. Op grond hiervan prefereren de schrijvers het gebruik van meerdere verdunningen, voor welk geval volgens hen dit bezwaar niet geldt (zie 4.4).

Het "Mathematisch Centrum" heeft het verband tussen V en ρ v nader onderzocht, terwijl het probleem van de maximale nauwkeurigheid ook met een andere methode benaderd is. Het een en ander zal in het aanvullend rapport worden behandeld.

4.4 Onderzoekingen van Halvorson en Ziegler omtrent de nauwkeurigheid van de aannemelijkste schatting bij gebruik van verschillende verdunningen.

In het geval van verschillende verdunningen kent men de kans op het optreden van proefresultaat (7) uit vergelijking (8). Bij ieder proefresultaat behoort één aannemelijkste schatting $\hat{r}$ van ρ , die verkregen wordt door de oplossing van vergelijking (9). Op deze wijze kan men bij iedere gegeven ρ de theoretische verdeling van $\hat{r}$ berekenen.

Zoals blijkt uit de opmerking onder (9a) is er steeds een positieve kans, dat $\hat{r}=\infty$ is. Hieruit volgt dat de verdeling van $\hat{r}$ geen eindige momenten en in bijzonder geen eindig gemiddelde noch een eindige spreiding bezit. Wil men toch iets over de momenten beweren, dan zal men de verdeling van $\hat{r}$ moeten afknotten door de ongedetermineerde en eventueel een aantal andere resultaten buiten beschouwing te laten. Dit heeft uiteraard alleen zin, als de weggelaten resultaten een zeer kleine waarschijnlijkheid hebben.

HALVORSON en ZIEGLER hebben (zie [12]) dergelijke afgeknotte verdelingen berekend, zonder er uitdrukkelijk de aandacht op te vestigen dat zij hiermee bovengenoemde moeilijkheden uit de weg gaan.

Genoemde schrijvers hebben deze berekeningen uitgevoerd voor $v_1=10$, $v_2=1$ en $v_3=0,1$ en

$$n = (n_1 = n_2 = n_3) = 10 ; \rho = 0,15 ; 0,25 ; 0,40 \text{ en } 1,5.$$

Zij bepaalden hierbij de variatiecoëfficiënten:

$$V = \frac{\text{spreiding van } \hat{r}}{\text{gemiddelde van } \hat{r}};$$

$$V' = \frac{(2\text{e moment van } \hat{r} \text{ t.o.v. } \rho)^{\frac{1}{2}}}{\rho}$$

Deze twee definities vertonen een analogie omdat de spreiding van $\hat{r}$ de wortel uit het 2e moment van $\hat{r}$ t.o.v. het gemiddelde van $\hat{r}$ is. De momenten in bovenstaande definities hebben uiteraard betrekking op de afgeknotte verdelingen. Bij hun berekeningen hebben Halvorson en Ziegler volgens bovenstaande definitie voor V' gewerkt,

doch. in de begeleidende tekst schrijven zij ten onrechte "modus van $\hat{r}$ " in plaats van ρ .

De resultaten kan men in het volgende tabelletje samenvatten:

$\rho = 0,15$	0,25	0,40	1,5
$V = 0,40$	0,41	0,41	0,42
$V' = 0,45$	0,49	0,47	0,47

Daarnaast hebben de schrijvers in een figuur een grafische voorstelling van de verdelingen van $\hat{r}$ bij deze vier waarden van ρ getekend. Aan de hand van de figuur en van de berekende variatiecoëfficiënten komen zij tot de conclusie, dat bij gebruik van drie verdunningen de nauwkeurigheid van de resultaten nagenoeg onafhankelijk is van de coli-dichtheid.

Verder voerden zij ook berekeningen uit voor:

$$\rho = 1,5 ; n = 5, 10, 20 \text{ en } 40$$

In deze gevallen hebben de schrijvers staarten van de verdeling van $\hat{r}$ bepaald die samen een waarschijnlijkheid ongeveer gelijk aan 0,03 bezitten, waarbij echter weer een aantal mogelijkheden met zeer kleine waarschijnlijkheid (met inbegrip van de ongedetermineerde resultaten) is weggelaten. De wijdte van het interval tussen de staarten (d.i. een voorspellingsinterval¹⁾ voor $\hat{r}$ behoudens een waarschijnlijkheid $\sim 0,03$) wordt door hen als maat voor de nauwkeurigheid van $\hat{r}$ gebruikt. Het blijkt dat de nauwkeurigheid van $\hat{r}$ toeneemt als n groter wordt, hetgeen te verwachten was. De invloed van de stijging van n is het sterkst bij kleine n . Verder constateren de schrijvers dat de verdeling van $\hat{r}$ meer symmetrisch wordt als n toeneemt.

Uit het bovenstaande blijkt, dat het mogelijk is om de verdeling van $\hat{r}$ bij gegeven n en ρ exact te berekenen. In de praktijk zal dit echter een zeer tijdrovend werk zijn omdat het aantal mogelijke proefresultaten zeer groot is; bij ieder van deze proefresultaten moeten omslachtige berekeningen worden uitgevoerd. Wil men resultaten bereiken, dan moet de verdeling van $\hat{r}$ bekend zijn bij een groot aan-

1) Een voorspellingsinterval voor een stochastische grootheid x , behoudens een waarschijnlijkheid α , is een vast interval waar x met de waarschijnlijkheid $1 - \alpha$ in valt. (Vergelijk het memorandum S 47 (M 18), "Betrouwbaarheidsintervallen", (Bijlage V)).

tal waarden van de coli-dichtheid, terwijl ook nog variatie mogelijk is in de opzet van de experimenten.

De waarde van de variatiecoëfficiënten als nauwkeurigheidsmaten is zeer dubieus, eerstens omdat deze coëfficiënten sterk afhankelijk zijn van de grenzen, waarbij de verdeling van $\hat{r}$ wordt afgeknot en verder omdat aan de spreiding en het tweede moment t.o.v. ρ geen waarschijnlijkheidstheoretische interpretatie kan worden gegeven, daar de verdeling van $\hat{r}$ een zeer onregelmatig karakter heeft en geenszins normaal is. Het is te betreuren dat de schrijvers niet meer van voorspellingsintervallen gebruik gemaakt hebben, zoals zij bij een deel van hun onderzoek gedaan hebben. In 4.5 en 4.6 zullen wij methoden bespreken, die niet aan de bezwaren van de variatiecoëfficiënten onderhevig zijn.

4.5 Bepaling van betrouwbaarheidsintervallen¹⁾ met behulp van de tweede schatting van Fisher.

Wij beschouwen verdunningsexperimenten, waarbij de monster-volumina $v_1, \dots, v_m$ zijn (m is het aantal verdunningen), terwijl van iedere verdunning een gelijk aantal monsters n genomen wordt. (Dit laatste is niet essentieel; wij voeren deze veronderstelling slechts in ter vereenvoudiging van de formules.) Het totale aantal positieve monsters van alle verdunningen stellen wij voor door s . Er geldt dus:

$$(25) \quad s = \sum_{i=1}^m p_i$$

s is geheel en wij hebben:

$$(26) \quad 0 \leq s \leq mn,$$

(hetgeen volgt uit $0 \leq p_i \leq n$ voor $i = 1, 2, \dots, m$).

De verdeling van s bij gegeven ρ is als volgt:

$$(27) \quad P[s = s | \rho] = \sum_{\sum_{i=1}^m p_i = s} \left\{ \prod_{i=1}^m \frac{n!}{p_i! (n-p_i)!} e^{-\rho v_i (n-p_i)} (1-e^{-\rho v_i})^{p_i} \right\}$$

Hulpstelling I: De kans $P[s \leq s | \rho] = \sum_{k=0}^s P[s = k | \rho]$

is een monotoon dalende continue functie van ρ als $s < mn$. Deze functie is gelijk aan 1 als $\rho = 0$ is en nadert tot 0 als ρ oneindig wordt. De functie is constant en gelijk aan 1, als $s = mn$ is.

De kans $P[s \geq s | \rho] = \sum_{l=s}^{mn} P[s = l | \rho]$ is een monotoon stijgende continue functie van ρ als $s > 0$ is. Deze functie is ge-

1) Voor het begrip betrouwbaarheidsinterval zie men memorandum S 47 (M 18) "Betrouwbaarheidsintervallen" (Bijlage V).

lijk aan 0, als $\rho=0$ is en nadert tot 1 voor $\rho \rightarrow \infty$. Voor $s=0$ is deze functie constant gelijk aan 1.

