OVERDRUK UIT:

DEEL XXXV (1) JANUARI 1958

HET VERZEKERINGS-ARCHIEF

WETENSCHAPPELIJK ORGAAN VAN DE
NEDERLANDSE VERENIGING TER BEVORDERING
VAN HET LEVENSVERZEKERINGWEZEN EN VAN DE
VERENIGING VOOR VERZEKERINGS-WETENSCHAP

ACTUARIEEL BIJVOEGSEL

'S-GRAVENHAGE
MARTINUS NIJHOFF
1958

MAXIMUM LIKELIHOOD ESTIMATION OF PARTIALLY OR COMPLETELY ORDERED PROBABILITIES [1]

by

CONSTANCE VAN EEDEN

1. *Introduction*

The problem to be treated in this paper concerns the maximum likelihood estimation of partially or completely ordered probabilities.

A description of this problem will be given in section 2 and in section 3 methods will be described by means of which the estimates may be found.

In this paper no proofs will be given; these may be found in [3], [4], [5] and [6].

2. *The problem*

Consider k independent series of independent trials, each trial resulting in a success or a failure. The i-th series consists of n_i trials with $\boldsymbol{a}_i$ [2] successes and $\boldsymbol{b}_i = n_i - \boldsymbol{a}_i$ failures; π_i is the (unknown) probability of a success for each trial of the i-th series ($i = 1, \ldots, k$) [3].

The probabilities $\pi_1, \ldots, \pi_k$ are partially or completely ordered, i.e. they satisfy the inequalities

$$\alpha_{i,j} (\pi_i - \pi_j) \leq 0 \qquad (i, j = 1, \ldots, k), \qquad (2;1)$$

---

[1]) Report S 220 (V15) of the Statistical Department of the Mathematical Centre, Amsterdam. Lecture delivered at the meeting of the „Actuarieel Genootschap" of November 27, 1956.

[2]) Random variables will be distinguished from numbers (e.g. from the values they take in an experiment) by printing their symbols in bold type.

[3]) Unless explicitely stated otherwise i, j and h run through the values 1, ..., k.

where $\alpha_{i,j}$ are given numbers satisfying

$$\begin{aligned} &1.\ \alpha_{i,j} = -\alpha_{j,i}, \\ &2.\ \alpha_{i,j} = 0, +1 \text{ or } -1, \\ &3.\ \alpha_{i,j} = 1 \text{ if } \alpha_{i,h} = \alpha_{h,j} = 1 \text{ for any } h. \end{aligned} \tag{2;2}$$

It will be supposed that the probabilities $\pi_1, \ldots, \pi_k$ are numbered in such a way that $\alpha_{i,j} \geq 0$ for each pair of values (i, j) with $i < j$.

If $\alpha_{i,j} = 1$ for each pair of values (i, j) with $i < j$ then (2;1) reduces to

$$\pi_i - \pi_j \leq 0 \text{ for each pair of values } (i, j) \text{ with } i < j. \tag{2;3}$$

Thus in this case $\pi_1, \ldots, \pi_k$ satisfy

$$\pi_1 \leq \ldots \leq \pi_k, \tag{2;4}$$

i.e. they are completely ordered.

Examples of ordered probabilities may be found in [1] (p. 641—642) and in [2] (p. 610).

The random variable $\boldsymbol{a}_i$ possesses a binomial probability distribution, i.e.

$$\mathrm{P}[\boldsymbol{a}_i = a_i] = \binom{n_i}{a_i} \pi_i^{a_i} (1 - \pi_i)^{b_i} \quad (i = 1, \ldots, k). \tag{2;5}$$

From the fact that $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ are distributed independently then follows

$$\mathrm{P}[\boldsymbol{a}_1 = a_1, \ldots, \boldsymbol{a}_k = a_k] = \prod_{i=1}^{k} \binom{n_i}{a_i} \pi_i^{a_i} (1 - \pi_i)^{b_i}, \tag{2;6}$$

i.e. the maximum likelihood estimates of the ordered probabilities $\pi_1, \ldots, \pi_k$ are the values of $y_1, \ldots, y_k$ which maximize

$$\prod_{i=1}^{k} \binom{n_i}{a_i} y_i^{a_i} (1 - y_i)^{b_i} \tag{2;7}$$

in the domain

$$D: \begin{array}{c} \alpha_{i,j} (y_i - y_j) \leq 0 \\ 0 \leq y_i \leq 1 \end{array} \quad (i, j = 1, \ldots, k). \tag{2;8}$$

