MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
STATISTISCHE AFDELING

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

Report S 207 (VP 9)

Maximum likelihood estimation of partially or completely ordered parameters

by

Constance van Eeden

September 1956

1. Introduction

The problem treated in this report concerns the maximum likelihood estimation of partially or completely ordered parameters of probability distributions. A special case of this problem, the maximum likelihood estimation of ordered probabilities, has been treated in [2] .

The problem will be formulated in section 2; in section 4 and 5 methods will be given by means of which the estimates may be found. For the proofs of the theorems we need some lemma's which will be proved in section 3 and in section 6 some examples will be given.

2. The problem

Consider k independent random variables $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_k$ 1) and n_i independent observations $x_{i,1}, x_{i,2}, \ldots, x_{i,n_i}$ of $\underline{x}_i$ $(i = 1, 2, \ldots, k)$. The distribution of $\underline{x}_i$ contains one unknown parameter θ_i $(i = 1, 2, \ldots, k)$ and its distribution function is

$$(2.1) \qquad F_i(x_i \mid \theta_i) \overset{\text{def}}{=} P[\underline{x}_i \leqq x_i \mid \theta_i] \qquad (i = 1, 2, \ldots k).$$

Two types of restrictions are imposed on the parameters $\theta_1, \theta_2, \ldots, \theta_k$. First let ϑ_i be a closed interval such that $F_i(x_i \mid y_i)$ is a distributionfunction for each value of $y_i \in \vartheta_i$ $(i = 1, 2, \ldots, k)$. By meansof the choice of ϑ_i restrictions of the type $c_i \leqq \theta_i \leqq d_i$ may be imposed. The second type of restrictions consists of a partial or complete ordering of the parameters, which may be described as follows. Let $\alpha_{i,j}$ $(i, j = 1, 2, \ldots, k)$ be numbers satisfying the conditions

$$(2.2) \quad \begin{cases} 1. \ \alpha_{i,j} = -\alpha_{j,i}, \\ 2. \ \alpha_{i,j} = 0 \text{ if the intersection } \vartheta_i \cap \vartheta_j \text{ contains at most} \\ \quad \text{one point,} \\ 3. \ \alpha_{i,j} = 0, +1 \text{ or } -1 \text{ in all other cases} \end{cases}$$

and

$$(2.3) \qquad \alpha_{i,j} = 1 \text{ if } \qquad \alpha_{i,h} = \alpha_{h,j} = 1 \quad \text{for any } h .$$

The restrictions imposed on $\theta_1, \theta_2, \ldots, \theta_k$ are then

$$(2.4) \qquad \begin{cases} 1. \ \alpha_{i,j}(\theta_i - \theta_j) \leqq 0 \\ 2. \ \theta_i \in \vartheta_i \end{cases} \qquad (i, j = 1, 2, \ldots, k).$$

---

1) Random variables will be distinguished from numbers (e.g. from the values they take in an experiment) by underlining their symbols.

and it will be supposed that the parameters $\theta_1, \theta_2, \ldots, \theta_k$ are numbered in such a way that

(2.5) $\qquad \alpha_{i,j} \geqq 0$ for each pair of values (i,j).

No other restrictions on $\theta_1, \theta_2, \ldots, \theta_k$ are admitted, such that all points $y_1, y_2, \ldots, y_k$ of the Cartesian product

(2.6) $$G \overset{\text{def}}{=} \prod_{i=1}^{k} \vartheta_i ,$$

satisfying

(2.7) $$\alpha_{i,j}(y_i - y_j) \leqq 0 \qquad (i,j = 1,2,\ldots,k)$$

belong to the parameterspace, which thus is a convex subdomain of G. This subdomain will be denoted by D.
Let

(2.8) $$\begin{cases} 1. & \alpha_{i,j} = 0 \text{ for } \varkappa_0 \text{ pairs of values } (i,j) \text{ with } i < j, \\ 2. & \alpha_{i,j} = 1 \text{ for } \varkappa_1 \text{ pairs of values } (i,j) \text{ with } i < j, \end{cases}$$

then

(2.9) $$\varkappa_0 + \varkappa_1 = \binom{k}{2}.$$

Let further $f_i(x_i \mid \theta_i)$ denote the density function of $\underline{x}_i$ if $\underline{x}_i$ possesses a continuous probability distribution and $P[\underline{x}_i = x_i \mid \theta_i]$ if $\underline{x}_i$ possesses a discrete probability distribution and let