De monotonie van genoemde functies kan men bewijzen door de afgeleide te bepalen. De andere beweringen van Hulpstelling I zijn direct in te zien.

Hulpstelling II. De vergelijking:

$$(28) \quad P[s \leq s | \rho] = \frac{\alpha}{2} \quad (0 < \alpha < 1; s < mn),$$

heeft één en slechts één oplossing naar ρ .

De vergelijking

$$(29) \quad P[s \geq s | \rho] = \frac{\alpha}{2} \quad (0 < \alpha < 1; s < mn),$$

heeft eveneens één en slechts één oplossing naar ρ .

Hulpstelling II volgt direct uit hulpstelling I.

Wij stellen de oplossing van (28) voor door R_s en de oplossing van (29) voor door r_s en definiëren $R_{mn} = \infty$ en $r_0 = 0$ (zie fig.2). Wij zien gemakkelijk in dat geldt:

$$(30) \quad r_s < R_s, \quad r_s < r_{s+1} \text{ en } R_s < R_{s+1}.$$

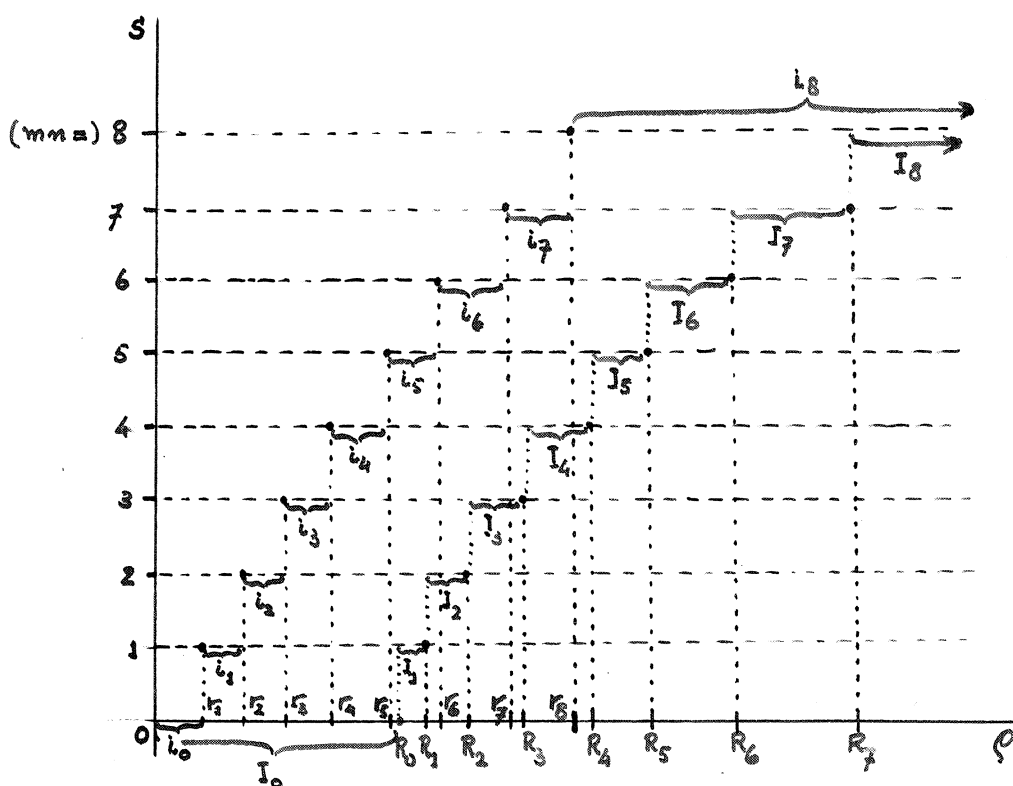


fig. 2: s = totaal aantal positieve monsters,

ρ = coli-dichtheid,

r_s is gedefinieerd door: $P[s \geq s | r_s] = \frac{\alpha}{2}$

R_s " " " : $P[s \leq s | R_s] = \frac{\alpha}{2}$

i_s is het interval tussen de punten: (r_s, s) en (r_{s+1}, s)

I_s " " " " " " (R_{s-1}, s) en (R_s, s)

$(R_0 = 0, r_{mn+1} = \infty)$

Beschouwen wij nu het totale aantal positieve monsters s weer als een stochastische grootheid, dan zijn dus ook $\underline{r}_s$ en $\underline{R}_s$ stochastisch. Wij bewijzen nu de

Hoofdstelling: Het interval $\underline{r}_s < \rho < \underline{R}_s$ is een betrouwbaarheidsinterval voor ρ met onbetrouwbaarheidsdrempel α d.w.z.

$$P[\underline{r}_s < \rho < \underline{R}_s | \rho] \geq 1 - \alpha.$$

Bewijs: Indien $\underline{r}_s < \rho < \underline{r}_{s+1}$ is, dan geldt wegens Hulpstelling I:

$$P[s \geq s+1 | \rho] \leq P[s \geq s+1 | \underline{r}_{s+1}] = \frac{\alpha}{2}$$

dus:

$$P[s > s | \rho] \leq \frac{\alpha}{2}$$

Op analoge wijze kan men bewijzen dat geldt:

$$P[s < s | \rho'] \leq \frac{\alpha}{2} \text{ als } \underline{R}_{s-1} \leq \rho' < \underline{R}_s$$

Wij stellen nu het lijnstuk, bestaande uit de punten (ρ, s) , waarvoor $\underline{r}_s < \rho \leq \underline{r}_{s+1}$ geldt als $s < mn$ is, en $\rho > \underline{r}_{mn}$ als $s = mn$ is, voor door i_s ; evenzo stellen wij het lijnstuk bestaande uit de punten (ρ', s) , waarvoor $\underline{R}_{s-1} \leq \rho' < \underline{R}_s$ geldt als $s > 0$ is, en $0 \leq \rho' < \underline{R}_0$ als $s = 0$ is voor door I_s . Iedere rechte evenwijdig aan de r -as, dus waarvoor $\rho = \text{constan}$ is, zal één punt gemeen hebben zowel met één der intervallen i_s als met één der intervallen I_s . Indien nu de werkelijke coli-dichtheid gelijk is aan ρ_0 en de rechte $\rho = \rho_0$ heeft één punt gemeen met elk van de beide intervallen I_{s_2} en i_{s_2} dan geldt:

$$\underline{1e.} \quad P[s_1 \leq s \leq s_2 | \rho_0] = 1 - P[s < s_1 | \rho_0] - P[s > s_2 | \rho_0] \geq 1 - \alpha$$

$$\underline{2e.} \quad \underline{r}_s < \rho_0 < \underline{R}_s \text{ geldt dan en slechts dan, als } s_1 \leq s \leq s_2 \text{ geldt, dus is:}$$

$$P[\underline{r}_s < \rho_0 < \underline{R}_s | \rho_0] = P[s_1 \leq s \leq s_2 | \rho_0],$$

en derhalve (zie 1e.):

$$P[\underline{r}_s < \rho_0 < \underline{R}_s | \rho_0] \geq 1 - \alpha,$$

waarmee de Hoofdstelling bewezen is.

Indien de opzet van de experimenten $(v_1, \dots, v_m \text{ en } n)$ eenmaal gegeven is, kan men een tabel vervaardigen van $r_0, r_1, \dots, r_{mn}, R_0, R_1, \dots, R_{mn}$. Bij iedere gevonden s leest men in deze tabel dan direct het betrouwbaarheidsinterval

(r_s, R_s) voor ρ af. Deze methode is betrekkelijk eenvoudig en heeft in zoverre haar verdiensten. Zoals wij zullen zien in 4.6 kan men op een andere wijze bij dezelfde onbetrouwbaarheidsdrempel nauwere betrouwbaarheidsintervallen voor ρ bepalen.

De hierboven beschreven methode is afkomstig van T.MATUSZEWSKI, J.NEYMAN en J.SUPINSKA (zie [16]). Wij hebben de bewijsvoering van deze schrijvers gewijzigd op sommige punten waar een verouderde redeneringswijze werd toegepast.

De stochastische grootheid $\underline{s} = \sum_{i=1}^m \underline{p}_i = mn - \sum_{i=1}^m \underline{q}_i$, die we hier gebruikt hebben is als het ware het complement van de stochastische grootheid $\sum_{i=1}^m \underline{q}_i$, die ten grondslag ligt aan de tweede schatting van Fisher (zie 3.2); de verdelingen van beide grootheden zijn symmetrisch ten opzichte van $\frac{1}{2}mn$. De breedte van de met bovenstaande methode verkregen betrouwbaarheidsintervallen kan ons dus tevens een inzicht verschaffen in de nauwkeurigheid van de tweede schatting van Fisher. In 4.6 zullen wij op grond hiervan tot de conclusie komen, dat de aannemelijkste schatting veel nauwkeuriger is.