In order to simplify the calculations we maximize the function

$$L = L(y_1, \ldots, y_k) \stackrel{\text{def}}{=} \sum_{i=1}^{k} \{a_i \ln y_i + b_i \ln (1 - y_i)\} \tag{2;9}$$

instead of the function (2;7).

In [3] and [4] it has been proved that the function $L(y_1, \ldots, y_k)$ possesses a unique maximum in the domain D. The values of $y_1, \ldots, y_k$ which maximize L in D, i.e. the maximum likelihood estimates of $\pi_1, \ldots, \pi_k$, will be denoted by $p_1, \ldots, p_k$.

Now let

$$L_i(y) \stackrel{\text{def}}{=} a_i \ln y + b_i \ln (1 - y) \quad (0 \leq y \leq 1 \,; i = 1, \ldots, k), \quad (2;10)$$

then

$$\left.\begin{aligned} \frac{dL_i(y)}{dy} &= \frac{a_i}{y} - \frac{b_i}{1-y} = \frac{a_i - n_i y}{y(1-y)} \\ \frac{d^2L_i(y)}{dy^2} &= \frac{-n_i y^2 + 2a_i y - a_i}{y^2(1-y)^2} < 0 \end{aligned}\right\} \quad (i = 1, \ldots, k). \quad (2;11)$$

Thus the function $L_i(y)$ possesses a unique maximum in the interval $(0,1)$. This maximum is attained for

$$y = f_i \stackrel{\text{def}}{=} \frac{a_i}{n_i} \quad (2;12)$$

and $L_i(y)$ is a monotone increasing function of y for $y < f_i$ and a monotone decreasing function of y for $y > f_i$, i.e. if (y', y'') is a pair of values satisfying

$$0 \leq y' < y'' < f_i \quad (2;13)$$

or

$$f_i < y'' < y' \leq 1 \quad (2;14)$$

then

$$L_i(y') < L_i(y'') < L_i(f_i). \quad (2;15)$$

From (2;12) it follows that the function

$$L(y_1, \ldots, y_k) = \sum_{i=1}^{k} L_i(y_i)$$

attains its maximum in the domain

$$G : 0 \leq y_i \leq 1 \;\; (i = 1, \ldots, k) \quad (2;16)$$

for $y_i = f_i \; (i = 1, \ldots, k)$.

Thus if the point $(f_1, \ldots, f_k)$ lies in D, i.e. if $f_1, \ldots, f_k$ satisfy

$$\alpha_{i,j}(f_i - f_j) \leq 0 \qquad (i, j = 1, \ldots, k) \quad (2;17)$$

then, D being a (convex) subdomain of G, the maximum of L in D coincides with the maximum of L in G, i.e.

$$p_i = f_i \; (i = 1, \ldots, k) \;\text{ if }\; (f_1, \ldots, f_k) \in D. \qquad (2;18)$$

3. *The maximum likelihood estimates*

3.1. *A complete ordering*

In this section we consider the case that $\pi_1, \ldots, \pi_k$ are completely ordered, i.e. the case that $\pi_1, \ldots, \pi_k$ satisfy the inequalities