(2.10) $$\begin{cases} 1. & L_i = L_i(y_i) \overset{\text{def}}{=} \sum_{\gamma=1}^{n_i} \lg f_i(x_{i,\gamma} \mid y_i) \qquad (i = 1,2,\ldots,k), \\ 2. & L = L(y_1, y_2, \ldots, y_k) \overset{\text{def}}{=} \sum_{i=1}^{k} L_i(y_i). \end{cases}$$

Then the maximum likelihood estimates of $\theta_1, \theta_2, \ldots, \theta_k$ are the values of $y_1, y_2, \ldots, y_k$ which maximize L in the domain D.
Unless explicitely stated otherwise L will only be considered in this domain D; the maximum likelihood estimates will throughout this paper be denoted by $t_1, t_2, \ldots, t_k$.
Further the restrictions $\theta_i \leqq \theta_j$ (i.e. $\alpha_{i,j} = 1$) satisfying

(2.11) $\qquad \alpha_{i,h} \cdot \alpha_{h,j} = 0$ for each h between i and j

will be denoted by $R_1, R_2, \ldots, R_s$. Each R_λ thus corresponds with one pair (i,j); this pair will be denoted by (i_λ, j_λ).
Because of the transitivity relations (2.3) the system $R_1, R_2, \ldots, R_s$ is equivalent to (2.4.1) and uniquely determined by (2.4.1).
The restrictions $R_1, R_2, \ldots, R_s$ will be called the essential restrictions.

Remark 1: H.D. BRUNK [1] described a method by means of which the estimates of $\theta_1, \theta_2, \ldots, \theta_k$ may be found if the distribution of $\underline{x}_i$ belongs to the "exponential family" $(i = 1, 2, \ldots, k)$ and if moreover $\mathcal{Y}_i$ is the set of all values of y_i for which $F_i(x_i \mid y_i)$ is a distribution function $(i = 1, 2, \ldots k)$. His method however leads to much more complicated computations than ours.

3. Lemma's

Definition: A function $\varphi(y)$ of a variable y will be called strictly unimodal in an interval $\mathcal{J}$ if there exists a value $y^* \in \mathcal{J}$ such that

$$(3.1) \qquad \varphi(y) < \varphi(z) < \varphi(y^*)$$

for each pair of values $(y, z) \in \mathcal{J}$ with

$$(3.2) \qquad y < z < y^*$$

and for each pair of values $(y, z) \in \mathcal{J}$ with

$$(3.3) \qquad y^* < z < y.$$

It follows at once from this definition that a strictly unimodal function $\varphi(y)$ is bounded in every closed subdomain of $\mathcal{J}$ not containing y^*.

Now let $\varphi_\kappa(y_\kappa)$ be a strictly unimodal function of y_κ in the interval $\mathcal{J}_\kappa$ $(\kappa = 1, 2, \ldots, k)$ and let further

$$(3.4) \qquad \Phi(y_1, y_2, \ldots, y_k) \stackrel{\text{def}}{=} \sum_{\kappa=1}^{k} \varphi_\kappa(y_\kappa),$$

then

Lemma I: $\Phi(y_1, y_2, \ldots, y_k)$ possesses a unique maximum in

$$(3.5) \qquad \Gamma \stackrel{\text{def}}{=} \prod_{\kappa=1}^{k} \mathcal{J}_\kappa.$$

Proof: Let $\varphi_\kappa(y_\kappa)$ attain its maximum in $\mathcal{J}_\kappa$ for $y_\kappa = y_\kappa^*$ $(\kappa = 1, 2, \ldots, k)$. Then it follows from the fact that $\Phi(y_1, y_2, \ldots, y_k)$ is the sum of the k functions $\varphi_\kappa(y_\kappa)$ and that Γ is the Cartesian product of the k intervals $\mathcal{J}_\kappa$, that $\Phi(y_1, y_2, \ldots, y_k)$ possesses a unique maximum in Γ and attain this maximum for $y_\kappa = y_\kappa^*$ $(\kappa = 1, 2, \ldots, k)$.