4.6 Bepaling van betrouwbaarheidsintervallen met behulp van de aannemelijkste schatting.

Zoals wij in 4.4 gezien hebben is het in principe mogelijk de verdeling van $\hat{r}$ te berekenen, als de coli-dichtheid ρ en het schema van het experiment (n_i en v_i voor alle verdunningen) gegeven zijn. Indien men deze verdeling voor een groot aantal waarden van ρ kent, kan men een methode aangeven om ~~an~~ betrouwbaarheidsintervallen voor ρ te bepalen. Het is echter in 4.4 eveneens gebleken, dat het ondoenlijk is om de verdeling van $\hat{r}$ bij een groot aantal waarden van ρ exact te berekenen. Men kan echter trachten de exacte berekening door een benadering te vervangen. Een dergelijke benadering kan men interpreteren als een onzekerheid in de grenzen van het onbetrouwbaarheidsinterval, maar desgewenst ook als een onzekerheid in de betrouwbaarheidsdrempel.

Een goede benadering van de verdeling van $\hat{r}$ kan men nu verkrijgen door de frequentie-verdeling in een voldoende grote en representatieve steekproef van $\hat{r}$ te bepalen.

Om een steekproef uit de verdeling van $\hat{r}$ (bij gegeven ρ) te verkrijgen kan men als volgt te werk gaan:

1e. Men berekent voor iedere verdunning apart de theoretische verdeling van het aantal positieve monsters volgens:

$$(31) \quad P_{i,k} = P[p_i = k] = \binom{n_i}{k} e^{-\rho v_i} (1 - e^{-\rho v_i})^k$$

Verg. (5), $i=1,2,\dots,m$; $k=0,1,2,\dots,n_i$.

2e. Men neemt m grote steekproeven van de uitgebreidheid N uit een homogene $(0,1)$ verdeling; dit is een verdeling waarvoor geldt:

$$P[\underline{x} \leq x] = \begin{cases} 0 & \text{als } x < 0 \\ x & \text{als } 0 \leq x \leq 1 \\ 1 & \text{als } x > 1 \end{cases}$$

We duiden deze steekproeven aan met:

$$x_{i,1}, x_{i,2}, \dots, x_{i,N} \quad (i=1,2,\dots,m).$$

In tabellen van "random numbers" (b.v. M.G.Kendall en B.Babington Smith, Random Sampling Numbers) vindt men zeer uitgebreide steekproeven uit een homogene verdeling.

3e. Men past op de steekproeven de volgende transformatie $y = g_i(x)$ toe:

$$\begin{aligned} y &= 0 & \text{als } 0 < x < P_{i,0} \\ y &= 1 & \text{als } P_{i,0} < x < P_{i,0} + P_{i,1} \\ &\vdots \\ y &= k & \text{als } \sum_{v=0}^{k-1} P_{i,v} < x < \sum_{v=0}^k P_{i,v} \\ &\vdots \\ y &= N & \text{als } 1 - P_{i,N} < x < 1 \end{aligned}$$

Wij verkrijgen zodoende m getransformeerde steekproeven

$$y_{i,1}, y_{i,2}, \dots, y_{i,N} \quad (i = 1, 2, \dots, m)$$

4e. Men substitueert in vergelijking (9) achtereenvolgens:

$$\begin{aligned} p_1 &= y_{1,1}, p_2 = y_{2,1}, \dots, p_m = y_{m,1} \\ p_1 &= y_{1,2}, p_2 = y_{2,2}, \dots, p_m = y_{m,2} \\ &\vdots \\ p_1 &= y_{1,N}, p_2 = y_{2,N}, \dots, p_m = y_{m,N} \end{aligned}$$

en verkrijgt zodoende N vergelijkingen. De oplossingen van deze vergelijkingen naar r noemen wij $r_1, r_2, \dots, r_N$.

Men kan nu zonder grote moeite aantonen, dat $r_1, \dots, r_N$ een steekproef van de uitgebreidheid N uit de verdeling van $\hat{r}$ is. Hiermee is, zoals wij gezien hebben, tevens een experimentele benadering voor de verdeling van $\hat{r}$ verkregen.

Op deze wijze kan men bij iedere ρ approximatief de verdeling van $\hat{r}$ bepalen, dus, eveneens approximatief, de kleinste door $\hat{r}$ aangenomen waarde $R(\rho)$ waarvoor geldt:

$$P[\hat{r} \geq R(\rho)] \leq \frac{\alpha}{2}$$

en de grootste door $\hat{r}$ aangenomen waarde $r(\rho)$, waarvoor geldt

$$P[\hat{r} \leq r(\rho)] \leq \frac{\alpha}{2}.$$

Men zal de hypothese $\rho = \rho_0$ kunnen verwerpen behoudens een waarschijnlijkheid $\leq \alpha$ als geldt:

$$\hat{r} \geq R(\rho_0) \text{ of } \hat{r} \leq r(\rho_0).$$

Een betrouwbaarheidsinterval voor ρ verkrijgt men dus¹⁾ door de verzameling te nemen van alle waarden ρ_0 van ρ , die op grond van de bij een experiment verkregen aannemelijkste schatting $\hat{r} = \hat{r}_0$ niet voor verwerping in aanmerking komen. D.i. dus de verzameling van die waarden ρ_0 van ρ , waarvoor geldt:

$$r(\rho_0) \leq \hat{r}_0 \leq R(\rho_0)$$

In de praktijk gaat men als volgt te werk. Men tekent in een $(\rho, \hat{r})$ -diagram een aantal puntenparen $(\rho, r(\rho))$ en $(\rho, R(\rho))$ en past bij de punten $(\rho, r(\rho))$ een kromme k , bij de punten $(\rho, R(\rho))$ een kromme K aan. Vervolgens bepaalt men de snijpunten $(\rho_1, \hat{r}_0)$ en $(\rho_2, \hat{r}_0)$ van de rechte $\hat{r} = \hat{r}_0$ met respectievelijk k en K . Het betrouwbaarheidsinterval voor ρ wordt nu begrensd door ρ_1 en ρ_2 (zie fig.3).

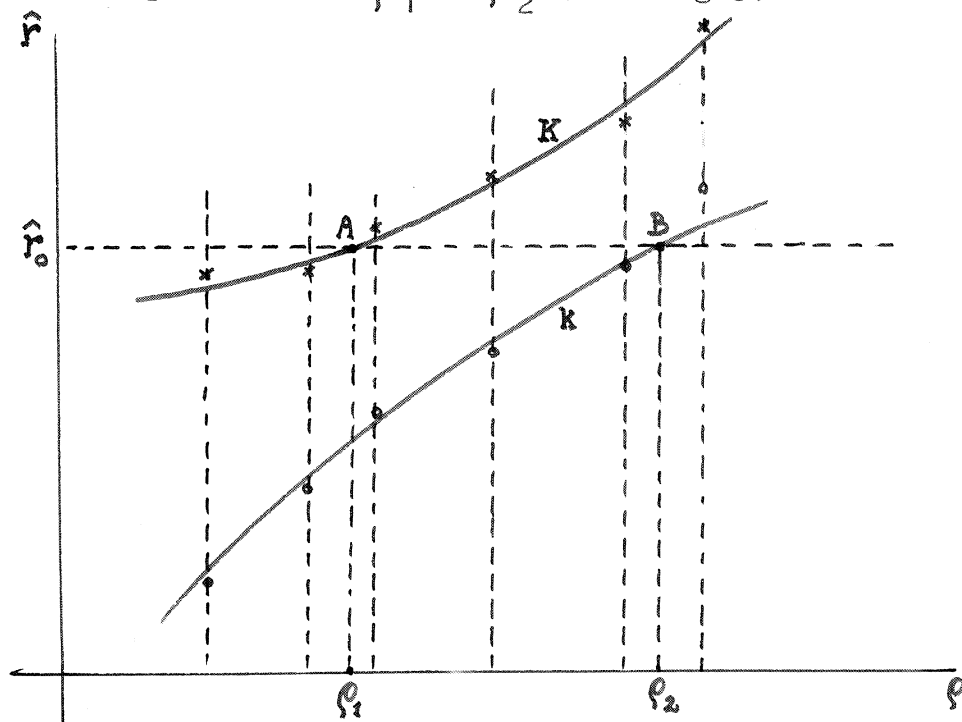


fig. 3. (Toelichting, zie blz 27)