$$\pi_1 \leq \ldots \leq \pi_k. \qquad (3.1;1)$$

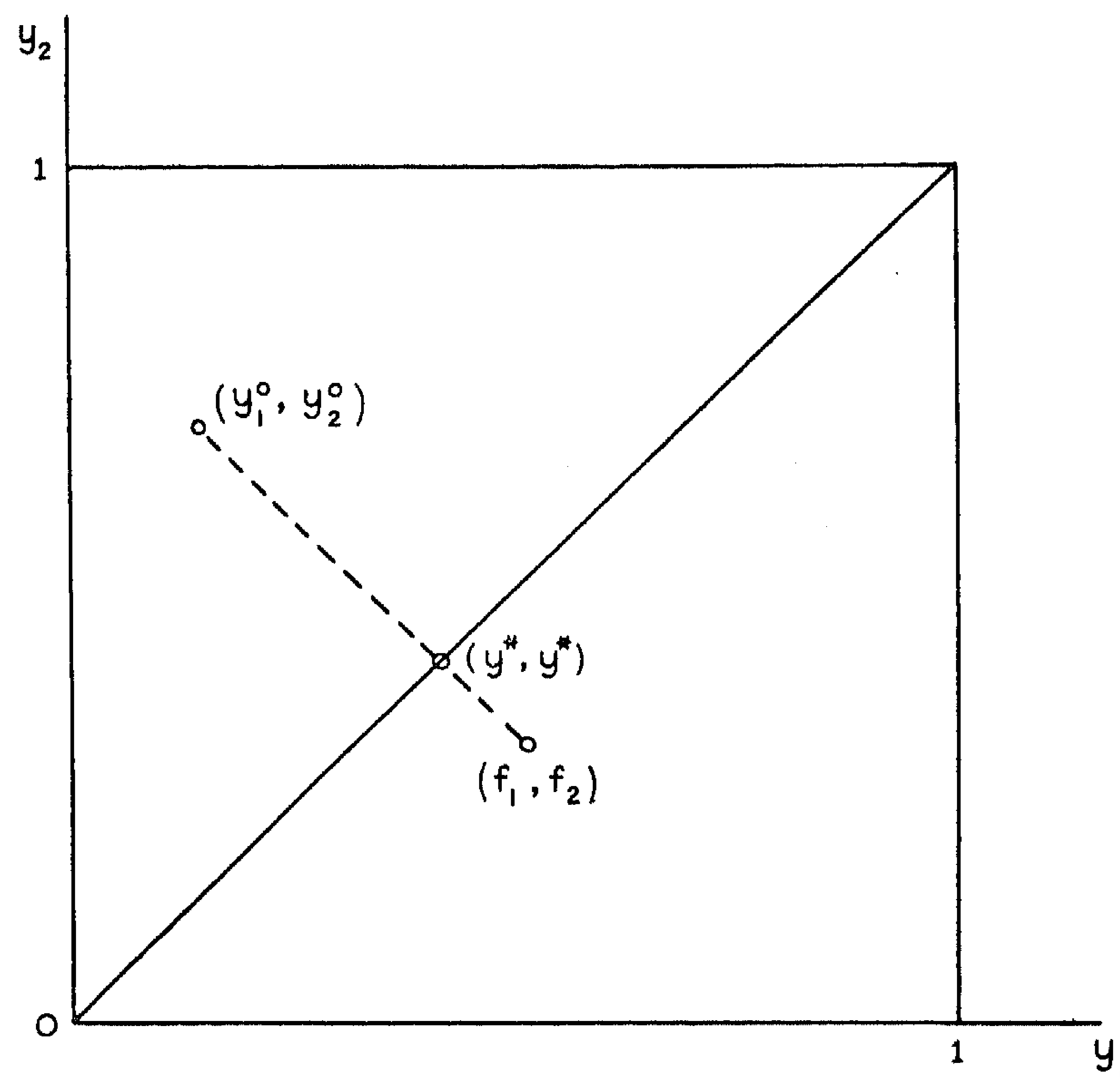

Fig. 1. A complete ordering with $k = 2$ and $f_1 > f_2$.