We now define a function V as follows.

Let $y_1^\circ, y_2^\circ, \ldots, y_k^\circ$ be a given point in Γ with $y_\kappa^\circ \neq y_\kappa^*$ for at least one value of κ and let

(3.6) $$\begin{cases} y_\kappa(\beta) \stackrel{\text{def}}{=} (1-\beta)\, y_\kappa^\circ + \beta\, y_\kappa^* \quad (\kappa = 1, 2, \ldots, k), \\ 0 \leqq \beta \leqq 1. \end{cases}$$

Then $\{ y_1(\beta), y_2(\beta), \ldots, y_k(\beta) \}$ is a point in Γ and V is defined by

(3.7) $$V(\beta) \stackrel{\text{def}}{=} \Phi\{ y_1(\beta), y_2(\beta), \ldots, y_k(\beta)\}.$$

Lemma II: $V(\beta)$ is a monotone increasing function of β in the interval $0 \leqq \beta \leqq 1$.

Proof: Consider a value of κ with

(3.8) $$y_\kappa^\circ = y_\kappa^*$$

then

(3.9) $$y_\kappa(\beta) = y_\kappa^* \quad \text{for each } \beta \text{ with } 0 \leqq \beta \leqq 1.$$

Thus in this case we have

(3.10) $$\varphi_\kappa(y_\kappa^\circ) = \varphi_\kappa\{y_\kappa(\beta)\} = \varphi_\kappa(y_\kappa^*) \quad \text{for each } \beta \text{ with } 0 \leqq \beta \leqq 1.$$

Now consider a value of κ with

(3.11) $$y_\kappa^\circ \neq y_\kappa^*,$$

then it follows from the fact that $\varphi_\kappa(y_\kappa)$ is, in the interval $\mathcal{J}_\kappa$, a strictly unimodal function of y_κ and attain its maximum in $\mathcal{J}_\kappa$ for $y_\kappa = y_\kappa^*$ that

(3.12) $$\varphi_\kappa(y_\kappa^\circ) < \varphi_\kappa\{y_\kappa(\beta_1)\} < \varphi_\kappa\{y_\kappa(\beta_2)\} < \varphi_\kappa(y_\kappa^*)$$

for each pair of values (β_1, β_2) with $0 < \beta_1 < \beta_2 < 1$.

From (3.4) and the fact that there exists at least one value of κ with (3.11) it follows then that

(3.13) $$V(0) < V(\beta_1) < V(\beta_2) < V(1)$$

for each pair of values (β_1, β_2) with $0 < \beta_1 < \beta_2 < 1$.

Lemma III: If C is a closed convex subdomain of Γ, not containing the point $(y_1^*, y_2^*, \ldots, y_k^*)$, then $\Phi(y_1, y_2, \ldots, y_k)$ attains its maximum in C only in one or more points on its border.

Proof: Consider any inner point $y_1^\circ, y_2^\circ, \ldots, y_k^\circ$ of C and let $y_\kappa(\beta)$ be defined by (3.6) $(\kappa = 1, 2, \ldots, k)$. Then, C being a closed convex domain not containing the point $(y_1^*, y_2^*, \ldots, y_k^*)$ there exists a value of β in the interval $0 < \beta < 1$, say β_0, such that $\{y_1(\beta_0), y_2(\beta_0), \ldots, y_k(\beta_0)\}$ is a border point of C. Further it follows from Lemma II that

$$(3.14) \qquad \Phi\{y_1(\beta_0), y_2(\beta_0), \ldots, y_k(\beta_0)\} > \Phi(y_1^\circ, y_2^\circ, \ldots, y_k^\circ).$$

Thus for each inner point $(y_1^\circ, y_2^\circ, \ldots, y_k^\circ)$ of C there exists a border point $(y_1, y_2, \ldots, y_k)$ of C with a larger value of Φ. Moreover Φ is bounded in C, because the point $(y_1^*, y_2^*, \ldots, y_k^*)$ is not contained in C. Thus Φ has a maximum in C, which can evidently only be attained in border points.