1) Zie memorandum S 47 (M 18), "Betrouwbaarheidsintervallen" (Bijlage V).

fig. 3. ρ = coli-dichtheid,
 $\hat{r}$ = aannemelijkste schatting van de coli-dichtheid,
 $\times$ = punten $(\rho, R(\rho))$,
 $\circ$ = punten $(\rho, r(\rho))$,
 $R(\rho)$ = kleinste door $\hat{r}$ aangenomen waarde, die voldoet aan $P[\hat{r} \geq R(\rho) | \rho] \leq \frac{\alpha}{2}$,
 $r(\rho)$ = grootste door $\hat{r}$ aangenomen waarde, die voldoet aan $P[\hat{r} \leq r(\rho) | \rho] \leq \frac{\alpha}{2}$,
 k is kromme aangepast bij de punten $\circ$,
 K " " " " " " " $\times$.
De rechte $\hat{r} = \hat{r}_0$ snijdt K in $A(\rho_1, r_0)$ en k in $B(\rho_2, r_0)$,
 $\rho_1 < \rho < \rho_2$ is een betrouwbaarheidsinterval voor ρ , als $\hat{r} = \hat{r}_0$, met een onbetrouwbaarheidsdrempel $\approx \alpha$.

De hierboven beschreven methode is afkomstig van Mej. J. SUPINSKA (zie [16]) en door haar uitgewerkt voor het geval $v_1=10$, $v_2=1$ en $v_3=0,1$. De resultaten van Mej. Supinska zijn verenigd in een grafiek, waarvan een fotocopie aan dit rapport is toegevoegd (Bijlage VI).

De in deze grafiek horizontaal uitgezette grootte λ' correspondeert met $\hat{r}$, de verticaal uitgezette grootte λ correspondeert met ρ , voor de aangepaste krommen k en K zijn hier parabolen genomen. Na het bovenstaande is het duidelijk hoe men deze grafiek kan gebruiken. De opmerking uit de laatste alinea van 2.5 is hier eveneens van toepassing; indien $v_1=k \times 10$, $v_2=k$ en $v_3=k \times 0,1$ is, vindt men met behulp van de grafiek betrouwbaarheidsintervallen voor $k \cdot \rho$.

Indien men de breedte van de betrouwbaarheidsintervallen van Mej. Supinska wil vergelijken met de breedte van betrouwbaarheidsintervallen berustend op het totaal aantal positieve monsters $\underline{s}$, is de grote moeilijkheid, dat de verschillende proefresultaten, waarvoor $\underline{s}$ hetzelfde is, verschillende betrouwbaarheidsintervallen leveren volgens de methode van Mej. Supinska, doch slechts één betrouwbaarheidsinterval, volgens de methode, die berust op de waarde van $\underline{s}$. De schrijvers van [16] hebben de breedten van de betrouwbaarheidsintervallen van 4.5 voor een aantal waarden van $\hat{r}$ berekend en het kleinste dezer intervallen vergeleken met het kleinste interval in de door Mej. Supinska geconstrueerde figuur 3. Daarbij hebben zij om een vergelijking mogelijk te maken, in beide gevallen de breedte gedeeld door de

onderste grens van het interval. Deze vergelijking valt sterk ten gunste van de methode van Mej. SUPINSKA uit. Men heeft hiermee een aanwijzing dat de aannemelijkste schatting nauwkeuriger is dan de tweede schatting van FISHER (zie 3.2 en 4.5). Dit is te verwachten op grond van het feit, dat de aannemelijkste schatting vollediger van de gegevens gebruik maakt dan de tweede schatting van Fisher.

W.G.COCHRAN veronderstelt om betrouwbaarheidsintervallen voor ρ te kunnen bepalen, dat $\log \hat{r}$ normaal verdeeld is met een spreiding $\sigma_{\log \hat{r}} \approx 0,55 \sqrt{\frac{\log a}{n}}$ (a = verdunningsfactor, n = aantal monsters per verdunning) (zie [30]). Hij vermeldt niet hoe deze formule afgeleid is. In ieder geval geldt ook hier het bezwaar genoemd in 4.2, dat de spreiding van $\log \hat{r}$ in het geheel niet bestaat. De verdeling van $\log \hat{r}$ kan dus in ieder geval niet exact normaal zijn; voor men de methode van Cochran als benaderingsmethode kan toepassen, zal men eerst moeten nagaan in hoeverre de verdeling van $\log \hat{r}$ door een normale verdeling geapproximeerd wordt.

4.7 Enige opmerkingen over de gunstigste opzet van de experimenten.

Indien men experimenteert met monsters van gelijk volume is slechts de keuze van dit volume van belang. (Ten aanzien van het aantal monsters n kan immers worden opgemerkt, dat vergroting van n steeds bijdraagt tot verhoging van de nauwkeurigheid van de schatting. De grootte van n wordt dus bepaald door de experimentele mogelijkheden en de eisen, die men aan de nauwkeurigheid stelt).

Zijn er a priori geen gegevens over de grootte-orde van ρ , dan kan men uiteraard geen aanwijzingen geven over de keuze van v (en verdient het aanbeveling met meerdere monster-volumina te werken). Beschikt men wel over gegevens dan staat men voor het dilemma beschreven in 4.3; de grootste nauwkeurigheid kan men verwachten als $\rho v \approx 1,2$ à $1,5$ is, doch de kans op een ongedetermineerd resultaat is het kleinste als $e^{-\rho v} = 0,5$, dus $\rho v \approx 0,7$ is. In de praktijk speelt dit dilemma geen grote rol, omdat de gegevens betreffende ρ in het algemeen van dien aard zijn, dat men slechts zeer globale aanwijzingen kan geven voor de keuze van v .

Indien men meerdere verdunningen wenst te gebruiken, zal men ook over de keuze van meerdere grootheden moeten beslissen. Als wij ons hierbij beperken tot het geval dat

de monster-volumina een meetkundige reeks vormen en dat van iedere verdunning een gelijk aantal van n monsters wordt genomen, terwijl het totale aantal te gebruiken monsters mn door de experimentele mogelijkheden begrensd is, dan kan men het kleinste monster-volume v_L , de verdunningsfactor a en n nog kiezen.

W.G.COCHRAN (zie [30]) heeft grenzen aangegeven voor de keuze van v_L en v_H (het kleinste resp. grootste monster-volume) en wel $v_L \leq \frac{1}{\rho_H}$, $v_H \geq \frac{1}{\rho_L}$, waarin ρ_H de grootste en ρ_L de kleinste coli-dichtheid is, die men voor het te onderzoeken materiaal mogelijk acht. Afgezien van het feit, dat het minstens even belangrijk is een benedengrens voor v_L en een bovengrens voor v_H aan te geven, kan men aan de door Cochran aangegeven grenzen geen bijzonder grote waarde toekennen. Het is een op grond van overwegingen betreffende de nauwkeurigheid van experimenten met één verdunning en betreffende de kans op een ongedetermineerd resultaat vrij willekeurig gekozen globale regel, waarnaast men met evenveel recht een groot aantal andere dergelijke regels zou kunnen stellen.

De zelfde schrijver is zoals wij gezien hebben in 4.2 uit nauwkeurigheidsoverwegingen tot de conclusie gekomen, dat het aanbeveling verdient met een kleine verdunningsfactor te werken. Dit betekent dat, als v_L en v_H eenmaal gekozen zijn, men a klein, het aantal verdunningen m groot en dus n klein zal moeten kiezen. Het is echter tevens van belang rekening te houden met de kans op een ongedetermineerd resultaat. Deze kans wordt gegeven door:

$$(32) \quad P[\underline{s} = 0 \text{ of } \underline{s} = mn | \rho] = \prod_{i=1}^m e^{-\rho v_H a^{-(i-1)n}} + \prod_{i=1}^m (1 - e^{-\rho v_H a^{-(i-1)n}})^n$$

(waarin $\underline{s}$ het totale aantal positieve monsters voorstelt).

Men kan nu bij gegeven v_H , a , m en n een interval voor ρ bepalen, waarbinnen geldt:

$$(33) \quad P[\underline{s} = 0 \text{ of } \underline{s} = mn | \rho] \leq 0,05$$

Naarmate het interval breder is, is het experiment uit dit gezichtspunt gunstiger te achten. MATUSZEWSKI, NEYMAN en SUPINSKA, van wie deze methode afkomstig is (zie [16]) noemen het door (33) gedefinieerde interval de "range of 95 percent efficacy". Deze range wordt kleiner als a afneemt en groter als n toeneemt. Terwijl wij om de nauwkeurigheid zo groot mogelijk te maken a klein en derhalve n groot zul-

len moeten kiezen, leidt de overweging van de kans op een ongedetermineerd resultaat tot de tegengestelde keuze.