We first consider the case that $k = 2$; then D is the domain

$$\begin{aligned} &y_1 \leq y_2, \\ &0 \leq y_i \leq 1 \qquad (i = 1,2). \end{aligned} \qquad (3.1;2)$$

From (2;18) it follows that

$$p_1 = f_1, \ p_2 = f_2 \text{ if } f_1 \leq f_2. \tag{3.1;3}$$

Now let $f_1 > f_2$ and let (y_1^0, y_2^0) be a point with $y_1^0 < y_2^0$. If (cf. fig. 1) this point (y_1^0, y_2^0) is connected with the point (f_1, f_2) by means of a straight line, then this line intersects the line $y_1 = y_2$ in a point (say) (y^*, y^*). From (2;15) then follows

$$L\,(y^*, y^*) > L\,(y_1^0, y_2^0), \tag{3.1;4}$$

i.e. for each point (y_1^0, y_2^0) with $y_1^0 < y_2^0$ a point (y^*, y^*) exists with a larger value of L. Consequently, if $f_1 > f_2$, then L attains its maximum in D for $y_1 = y_2$. Substituting this into L we obtain

$$L(y_1, y_1) = (a_1 + a_2) \ln y_1 + (b_1 + b_2) \ln (1 - y_1) \tag{3.1;5}$$

and this function attains its maximum in the interval (0,1) for

$$y_1 = \frac{a_1 + a_2}{n_1 + n_2}. \tag{3.1;6}$$

Thus for a complete ordering with $k = 2$ we have

$$\begin{aligned} &p_1 = f_1, \ p_2 = f_2 \ \text{ if } f_1 \leq f_2, \\ &p_1 = p_2 = \frac{a_1 + a_2}{n_1 + n_2} = \frac{n_1 f_1 + n_2 f_2}{n_1 + n_2} \quad \text{if} \quad f_1 > f_2. \end{aligned} \tag{3.1;7}$$

In an analogous way the following theorem may be proved for a complete ordering with $k \geq 2$ (cf. theorem IV in [3]).

Theorem I. If $\pi_1, \ldots, \pi_k$ *are completely ordered and if*

$$f_{i_0} > f_{i_0+1} \tag{3.1;8}$$

then L *attains its maximum in* D *for* $y_{i_0} = y_{i_0+1}$.

Substituting $y_{i_0} = y_{i_0+1}$ into L the two terms with $i = i_0$ and $i = i_0 + 1$ reduce to

$$(a_{i_0} + a_{i_0+1}) \ln y_{i_0} + (b_{i_0} + b_{i_0+1}) \ln (1 - y_{i_0}), \tag{3.1;9}$$

i.e. L reduces to a sum of $k - 1$ terms of the same form as L and D reduces to

$$\begin{aligned} y_1 \leq \ldots \leq y_{i_0} \leq y_{i_0+2} \leq \ldots \leq y_k, \\ 0 \leq y_i \leq 1 \quad (i \neq i_0 + 1). \end{aligned} \tag{3.1;10}$$

Thus the problem may be solved by repeatedly applying theorem I and (2 ;18).

Further it has been proved in [2] and [5] that

$$p_i = \max_{1 \leq r \leq i} \min_{i \leq s \leq k} \frac{a_r + \ldots + a_s}{n_r + \ldots + n_s}, \qquad (3.1\,;11)$$

but this formula is not recommended for calculation.

Example 1

Suppose $k = 4$ and

i	1	2	3	4
a_i	4	3	10	8
n_i	10	5	30	15
f_i	0,4	0,6	0,33	0,53.

(3.1 ;12)

Then $f_2 > f_3$ and from theorem I then follows that L attains its maximum in D for $y_2 = y_3$. Substituting this into L the problem reduces to the case of $k - 1 = 3$ series of trials with (cf. (3.1 ;9))

i	1	2 (+3)	4
a'_i	4	13	8
n'_i	10	35	15
f'_i	0,4	0,37	0,53.

(3.1 ;13)

From $f'_1 > f'_2$ and theorem I then follows that L attains its maximum in D for $y_1 = y_2$ and this reduces the problem to the case of $k - 2 = 2$ series of trials with

i	1 (+2 +3)	4
a''_i	17	8
n''_i	45	15
f''_i	0,38	0,53.