4. The maximum likelihood estimates of $\theta_1, \theta_2, \ldots, \theta_k$

Let M be a subset of the numbers $1, 2, \ldots, k$; let further

$$(4.1) \qquad \vartheta_M \overset{\text{def}}{=} \bigcap_{i \in M} \vartheta_i$$

and if $\vartheta_M \neq 0$

$$(4.2) \qquad L_M(z) \overset{\text{def}}{=} \sum_{i \in M} L_i(z) \qquad z \in \vartheta_M.$$

Throughout this report it will be supposed that the following condition is satisfied

(4.3) Condition: For each M with $\vartheta_M \neq 0$ the function $L_M(z)$ is strictly unimodal in the interval ϑ_M.

Now let M_ν $(\nu = 1, 2, \ldots, N)$ be subsets of the numbers $1, 2, \ldots, k$ with

$$(4.4) \quad \begin{cases} 1. \ \bigcup_{\nu=1}^{N} M_\nu = \{1, 2, \ldots, k\}, \\ 2. \ M_{\nu_1} \cap M_{\nu_2} = 0 \text{ for each pair of values } \nu_1, \nu_2 = 1, 2, \ldots, N \text{ with } \nu_1 \neq \nu_2, \\ 3 \ \vartheta_{M_\nu} \neq 0 \text{ for each } \nu = 1, 2, \ldots, N, \end{cases}$$

where

$$(4.5) \qquad \vartheta_{M_\nu} \overset{\text{def}}{=} \bigcap_{i \in M_\nu} \vartheta_i \qquad (\nu = 1, 2, \ldots, N).$$

Let further

$$(4.6) \qquad G_N \overset{\text{def}}{=} \prod_{\nu=1}^{N} \vartheta_{M_\nu}$$

and

$$(4.7) \qquad L_{M_\nu}(z_\nu) \overset{\text{def}}{=} \sum_{i \in M_\nu} L_i(z_\nu) \qquad z_\nu \in \mathfrak{I}_{M_\nu} \quad (\nu = 1, 2, \ldots, N).$$

Then for all points in G_N $L(y_1, y_2, \ldots, y_k)$ reduces to a function of N variables $z_1, z_2, \ldots, z_N$; we denote this function by $L'(z_1, z_2, \ldots, z_N)$ and thus have

$$(4.8) \qquad L'(z_1, z_2, \ldots, z_N) = \sum_{\nu=1}^{N} L_{M_\nu}(z_\nu),$$

which is according to (4.3), a sum of strictly unimodal functions.

<u>Theorem I:</u> <u>L possesses a unique maximum in D</u>

<u>Proof:</u> This theorem will be proved by induction.

Let $M_1, M_2, \ldots, M_N$ be an arbitrary set of subsets of the numbers $1, 2, \ldots, k$ satisfying (4.4) and let

$$(4.9) \qquad D_{N,s} \overset{\text{def}}{=} D \cap G_N,$$

where s denotes the number of essential restrictions defining D and where G_N is defined by (4.6). Then $D_{N,s}$ is convex and:

for $N = k$ we have $\mathfrak{I}_{M_\nu} = \mathfrak{I}_\nu$ $(\nu = 1, 2, \ldots, N)$, thus $G_N = G$ and $D_{N,s} = D$

for $s = 0$ we have $D = G$ thus $D_{N,0} = G_N$.

We shall say that the function $L'(z_1, z_2, \ldots, z_N)$ can be monotonously traced to its maximum in $D_{N,s}$ if

(4.10)
1. $L'(z_1, z_2, \ldots, z_N)$ possesses a unique maximum in $D_{N,s}$,
2. every point of $D_{N,s}$ can be connected with the point in $D_{N,s}$ where L' assumes its maximum by means of a line in $D_{N,s}$ such that L' increases monotonously along this line. (Such a line will be called a trace)

For $s = 0$ $L'(z_1, z_2, \ldots, z_N)$ has this property for every set $M_1, M_2, \ldots, M_N$ satisfying (4.4) and every N. This follows from the fact that L' is the sum of strictly unimodal functions and that $D_{N,0}$ is the Cartesian product of the intervals $\mathfrak{I}_{M_\nu}$ $(\nu = 1, 2, \ldots, N)$, so that the Lemma's I and II may be applied.