5. Conclusies.

5.1 Proeven met verschillende verdunningen.

Indien men een proef verricht heeft met verschillende verdunningen is de aannemelijkste schatting, die men verkrijgt door vergelijking (9) op te lossen, de nauwkeurigste puntschatting van de coli-dichtheid. De berekening van deze schatting is vrij bewerkelijk, doch zij is voor een aantal veel voorkomende gevallen getabelleerd (zie 2.5). Wenst men een interval-schatting, dan is de methode van Mej. J.SUPINSKA (zie 4.6), die op de aannemelijkste schatting berust, het meeste aan te bevelen. Deze methode is echter slechts voor één geval volledig uitgewerkt; het zal veel werk vergen om met deze methode in andere gevallen betrouwbaarheidsintervallen te berekenen. De tweede schatting van FISHER (zie 3.2) is, indien men beschikt over de "Statistical tables" van Fisher en Yates, veel minder bewerkelijk doch ook minder nauwkeurig dan de aannemelijkste schatting. Op grond van deze methode kan men betrouwbaarheidsintervallen voor de coli-dichtheid aangeven; deze zijn gemakkelijker te bepalen maar ook breder dan de betrouwbaarheidsintervallen van Mej. J.SUPINSKA. De grafische methoden van Ir LEEFLANG en Dr Ir HASPERS (zie 3.3 en 3.4) zijn slechts als oriënterende methoden te gebruiken zolang men niet aangeeft, hoe men de S-krommen van Ir Leeftlang en de rechten van Dr Ir Haspers zal aanpassen. Deze methoden zijn zeker niet nauwkeuriger dan de aannemelijkste schatting en het staat te bezien of zij sneller tot een resultaat voeren dan de tweede schattingsmethode van Fisher. Het is vrijwel ondoenlijk een algemene regel aan te geven, hoe groot men bij een proef met meerdere verdunningen het aantal verdunningen, de verdunningsfactor en het aantal monsters per verdunning moet kiezen. Dit zal van tal van praktische overwegingen afhangen, die van geval tot geval bekeken zullen moeten worden.

5.2 Proeven met één verdunning.

De aannemelijkste en tevens nauwkeurigste schatting van de coli-dichtheid is in de gevallen, waarbij gewerkt wordt met één verdunning zo gemakkelijk te berekenen, dat andere schattingen niet in aanmerking komen. Het is tevens

niet moeilijk om hier betrouwbaarheidsintervallen voor de coli-dichtheid te bepalen, hoewel men hierover in de literatuur niets aantreft. Dit onderwerp wordt in ons aanvullend rapport behandeld. De beschouwing betreffende de variatie-coëfficiënt van HALVORSON en ZIEGLER (zie 4.3), stelt ons in staat een gunstig monster-volume te kiezen, als men over enige voorkennis van ρ beschikt.

Het komt ons voor, dat deze voordelen van het werken met één verdunning niet voldoende beseft worden. Nagenoeg alle schrijvers houden zich hoofdzakelijk bezig met de proeven, waarbij meerdere verdunningen worden gebruikt. De methode met één verdunning wordt toegepast bij de samenstelling van norm-voorschriften voor drinkwater. Dit onderwerp zal echter in het aanvullend rapport behandeld worden. In het bijzonder indien men een betrouwbaarheidsinterval wil bepalen, verdient het werken met één verdunning de voorkeur. De moeilijkheid van de keuze van het monster-volume kan opgelost worden doordat men vooraf een globale bepaling verricht met meerdere verdunningen, waarbij men kan volstaan met een eenvoudige schattingsmethode (b.v. de tweede schatting van Fisher).

Litteratuur en bijlagen.

De nummers tussen vierkante haken zijn verwijzingen naar de hier bijgevoegde litteratuurlijst (Bijlage I). Een systematisch overzicht van deze litteratuur treft men aan in Bijlage II. Bijlage III bestaat uit een lijst van een aantal oudere artikelen, die niet in dit rapport, noch in het litteratuuroverzicht van Bijlage II verwerkt zijn. Het betreft hier artikelen die inmiddels door latere publicaties achterhaald zijn.

De drie andere bijlagen zijn een memorandum (een kort oriënterend referaat) over "Enige definities uit de schattingstheorie" (Bijlage IV), een memorandum over "Betrouwbaarheidsintervallen" (Bijlage V) en een fotocopie van de grafiek van Mej. Supinska (Bijlage VI).

Schrijver(s)	Titel	Tijdschrift, Nr, Pag.
[1] M.H. Mac Crady (1915)	The numerical interpretation of fermentation-tube results	J. Inf. Dis. <u>17</u> (1915) 183 - 212
[2] A. Wolman and W. Weaver (1917)	A modification of the Mac Crady method of the numerical interpretation of fermentation-tube results	J. Inf. Dis. <u>21</u> (1917) 287 - 291
[3] M. Greenwood and G.U. Yule (1917)	On the statistical interpretation of some bacteriological methods employed in water analysis	J. Hyg. Camb. <u>21</u> (1917) 36 - 54
[4] M.F. Stein (1919)	The interpretation of B-coli-test results on a numerical and comparative basis as applied to the investigations of water and servage	Journ. Bact. <u>4</u> (1919) 243 - 265
[5] R.A. Fisher (1921 ^a)	Distribution of observations in a Dilution series (N.B. § 3 van het artikel (1921 ^d) On the mathematical foundation ^d of theoretical statistics	Phil. trans. Roy. Soc. <u>A 222</u> (1921) 363 - 366
[6] Advisory Committee on official water standards (1925)	Report of advisory committee on official water standards	Public Health Reports <u>40</u> (1925) 696 - 721
[7] L.J. Reed (1925)	B-coli-densities as determined from various types of samples N.B.: Appendix III by Advisory Committee on official water standards (1925)	Public Health Reports <u>40</u> (1925) 704 - 716
[8] J.H. Linschoten (1929)	Een nomografische methode voor de berekening van het aantal coli-kiemen in water	Water en Gas <u>3</u> (1929) 169-171

Schrijver(s)	Titel	Tijdschrift, Nr, Pag.
[9] E.S.Allen (1932)	The accuracy of the dilution method of estimating the density of a population of micro-organisms	Iowa St.Coll.J. of Sc. <u>6</u> (1932) 251 - 262
[10] H.O.Halvorson and N.R.Ziegler (1933 ^a)	Application of Statistics to problems in Bacteriology <u>I.</u> A means of determining Bacteriological population by the Dilution method	Journ.Bact. <u>25</u> (1933) 101 - 121
[11] H.O.Halvorson and N.R.Ziegler (1933 ^b)	<u>II.</u> A consideration of the accuracy of Dilution Data obtained by using a single dilution	Journ.Bact. <u>26</u> (1933) 331 - 339
[12] H.O.Halvorson and N.R.Ziegler (1933 ^c)	<u>III.</u> A consideration of the accuracy of Dilution Data using several dilutions	Journ.Bact. <u>26</u> (1933) 559 - 567
[13] J.K.Hoskins (1933)	The most probable number of B-coli in water analysis	J.Am.Water Works Ass. <u>25</u> (1933) 867 - 877
[14] J.K.Hoskins (1934) (zie ook: [22])	Most probable numbers for evaluation of coliaerogenes tests by fermentation-tube method	Public Health Reports <u>49</u> (1934) 393 - 405
[15] J.K.Hoskins and C.T.Butterfield (1935)	Determining the Bacteriological quality of drinking water	Journ.Am.Water Works Ass. <u>27</u> (1935) 1101 - 1109
[16] T.Matuszewski, J.Neyman and J.Supinska (1935)	Statistical studies in questions of bacteriology. I. Supinska The accuracy of the Dilution method.	J. Supinska Journ.Roy.Stat.Soc. <u>2</u> (1935) 63 - 82