(3.1 ;14)

From (3.1 ;14) and (2 ;18) then follows

$$(3.1\,;15) \qquad p_1 = p_2 = p_3 = 0{,}38\,;\; p_4 = 0{,}53.$$

3.2. *A partial or complete ordering*

The restrictions $\pi_i \leq \pi_j$ satisfying

$$\alpha_{i,h} \cdot \alpha_{h,j} = 0 \text{ for each } h \text{ between } i \text{ and } j \qquad (3.2\,;1)$$

will be called the essential restrictions defining D and will be denoted by $R_1, \ldots, R_s$. Then each R_λ corresponds to one pair of values (i, j) ; this pair will be denoted by (i_λ, j_λ) $(\lambda = 1, \ldots, s)$. If $\pi_1, \ldots, \pi_k$ are completely ordered then $s = k - 1$ and the essential restrictions are $\pi_i \leq \pi_{i+1}$ $(i = 1, \ldots, k - 1)$.

In [3] the following theorem has been proved.

Theorem II: *If* $p_1', \ldots, p_k'$ *are the values of* $y_1, \ldots, y_k$ *which maximize L under the restrictions* $R_1, \ldots, R_{s-1}$ *then*

$$\begin{aligned} &1.\ p_i = p_i' \ (i = 1, \ldots, k) \text{ if } p_{i_s}' \leq p_{j_s}', \\ &2.\ p_{i_s} = p_{j_s} \text{ if } p_{i_s}' > p_{j_s}'. \end{aligned} \qquad (3.2;2)$$

Thus if $p_1', \ldots, p_k'$ are known then the problem is solved if $p_{i_s}' \leq p_{j_s}'$ and if $p_{i_s}' > p_{j_s}'$ (by substituting $y_{i_s} = y_{j_s}$) the problem is reduced to the case of $k - 1$ series of trials with $s - 1$ or less essential restrictions, i.e. by means of theorem II the problem is reduced to the case of k and $k - 1$ series of trials under $s - 1$ or less essential restrictions. The solution for the case that $s = 0$ being known ($p_i = f_i$ $(i = 1, \ldots, k)$) the problem may be solved by repeatedly applying theorem II.

In many cases however this procedure may be simplified by applying one of the following theorems (cf. the theorems III, IV and V in [3]).

Theorem III: *If* (i, j) *is a pair of values with*

$$\begin{aligned} &1.\ \alpha_{i,j}(f_i - f_j) > 0, \\ &2.\ \alpha_{i,h} = \alpha_{j,h} \quad \text{for each } h \text{ with } h \neq i,\ h \neq j, \end{aligned} \qquad (3.2;3)$$

then L attains its maximum in D for $y_i = y_j$.

By means of this theorem the problem may be reduced to the case of $k - 1$ series of trials with $s - 1$ or less restrictions by substituting $y_i = y_j$.

Remark 1 :

If $\pi_1, \ldots, \pi_k$ are completely ordered then for each pair of values (i, j) with $j = i + 1$ we have

$$\begin{aligned} \alpha_{h,i} = \alpha_{h,j} = 1 \quad &\text{for each } h < i, \\ \alpha_{i,h} = \alpha_{j,h} = 1 \quad &\text{for each } h > j. \end{aligned} \qquad (3.2;4)$$

Thus theorem I is a special case of theorem III.

Theorem IV: If (i, j) is a pair of values with

$$\begin{aligned} &1.\ \alpha_{i,j} = 0, \\ &2.\ f_i \leq f_j, \\ &3.\ \alpha_{i,h} \geq \alpha_{j,h} \text{ for each } h \text{ with } h \neq i,\ h \neq j, \end{aligned} \tag{3.2;5}$$

then L attains its maximum in D for $y_i \leq y_j$.

Thus if (i, j) is a pair of values satisfying (3.2;5) then the restriction $y_i \leq y_j$ may be added.

Theorem V: If M is a subset of the numbers 1, ..., k with

$$\alpha_{i,j} = 0 \text{ for each pair of values } (i, j) \text{ with } i \in M,\ j \notin M, \tag{3.2;6}$$

then the maximum likelihood estimates may be found by seperately maximizing $\sum_{i \in M} L_i(y_i)$ and $\sum_{i \notin M} L_i(y_i)$.

The theorems will now be illustrated by means of the following examples.