Let us now suppose that it has been proved that L' can be monotonously traced to its maximum for all values of $s \leq s_0$ for every set $M_1, M_2, \ldots, M_N$ satisfying (4.4) and for every N . We then prove that the same holds for $s_0 + 1$ essential restrictions.

Consider, for a given set $M_1, M_2, \ldots, M_N$, satisfying

(4.4), a domain D_{N,s_0+1} and the domain D_{N,s_0} which is obtained by omitting one of the essential restrictions defining D_{N,s_0+1}. Let this be the restriction $R_\lambda : z_{i_\lambda} \leqq z_{j_\lambda}$. Then clearly

(4.11) $$D_{N,s_0+1} \subset D_{N,s_0}.$$

Now L' has a unique maximum in D_{N,s_0}, attained in (say) the point T. We first consider the case that T is outside D_{N,s_0+1}. Then an arbitrary point P of D_{N,s_0+1} with $z_{i_\lambda} < z_{j_\lambda}$ can be connected with T by means of a trace in D_{N,s_0} and this trace must contain at least one border point of D_{N,s_0+1} with $z_{i_\lambda} = z_{j_\lambda}$, because within D_{N,s_0+1} we have: $z_{i_\lambda} < z_{j_\lambda}$ and outside D_{N,s_0+1}: $z_{i_\lambda} > z_{j_\lambda}$. The first of these points when following the trace be denoted by U; then L' assumes a larger value in U than in P. Now U lies in a domain $D_{N',s_0'}$, where $N' = N-1$ and $s_0' \leqq s_0$ and L' can thus monotonously be traced from U to its unique maximum in $D_{N',s_0'}$ by means of a trace within $D_{N',s_0'}$. The trace from P to U in D_{N,s_0+1} and from U to the maximum of L' in $D_{N',s_0'}$ together form a trace from P to the maximum of L' in D_{N,s_0+1}.

Consider next the case where T is a point of D_{N,s_0+1}. Then L' attains a unique maximum in D_{N,s_0+1} in T. If T is the maximum of L' in G_N then, according to Lemma II, L' can be monotonously traced to its maximum from every point of D_{N,s_0+1} by means of a straight line, connecting this point with T. If T is not the maximum of L' in G_N then it follows from Lemma III that T is a border point of D_{N,s_0+1} where at least two z_ν from $z_1, z_2, \ldots, z_N$ corresponding to an essential restriction for D_{N,s_0+1} are equal. Let this pair be

(4. 12) $$z_{i_\mu} = z_{j_\mu},$$

then we consider the domain D'_{N,s_0} which is obtained from D_{N,s_0+1} by omitting the restriction $R_\mu : z_{i_\mu} \leqq z_{j_\mu}$ from the essential restrictions defining D_{N,s_0+1}. The maximum of L' in D'_{N,s_0} then exists and the point where it is attained is a point of D'_{N,s_0} with $z_{i_\mu} \geqq z_{j_\mu}$. The rest of the proof for this case is then the same as for the first case considered.

Thus L' can be monotonously traced to its maximum in every $D_{N,s}$, one of which is D.

<u>Remark 2:</u> For $s=0$ and $N=k$ we have $D_{N,s} = G$. Thus L attains a unique maximum in G in a point which will be denoted by $v_1, v_2, \ldots, v_k$.

Theorem II: If $t'_1, t'_2, \ldots, t'_k$ are the values of $y_1, y_2, \ldots, y_k$ which maximixe L in G and under the restrictions $R_1, \ldots, R_{\lambda-1}, R_{\lambda+1}, \ldots, R_s$ then

$$(4.13)\quad \begin{cases} 1.\ t_i = t'_i \quad (i=1,2,\ldots,k) \quad \text{if} \quad t'_{i_\lambda} \leqq t'_{j_\lambda}, \\ 2.\ t_{i_\lambda} = t_{j_\lambda} \quad \text{if} \quad t'_{i_\lambda} > t'_{j_\lambda}. \end{cases}$$