Schrijver(s)	Titel	Tijdschrift, Nr, Pag.
[17] H. Barkworth and J.O. Irwin (1938)	Distribution of coli-form organisms in milk and the accuracy of the presump- tive coli-form test	J. Hyg. Camb. <u>38</u> (1938) 446 - 457
[18] S. Swaroop (1938 ^b)	Numerical estimation of B-coli by Dilution method	Ind. Journ. Med. Res. <u>26</u> (1938) 353 - 378
[19] R.D. Gordon (1939)	Estimating bacterial po- pulations by the dilution method	Biometrika <u>31</u> (1939) 168 - 180
[20] E.S. Pearson (1939)	Note on the inverse and direct methods of estima- ting in R.D. Gordon's pro- blem	Biometrika <u>31</u> (1939) 181 - 186
[21] Min. of Health London H.M.S.O. (1939)	The bacteriological exa- mination of water supplies VI. Suggested classifica- tion of water	
[22] J.K. Hoskins (1940)	Most probable numbers for evaluation of coli-aero- genes tests by fermenta- tion-tube results Verbeterde uitgave van no. [14]	Public Health Re- ports Reprint 1621 (1940)
[23] R. Pomeroy (1940)	Calculating most probable numbers in coli-aerogenes analysis	J. Am. Water Works Ass. <u>32</u> (1940) 478 - 486
[24] Drinking Water Standards (1946)	Drinking water standards 3. As to bacteriological quality	Public Health Re- ports Reprint 2697 (1946)
[25] Ir K.W.H. Leeftang (1946)	Een eenvoudige methode ter bepaling van het meest waarschijnlijke aantal B-coli in water	"Water" 22 Aug. '46
[26] W.L. Stevens (1947 ^a)	Densities of organisms es- timated by the Dilution method	Table VIII ² in "Sta- tistical tables" van Fisher and Yates 3rd edition (vervolg zie blz 4)

Schrijver(s)	Titel	Tijdschrift, Nr, Pag.
vervolg [26]		Oliver and Boyd (1949) Toelichting: B12 . 6 t/m 9 Tabel B12 . 49
[27] D.J.Finney (1947 ^c)	The principles of biological assay	Sup.Journ.Roy.Stat. Soc. IX (1947) 46 - 76
[28] Dr Ir J.H.Haspers (1949 ^a)	Grafische bepaling van het meest waarschijnlijke aantal B-coli in water	) Niet in druk gepubliceerde aantekeningen bij [25] .
[29] Dr Ir J.H.Haspers (1949 ^b)	Eenvoudige rekenkundige bepaling van het meest waarschijnlijke aantal B-coli in water	
[30] W.G.Cochran (1950)	Estimation of bacterial densities by means of the "most probable number"	Biometrics <u>6</u> (1950) 105 - 116
[31] J.W.Fertig and A.N.Heller (1950)	The application of statistical techniques to ser- vage treatment processes	Biometrics <u>6</u> (1950) 127 - 135

Afkortingen uit Litteratuurlijst.

J.Inf.Dis.	= Journal of Infectious Diseases
J.Hyg.Camb.	= Journal of Hygiene Cambridge
Journ.Bact.	= Journal of Bacteriology
Phil.trans.Roy.Soc.	= Philisophical Transactions of the Royal Society of London
Iowa St.Coll.J.of Sc.	= Iowa State College Journal of Science
J.Am.Water Works Ass.	= Journal of the American Water Works Associatio
Sup.Journ.Roy.Stat. Soc.	= Supplement to the Journal of the Royal Statis- tical Society
Ind.Journ.Med.Res.	= Indian Journal of Medical Research.

Systematisch litteratuuroverzicht.

Dit overzicht bevat een indeling van de litteratuur voorkomende in Bijlage I, naar de onderwerpen behandeld in het overzichtsrapport. De hieronder voorkomende nummers tussen vierkante haken corresponderen met de nummers in Bijlage I. De onderstreepte nummers corresponderen met de nummers der paragrafen in het rapport.

2. Aannemelijkste schatting van de coli-dichtheid.

- 2.1 Verdeling van de aantallen coli-bacteriën per monster: [3], [10], [30].
- 2.2 Aannemelijkste schatting van de coli-dichtheid (afleiding):
- 2.3 [5], [10], [30].
- 2.4 Enige opmerkingen betreffende de afleiding van verg.(9) (waaraan de aannemelijkste schatting voldoet): [5]; (afleiding volgens de inverse methode): [3].
- 2.5 Berekening van de aannemelijkste schatting: (met behulp van verg.(9)): [10], [14], [18], [23];
(Hulptabellen): [10], [14], [18];
(Tabellen van $\hat{r}$): [10], [14];
(Zonder verg.(9)): [13].
- 2.6 Successieve approximatie der aannemelijkste schatting volgens Finney (onjuiste methode): [27].

3. Diverse andere schattingen van de coli-dichtheid.

- 3.1 Schatting van Mac Crady: [1], [2], [8];
- 3.2 Tweede schattingsmethode van Fisher: [5], [26];
- 3.3 Grafische methode van Leeftang: [25];
- 3.4 Grafische methode van Haspers: [28].

4. Nauwkeurigheid van de schattingen.

- 4.1 Inverse methode: [4];
- 4.2 Asymptotische formule van Fisher: [5], [9], [18], [30];
- 4.3 Variatiecoëfficiënt van Halvorson en Ziegler voor één verdunning: [11];
(Kans op een ongedetermineerd resultaat): [15].
- 4.4 Onderzoekingen van Halvorson en Ziegler betreffende de nauwkeurigheid van de aannemelijkste schatting bij het gebruik van meerdere verdunningen: [12].
- 4.5 Bepaling van betrouwbaarheidsintervallen met behulp van de tweede schatting van Fisher: [16].
- 4.6 Bepaling van betrouwbaarheidsintervallen met behulp van de aannemelijkste schatting: (Experimentele benaderingsmethode): [16];
(Suggesties): [20];
(Benadering met normale verdeling van $\log \hat{r}$): [30].

4.7 Enige opmerkingen over de gunstigste opzet van de experimenten: [15], [30], [31];
(Range of 95 percent efficacy): [16].

Normvoorschriften (komen in het aanvullingsrapport ter sprake):
[6], [7], [21], [24].

Overzichtsartikelen: [10], [30].

Tabellen: Zie onder 2.5 en [26] (tweede schatting van Fisher).

De volgende, over het algemeen oudere artikelen
zijn niet in het rapport opgenomen.

No.	Schrijver(s)	Titel	Tijdschr., Jaarg., Pag.
1	E.B.Phelps (1908)	A method for calculating the number of B-coli from the results of dilution tests.	Am.J.Pub.Hyg. <u>4</u> (1908) 141 - 145
2	G.C.Whipple	The element of chance in sanitation	Journ.Franklin Inst. <u>182</u> (1916) nr 1 en 2
3	M.F.Stein (1917)	Making the B-coli test tell more	Eng.News-Rec. <u>78</u> (1917) 391
4	M.F.Stein (1918)	A critical study of the bacteriological count in water and sewage	Am.J.Pub.Health <u>8</u> (1918) 820 - 829
5	M.H.Mc Crady [1918]	Tables for rapid interpretation of fermentation-tube results	The public Health Journal (Canada) <u>9</u> (1918) no 5.
6	M.F.Stein (1919)	The numerical interpretation of bacteriological tests	Eng.News ^{Rec.} <u>82</u> (1919) 1106 - 1109
7	A.Wolman (1919)	Index numbers and scoring of water supplies	Am.Water Works Ass. <u>6</u> (1919) no 3.
8	W.F.Wells (1919 ^a)	The bacteriological dilution scale and the dilution as a bacteriological unit	Am.Journ.Pub.Health <u>9</u> (1919) no 9
9	W.F.Wells (1919 ^b)	On a standard system of bacteriological dilutions	Am.Journ.Pub.Health <u>9</u> (1919) no 12.
10	American Public Health Association (1920), (1941)	Standard methods for the examination of water and sewage	(Boek): Boston 4th Ed. 1920 8th Ed. 1941
11	P.V.Wells & } (1922) W.F.Wells }	Standard bacterial Index	Am.Water Works Ass. <u>9</u> (1922) no 3.
12	M.F.Stein (1922)	The bacteria-coli test. Relation between positive results in samples of one and ten cubic centimeters	Engineering and Contracting <u>57</u> (1922) 445 - 446
13	H.W.Streeter (1933)	The bacterial efficiency of certain intermediate stages of water treatment	Public Works <u>64</u> (1933) 17 - 20
14	Ministry of Health (1934)	Report of the bacteriological examination of water-supply.	Min.of Health (1934) Report no 71

Afkortingen:

Am.Journ.Pub.Hyg.	= American Journal of Public Hygiene
Journ. Franklin Inst.	= Journal of the Franklin Institute
Eng.News-Rec.	= English News-Record
Am.Journ.Pub.Health	= American Journal of Public Health
J.Am.Water Works Ass.	= Journal of the American Water Works Association
Min.of Health	= Ministry of Health