Example 2:

Suppose $k = 3$ and

i	1	2	3
a_i	8	8	26
n_i	15	20	40
f_i	0,53	0,4	0,65

(3.2;7)

and

$$\alpha_{1,2} = \alpha_{1,3} = 1,\ \alpha_{2,3} = 0. \tag{3.2;8}$$

Then the pair $i = 1, j = 2$ satisfies (3.2;3.1) but

$$\alpha_{1,3} \neq \alpha_{2,3}, \tag{3.2;9}$$

i.e. the pair (1,2) does not satisfy (3.2;3.2). Thus theorem III cannot be applied.

Further the pair $i = 2, j = 3$ satisfies (3.2;5); we have

$$\begin{aligned} &1.\ \alpha_{2,3} = 0, \\ &2.\ f_2 \leq f_3, \\ &3.\ \alpha_{2,1} = \alpha_{3,1}. \end{aligned} \tag{3.2;10}$$

From theorem IV then follows that L attains its maximum in D for $y_2 \leq y_3$, which reduces the problem to the case of a complete ordering. Applying theorem I to this complete ordering we find that the maximum of L is attained for $y_1 = y_2$. This reduces the problem to the case of $k - 1 = 2$ series of trials with

i	1 (+2)	3
a_i'	16	26
n_i'	35	40
f_i'	0,46	0,65

(3.2 ;11)

and from (2 ;18) then follows

$$p_1 = p_2 = 0{,}46\,;\ p_3 = 0{,}65. \qquad (3.2\,;12)$$

Example 3 :

Suppose $k = 4$ and

i	1	2	3	4
a_i	7	18	13	10
n_i	10	30	20	25
f_i	0,7	0,6	0,65	0,4

(3.2 ;13)

and

$$\begin{aligned} &1.\ \alpha_{1,2} = \alpha_{1,4} = \alpha_{3,4} = 1, \\ &2.\ \alpha_{1,3} = \alpha_{2,3} = \alpha_{2,4} = 0. \end{aligned} \qquad (3.2\,;14)$$

Then the pairs (1,2), (1,4) and (3,4) satisfy (3.2 ; 3.1), but they do not satisfy (3.2 ;3.2). E.g. for the pair (1,2) we have

$$\alpha_{1,4} \neq \alpha_{2,4}. \qquad (3.2\,;15)$$

Thus theorem III cannot be applied.

Further the pairs (3,1), (2,3) and (4,2) satisfy (3.2 ; 5.1) and (3.2 ; 5.2), but they do not satisfy (3.2 ; 5.3). E.g. for the pair (2,3) we have

$$\alpha_{2,4} < \alpha_{3,4}, \qquad (3.2\,;16)$$

i.e. theorem IV cannot be applied and therefore we use theorem II. Omitting the restriction $\pi_1 \leq \pi_4$, i.e. taking

$$i_s = 1,\ j_s = 4 \qquad (3.2\,;17)$$

the estimates $p_1', \ldots, p_4'$ are the values of $y_1, \ldots, y_4$ which maximize L in the domain

$$D': \begin{array}{l} y_1 \leq y_2,\ y_3 \leq y_4, \\ 0 \leq y_i \leq 1 \qquad (i = 1, \ldots, 4). \end{array} \tag{3.2;18}$$

From theorem V it then follows that $p_1', \ldots, p_4'$ may be found by separately maximizing $L_1(y_1) + L_2(y_2)$ and $L_3(y_3) + L_4(y_4)$ and from theorem I then follows

$$\begin{aligned} p_1' = p_2' &= \frac{7 + 18}{10 + 30} = 0{,}63, \\ p_3' = p_4' &= \frac{13 + 10}{20 + 25} = 0{,}51. \end{aligned} \tag{3.2;19}$$