Proof: The R_λ have not been arranged in a special order, thus we may take without any loss of generality $\lambda = s$. First consider the case that $t'_{i_s} \leqq t'_{j_s}$; then $t'_1, t'_2, \ldots, t'_k$ satisfy all restrictions $R_1, R_2, \ldots, R_s$; thus in this case we have

$$(4.14)\qquad t_i = t'_i \quad (i = 1, 2, \ldots, k).$$

If $t'_{i_s} > t'_{j_s}$ then (4.13.2) may be proved as follows. The domain defined by the essentional restrictions $R_1, R_2, \ldots, R_{s-1}$ will be denoted by D'. Then for each point $(y_1, y_2, \ldots, y_k)$ in D with $y_{i_s} < y_{j_s}$ there exists a trace in D' from the point $(y_1, y_2, \ldots, y_k)$ to the point $(t'_1, t'_2, \ldots, t'_k)$ and this trace contains a point $(y'_1, y'_2, \ldots, y'_k)$ with

$$(4.15)\quad \begin{cases} 1.\ y'_{i_s} = y'_{j_s}, \\ 2.\ L(y'_1, y'_2, \ldots, y'_k) > L(y_1, y_2, \ldots, y_k). \end{cases}$$

Thus, if $t'_{i_s} > t'_{j_s}$, then $L(y_1, y_2, \ldots, y_k)$ attains its maximum in D for $y_{i_s} = y_{j_s}$; (4.13.2) then follows from the uniqueness of this maximum

Remark 3:

If

$$(4.16)\qquad P[x_i = 1] = \theta_i \quad , \quad P[x_i = 0] = 1 - \theta_i \quad (i = 1, 2, \ldots, k)$$

and

$$(4.17)\qquad a_i \overset{\text{def}}{=} \sum_{\gamma=1}^{n_i} x_{i,\gamma} \quad , \quad b_i \overset{\text{def}}{=} n_i - a_i \quad (i = 1, 2, \ldots, k)$$

then

$$(4.18)\qquad L(y_1, y_2, \ldots, y_k) = \sum_{i=1}^{k} \{ a_i \lg y_i + b_i \lg (1 - y_i).$$

In [2] it has been proved that, if G_i is the interval (0,1), this function L satisfies the following condition.

(4.19) Condition: If $(y_1, y_2, \ldots, y_k)$ and $(z_1, z_2, \ldots, z_k)$ are any two points in G with $y_i \neq z_i$ for at least one value of i and if

$$y_i(\beta) \stackrel{\mathrm{def}}{=} (1-\beta)\, y_i + \beta z_i \qquad (i = 1, 2, \ldots, k)$$

then $L\{y_1(\beta), y_2(\beta), \ldots, y_k(\beta)\}$ is a strictly unimodal function of β in the interval $0 \leqq \beta \leqq 1$.

This condition is stronger than condition (4.3) and the theorems I and II of this report have been proved in [2] by using condition (4.19).

Further if condition (4.19) is satisfied then theorem I of this report may be proved in a more simple way than we did in [2] as follows. Consider any two points $(y_1, y_2, \ldots, y_k)$ and $(z_1, z_2, \ldots, z_k)$ in D with $y_i \neq z_i$ for at least one value of i and

(4.20) $$L(y_1, y_2, \ldots, y_k) = L(z_1, z_2, \ldots, z_k).$$

Then it follows from condition (4.19) that there exists a point $(y_1, y_2, \ldots, y_k)$ in D with

(4.21) $$L(y_1, y_2, \ldots, y_k) > L(y_1, y_2, \ldots, y_k).$$

Thus L possesses a unique maximum in D.

The maximum likelihood estimates of $\theta_1, \theta_2, \ldots, \theta_k$ may always be found by repeatedly applying theorem II. This follows from the fact that $L'(z_1, z_2, \ldots, z_N)$ is a sum of strictly unimodal functions and that $D_{N,s}$ is a convex subdomain of the Cartesian product of the intervals $\mathcal{Y}_{M_\nu}$ $(\nu = 1, 2, \ldots, N)$ for each set $M_1, M_2, \ldots, M_N$ and each N.

This leads however to a rather complicated procedure which may often be simplified by using one of the theorems of the following section.

5. Some special theorems

The theorems III-VI in this section may be proved in precisely the same way as the theorems II-V in [2].