MATHEMATISCH CENTRUM,
2de Boerhaavestr. 49,
A m s t e r d a m - 0.
Statistische Afdeling
S 53 (M 21)

Enige definities uit de schattingstheorie.¹⁾

1. Inleiding.

Een veel voorkomend probleem in de statistiek is het schatten van één of meer onbekende parameters (b.v. het gemiddelde of de spreiding) van een overigens bekende waarschijnlijkheidsverdeling W , indien een aantal waarnemingen gegeven zijn. Men schat de onbekende parameter dan door een functie van de waarnemingen. Zijn $x_1, \dots, x_n$ de waarnemingen en zijn deze onderling onafhankelijk dan kan men b.v. het gemiddelde van de waarschijnlijkheidsverdeling W schatten door de grootheid

$$(1) \quad \bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i,$$

dus door het gemiddelde van de waarnemingen. Een dergelijke schatting geeft in het algemeen niet de werkelijke waarde van de gezochte parameter, maar een daarbij in de buurt liggende waarde. Men heeft steeds de keus tussen een aantal verschillende functies van $x_1, \dots, x_n$ als schattingen voor dezelfde onbekende parameter. Zo kan men b.v. als schatting voor het gemiddelde van de waarschijnlijkheidsverdeling ook nemen de functie

$$(2) \quad \bar{x}^* = \frac{2}{n(n+1)} (x_1 + 2x_2 + \dots + nx_n) = \frac{2}{n(n+1)} \sum_{i=1}^n i \cdot x_i,$$

Men kiest nu uit deze mogelijkheden naar aanleiding van de eigenschappen, die de verschillende schattingen bezitten. Daartoe denkt men zich het experiment vele malen herhaald, d.w.z. men onderstelt, dat vele reeksen van onderling onafhankelijke waarnemingen $x_1, \dots, x_n$ zijn verricht en berekent voor ieder van deze reeksen de verschillende schattingen. Van ieder dezer schattingen verkrijgt men zodoende een groot aantal waarden, waaruit men eigenschappen omtrent de nauwkeurigheid der verschillende schattingen kan afleiden. In mathematische formulering komt dit daarop neer, dat men ieder

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

der waarnemingen $x_1, \dots, x_n$ beschouwt als een stochastische grootheid¹⁾, die verdeeld is volgens de waarschijnlijkheidsverdeling W , waarvan de onbekende parameter geschat moet worden. Deze stochastische grootheden $x_1, \dots, x_n$ zijn, indien de waarnemingen onafhankelijk zijn, bovendien onafhankelijk verdeeld, alle volgens deze zelfde waarschijnlijkheidsverdeling W . Wij beperken ons tot de schatting van één onbekende parameter van W , die wij met θ angeven en laten het geval van de gezamenlijke schatting van verschillende parameters buiten beschouwing.

De schatting van θ uit n waarnemingen $x_1, \dots, x_n$ geven we aan met

$$(3) \quad u_n = u_n(x_1, \dots, x_n)$$

daar u_n een functie van $x_1, \dots, x_n$ is en dus zelf ook een waarschijnlijkheidsverdeling bezit.

De definities van enkele eigenschappen uit de schattingstheorie geven wij eerst in populaire vorm en vervolgens (in § 3) in streng mathematische vorm.

2. Definities (populair).

A. Een schatting²⁾ u_n van een parameter θ heet zuiver²⁾, indien het gemiddelde van de waarschijnlijkheidsverdeling van u_n gelijk aan θ is.

Voorbeelden: De door (1) en (2) gegeven schattingen van het gemiddelde van de waarschijnlijkheidsverdeling W zijn beide zuivere schattingen. De schatting

$$(4) \quad s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

waarin $\bar{x}$ het door (1) gegeven gemiddelde der waarnemingen voorstelt, is een zuivere schatting van de variantie (het spreidingskwadraat) van W .

B. Een schatting u_n van θ heet asymptotisch zuiver³⁾, indien het gemiddelde van de waarschijnlijkheidsverdeling van u_n voor $n \rightarrow \infty$ tot θ nadert.

1) Een stochastische grootheid is een grootheid, die een waarschijnlijkheidsverdeling bezit; wij geven een dergelijke grootheid aan door onderstreepte letter, terwijl niet onderstreepte letters gewone algebraïsche grootheden of door stochastische grootheden aangenomen waarden voorstellen.

2) Engels: "estimate" resp. "unbiased". De vertaling der termen is van Prof. Dr D. van Dantzig; zie de litteratuurverwijzing aan het einde van dit memorandum.

3) Engels: "asymptotically unbiased".

Voorbeeld:

$$(5) \quad \underline{s}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

is een voor eindige n onzuivere, maar asymptotisch zuivere, schatting van de variantie van W .

C. Een schatting u_n van θ heet bruikbaar¹⁾, indien men, door het verrichten van een voldoende groot aantal waarnemingen kan zorgen dat, op een willekeurig kleine gegeven waarschijnlijkheid na, het verschil tussen de gevonden waarde van u_n en de gezochte waarde θ kleiner is dan een willekeurig klein gegeven getal. Anders uitgedrukt: indien men, op een willekeurig kleine waarschijnlijkheid na, iedere gewenste nauwkeurigheid kan bereiken door n voldoende groot te kiezen.

Voorbeeld: Vrijwel alle gebruikelijke schattingen zijn bruikbaar. De door (1) gegeven schatting van het gemiddelde van W gaat b.v. in een niet bruikbare schatting over, indien wij de factor $\frac{1}{n}$ vervangen door $\frac{1}{2n}$. Bij een toenemend aantal waarnemingen zal dan de schatting vrijwel zeker vlak bij de helft van het gemiddelde terecht komen.

D. Een schatting u_n van θ heet de doeltreffendste²⁾ schatting van θ indien zij zuiver is en bovendien van alle zuivere schattingen de kleinste spreiding bezit.

Voorbeeld: De door (1) gegeven schatting van het gemiddelde van W is, als W een normale waarschijnlijkheidsverdeling is, de doeltreffendste schatting van het gemiddelde.

E. Een schatting u_n van θ heet een asymptotisch meest doeltreffende³⁾, indien zij bruikbaar is en onder alle bruikbare schattingen voor $n \rightarrow \infty$ de kleinste spreiding bezit.

Voorbeeld: de schatting (5) is evenals (4) een asymptotisch meest doeltreffende schatting van de variantie van W , indien W een normale waarschijnlijkheidsverdeling is. Daarentegen voldoet (4) wel, maar (5) niet aan de definitie D. .

1) Engels: "consistent".

2) Engels: "most efficient estimate". Deze naam wordt ook vaak gegeven aan de onder E genoemde schattingen.

3) Engels: "asymptotically most efficient estimate". Vgl. ook voetnoot 2) van deze pagina.

F. Een schatting u_n van θ heet de aannemelijkste schatting¹⁾ van θ , indien de waarschijnlijkheid, of bij continue $\mathbb{V}$ de waarschijnlijkheidsdichtheid van het gevonden resultaat (d.i. van de gevonden waarnemingen $x_1, \dots, x_n$) maximaal wordt bij substitutie van u_n voor θ .

Voorbeeld: Indien onder n worpen met een munt a maal kruis voorkomt, is a/n de aannemelijkste schatting van de kans op kruis.

Opmerking: In de volgende paragraaf worden deze definities gepreciseerd en nog een enkele verdere definitie toegevoegd. Verder worden enkele eigenschappen vermeld.

3. Precisering van de definities.

Zij $F(x/\theta)$ de verdelingsfunctie en $f(x/\theta)$, indien deze bestaat, de waarschijnlijkheidsdichtheid van de waarschijnlijkheidsverdeling W met de onbekende parameter θ .

Zij verder

$$(6) \quad \ell_{\theta} \varphi(x) = \int_0^1 \varphi(x) dF(x|\theta) = \int_{-\infty}^{\infty} \varphi(x) f(x|\theta) dx$$

Dan is:

A. u_n een zuivere schatting van θ , indien geldt:

$$(7) \quad \ell_{\theta} u_n = \theta$$

B. u_n een asymptotisch zuivere schatting van θ , indien geldt;

$$(8) \quad \lim_{n \rightarrow \infty} \ell_{\theta} u_n = \theta$$

C. u_n een bruikbare schatting van θ , indien voor iedere $\varepsilon > 0$ geldt:

$$(9) \quad \lim_{n \rightarrow \infty} P[|u_n - \theta| > \varepsilon] = 0$$

D. u_n de doeltreffendste schatting van θ , indien u_n aan (7) voldoet en

$$(10) \quad \sigma_{\theta}(u_n) \leq \sigma_{\theta}(v_n) \quad 2)$$

is voor iedere v_n , waarvoor (7) geldt.