Thus (cf. (3.2;17)) $p_{i_s}' > p_{j_s}'$ and from theorem II then follows that L attains its maximum in D for $y_{i_s} = y_{j_s}$, i.e. for $y_1 = y_4$. This reduces the problem to the case of $k - 1 = 3$ series of trials with

i	3	1 (+4)	2	(3.2;20)
a_i'	13	17	18	
n_i'	20	35	30	
f_i'	0,65	0,49	0,6	

and

$$\alpha_{3,1}' = \alpha_{1,2}' = 1. \tag{3.2;21}$$

This complete ordering may be solved by means of theorem I and we obtain

$$p_1 = p_3 = p_4 = 0{,}55\,;\ p_2 = 0{,}6. \tag{3.2;22}$$

Now let

$$\begin{array}{l} S_i \overset{\text{def}}{=} i \cup Ens\,\{j \mid \alpha_{j,i} = 1\}\,^{1)} \\ T_i \overset{\text{def}}{=} i \cup Ens\,\{j \mid \alpha_{i,j} = 1\} \end{array} \qquad (i = 1, \ldots, k) \tag{3.2;23}$$

and let, for a subset M of the numbers $1, \ldots, k$,

$$\begin{aligned} S &\overset{\text{def}}{=} \bigcup_{i \in M} S_i, \\ T &\overset{\text{def}}{=} \bigcup_{i \in M} T_i, \end{aligned} \tag{3.2;24}$$

---

[1]) $Ens\,\{j \mid \alpha_{j,i} = 1\}$ is the set of all values j for which $\alpha_{j,i} = 1$.

then it has been proved in [2] and [5] that

$$p_i = \max_{\substack{T \\ i \in T \cap S}} \min_{S} \frac{\sum_{j \in T \cap S} a_j}{\sum_{j \in T \cap S} n_j} \quad (i = 1, \ldots, k). \qquad (3.2;25)$$

If $\pi_1, \ldots, \pi_k$ are completely ordered then S_i consists of the numbers $1, \ldots, i$ and T_i of the numbers $i, \ldots, k$. Formula (3.2;25) then reduces to (3.1;11).

Further it has been proved in [6] that the maximum likelihood estimates of $\pi_1, \ldots, \pi_k$ in D are identical with the least squares estimates in D, i.e. with the values of $y_1, \ldots, y_k$ which maximize

$$Q(y_1, \ldots, y_k) \stackrel{\text{def}}{=} \sum_{i=1}^{k} n_i (y_i - f_i)^2 \qquad (3.2;26)$$

in D.

Remarks:

2. The problem treated in this paper may be generalized to the case of partially or completely ordered parameters of other probability distributions (cf. [2] and [4]), e.g. parameters of Poisson distributions and means of normal distributions. Further the problem may be generalized by introducing inequalities of the form $c_i \leq \pi_i \leq d_i$, where (c_i, d_i) is a given closed subinterval of the interval (0,1) (cf. [3] and [4]).

3. As stated in section 2 the procedure described in this paper may be applied if the probabilities $\pi_1, \ldots, \pi_k$ are known to satisfy the inequalities (2.1). However the method may also be applied if we want to obtain estimates $p_1, \ldots, p_k$ for the probabilities $\pi_1, \ldots, \pi_k$ satisfying the inequalities $\alpha_{i,j}\ (p_i - p_j) \leq 0$ $(i, j = 1, \ldots, k)$.

REFERENCES

[1] Ayer, Miriam, H. D. Brunk, G. M. Ewing, W. T. Reid and Edward Silverman, An empirical distribution function for sampling with incomplete information, Ann. Math. Stat. **26** (1955), 641—647.

[2] Brunk, H. D., Maximum likelihood estimates of monotone parameters Ann. Math. Stat. **26** (1955), 607—616.

[3] Van Eeden, Constance, Maximum likelihood estimation of ordered probabilities, Proc. Kon. Ned. Akad. v. Wet. A 59 (1956), Indagationes Mathematicae, 18 (1956), 444—455.

[4] Van Eeden, Constance, Maximum likelihood estimation of partially or completely ordered parameters, Proc. Kon. Ned. Akad. v. Wet. A **60** (1957), Indagationes Mathematicae **19** (1957), 128—136 and 201—211.
[5] Van Eeden, Constance, Note on two methods for estimating ordered parameters of probability distributions, to be published in Proc. Kon. Ned. Akad. v. Wet. and in Indagationes Mathematicae.
[6] Van Eeden, Constance, A least squares inequality for maximum likelihood estimates of ordered parameters, to be published in Proc. Kon. Ned. Akad. v. Wet. and in Indagationes Mathematicae.

Amsterdam, July 1957.