E. u_n een asymptotisch doeltreffendste schatting van θ , indien u_n aan (9) voldoet, terwijl voor iedere v_n , die aan (9) voldoet geldt:

1) Engels: "maximum likelihood estimate". Dit begrip is afkomstig uit de theorie der aannemelijkste schattingen van R.A.Fisher.

$$2) \quad \sigma_{\theta} \{ \psi(x) \} = \left[\ell_{\theta} \{ \psi(x) - \ell_{\theta} \psi(x) \}^2 \right]^{1/2}$$

$$(11) \quad \lim_{n \rightarrow \infty} \frac{\sigma_{\theta}^2(u_n)}{\sigma_{\theta}^2(\hat{u}_n)} \leq 1.$$

D', E'. Men noemt u_n doeltreffender resp. asymptotisch doeltreffender dan $\hat{u}_n$ als schatting van θ , indien (10) resp. (11) geldt met het kleiner-teken.

Men noemt de grootheid:

$$(12) \quad \lim_{n \rightarrow \infty} \frac{\sigma_{\theta}^2(u_n)}{\sigma_{\theta}^2(\hat{u}_n)}$$

de asymptotische doeltreffendheid¹⁾ van $\hat{u}_n$, indien u_n een asymptotisch meest doeltreffende schatting van θ is.

F. De aannemelijkheid²⁾ van het stelsel waarnemingen $x_1, \dots, x_n$ wordt gedefinieerd door

$$(13) \quad \mathcal{L} = \sum_{i=1}^n \ln f(x_i | \theta)$$

of, indien W een discontinue waarschijnlijkheidsverdeling is, door:

$$(14) \quad \mathcal{L}^* = \sum_{i=1}^n \ln P[x_i = x_i | \theta]$$

$\mathcal{L}$ en $\mathcal{L}^*$ worden in de theorie der aannemelijkste schattingen beschouwd als functies van θ en de aannemelijkste schatting u_n van θ , bij gegeven $x_1, \dots, x_n$, wordt gedefinieerd als die waarde van θ , waarvoor $\mathcal{L}$ resp. $\mathcal{L}^*$ een absoluut maximum bezitten.

Enkele eigenschappen.

Een bruikbare schatting is altijd asymptotisch zuiver.

Een zuivere schatting is bruikbaar, indien geldt:

$$(15) \quad \lim_{n \rightarrow \infty} \sigma_{\theta}^2(u_n) = 0 \quad 3)$$

Onder vrij algemene voorwaarden geldt, dat de aannemelijkste schatting bruikbaar en asymptotisch meest doeltreffend is. Eén van de voorwaarden, waaronder deze stelling bewezen is, is, dat W een continue waarschijnlijkheidsverdeling is.

Litteratuur: D. van Dantzig, Kadercursus Mathematische Statistiek, Mathematisch Centrum, Amsterdam 1946-1950, Hoofdstuk 4.

1) Engels: "efficiency"; ook in het Nederlands laat men de toevoeging "asymptotisch" meestal weg.

2) Engels: "likelihood function" of "likelihood".

3) Deze voorwaarde is voldoende, maar niet noodzakelijk.

MATHEMATISCH CENTRUM
 Amsterdam.
 Statistische Afdeling.
 S47(M18)

Betrouwbaarheidsintervallen (algemeen).¹⁾

Zij x een stochastische grootheid, die een verdelingsfunctie bezit die, op een onbekende parameter θ na, geheel bekend is (θ kan bv. het gemiddelde van x zijn, of de spreiding of iets dergelijks), dan kan men de vraag stellen uit een aantal waarnemingen van x een schatting voor θ af te leiden.

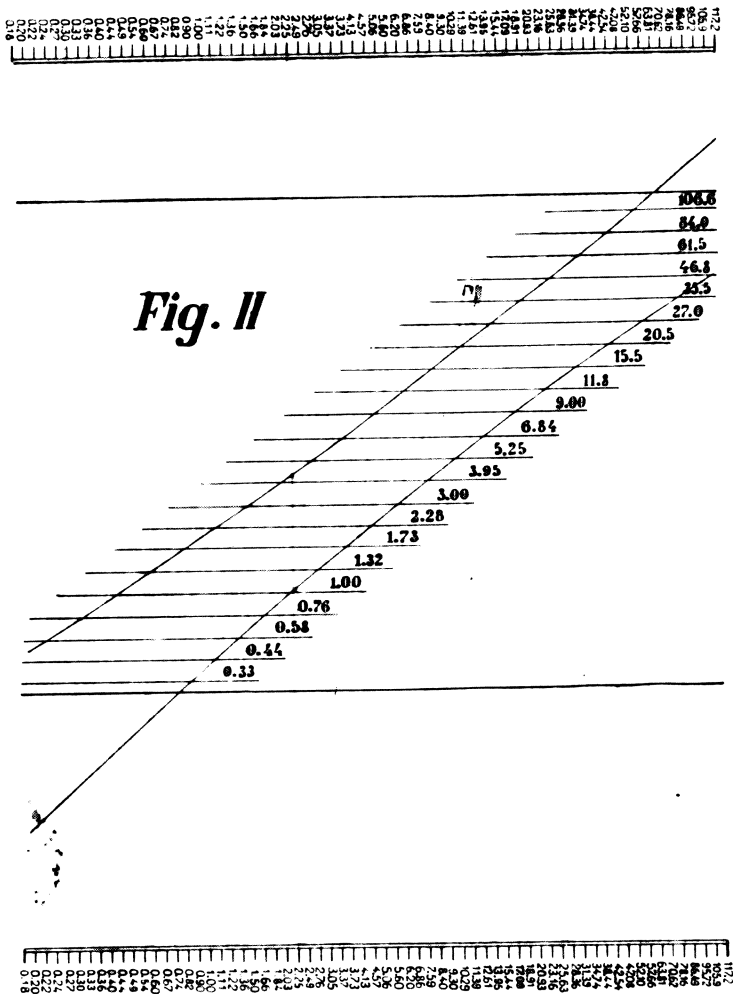
Een betrouwbaarheidsinterval $\mathcal{I}$ voor θ is een interval, waarvan de grenzen afhankelijk zijn van de waarnemingen $x_1, \dots, x_n$ van x , en dat de eigenschap bezit, behoudens een zekere gegeven onbetrouwbaarheid α , de juiste waarde van θ te bevatten. Dit betekent, dat bij een serie bepalingen van betrouwbaarheidsintervallen slechts in ongeveer een fractie α van deze gevallen het interval $\mathcal{I}$ zo zal uitvallen, dat het θ niet bevat. Hierbij is dus θ constant en het interval $\mathcal{I}$ veranderlijk (en wel stochastisch). Hierin ligt het grote verschil met een zgn. voorspellingsinterval, d.i. een gegeven vast interval, waar een stochastisch punt met een zekere waarschijnlijkheid in valt.

Het algemene principe ter bepaling van een betrouwbaarheidsinterval is het volgende: zij $\mathcal{T}$ een toets voor de hypothese $\theta = \theta_0$ (vgl. S47(M6)), dan is $\mathcal{I}$ de verzameling van die waarden θ_0 die bij toepassing van $\mathcal{T}$ op grond van de gevonden waarnemingen $x_1, \dots, x_n$ niet voor verwerping in aanmerking komen. Is $\mathcal{T}$ toegepast met een onbetrouwbaarheidsdrempel α , dan is dit ook de onbetrouwbaarheidsdrempel van het betrouwbaarheidsinterval.

Litteratuur:

- M.G. Kendall, The Advanced Theory of Statistics, London 1946, deel II, p.62-84.
 A.M. Mood, Introduction to the theory of Statistics, London 1950, p.220.
 J. Neyman, First course in probability and statistics, N.Y. 1950.

1) Dit memorandum is slechts bedoeld ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.



Notice.—The two horizontal scales in the diagram are those of λ' , the maximum likelihood estimate of λ as read up from the table of Halvorsen and Ziegler. The vertical scale, given in steps, refers to the ends of confidence intervals. To obtain the confidence interval corresponding to any given value of λ' find this value of λ' in the two horizontal scales and connect the corresponding points by a straight line. The points of intersection of the same with the two parabolae will determine the confidence interval required, e.g. if $\lambda' = 2.25$, then the confidence interval will extend from $\lambda = 1.0$ to about $\lambda = 4.3$.