Theorem III: If $\alpha_{i,j}(v_i - v_j) \leqq 0$ for each pair of values (i, j) then

(5.1) $$t_i = v_i \qquad (i = 1, 2, \ldots, k).$$

Theorem IV: If $\ell_1, \ell_2, \ldots, \ell_m$ is a set of values satisfying

(5.2) $$\alpha_{i,\ell_1} = \alpha_{i,\ell_2} = \ldots = \alpha_{i,\ell_m} = 0 \quad \text{for each } i \neq \ell_1, \ell_2, \ldots, \ell_m$$

then the maximum likelihood estimates of $\theta_{\ell_1}, \theta_{\ell_2}, \ldots, \theta_{\ell_m}$ are the values of $y_{\ell_1}, y_{\ell_2}, \ldots, y_{\ell_m}$ which maximize $L_{\ell_1} + L_{\ell_2} + \ldots + L_{\ell_m}$ in the domain

$$(5.3)\qquad D_1 \begin{cases} \alpha_{i,j}(y_i - y_j) \leqq 0 \\ y_i \in \theta_i \end{cases} \qquad (i,j = \ell_1, \ell_2, \ldots, \ell_m).$$

Theorem V: If for some pair of values (i,j) with $i<j$

$$(5.4)\qquad \alpha_{i,j}(v_i - v_j) > 0$$

and

$$(5.5)\qquad \begin{cases} 1.\ \alpha_{i,h} = \alpha_{h,j} = 0 & \text{for each } h \text{ between } i \text{ and } j, \\ 2.\ \alpha_{h,i} = \alpha_{h,j} & \text{for each } h < i, \\ 3.\ \alpha_{i,h} = \alpha_{j,h} & \text{for each } h > j, \end{cases}$$

then

$$(5.6)\qquad t_i = t_j .$$

Theorem VI: If (i,j) is a pair of values satisfying

$$(5.7)\qquad v_i \leqq v_j$$

and

$$(5.8)\qquad \begin{cases} 1.\ \alpha_{i,j} = 0, & \\ 2.\ \alpha_{h,i} \leqq \alpha_{h,j} & \text{for each } h < i, \\ 3.\ \alpha_{i,h} \geqq \alpha_{j,h} & \text{for each } h > j, \end{cases}$$

then

$$(5.9)\qquad t_i \leqq t_j .$$

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

$$(5.10)\qquad \alpha_{i,j} = 0,$$

if D' is the subdomain of D where $y_i \leqq y_j$ and if $(t_1', t_2', \ldots, t_k')$ is the point where L assume its maximum in D' then

$$(5.11)\qquad \begin{cases} 1.\ t_1 = t_1', t_2 = t_2', \ldots, t_k = t_k' & t_i' < t_j', \\ 2.\ t_i \geqq t_j & t_i' = t_j'. \end{cases}$$

Proof: The proof of this theorem differs from the one given for theorem VI in [2] only in the form of the trace from a point in D' to the maximum in D. This trace which is a straight l line in [2], need not be straight now(cf. the proof of theorem II of the present report).

6. Examples

In this section the pooled samples of $\underline{x}_i$ and $\underline{x}_j$ will be denoted by $x_{i,\gamma}'$ $(\gamma = 1, 2, \ldots, n_i')$, where $n_i' = n_i + n_j$.

6.1 $\underline{x}_i$ possesses a normal distribution with mean θ_i and known variance $(i = 1, 2, \ldots, k)$.

Without any loss of generality we may suppose that $\sigma^2\{\underline{x}_i\} = 1$ for all i; then

(6.1.1) $\quad L_i(y_i) = -\frac{1}{2} n_i \lg 2\pi - \frac{1}{2} \sum_{\gamma=1}^{n_i} (x_{i,\gamma} - y_i)^2 \qquad (i = 1, 2, \ldots, k).$

From (6.1.1) it follows that

(6.1.2) $$\frac{d L_i(y_i)}{d y_i} = \sum_{\gamma=1}^{n_i} (x_{i,\gamma} - y_i) \qquad (i = 1, 2, \ldots, k),$$

thus, if

(6.1.3) $$m_i \stackrel{\text{def}}{=} \frac{1}{n_i} \sum_{\gamma=1}^{n_i} x_{i,\gamma} \qquad (i = 1, 2, \ldots, k),$$

then

(6.1.4) $$\frac{d L_i(y_i)}{d y_i} \begin{cases} > 0 & \text{if} \quad y_i < m_i, \\ = 0 & \text{if} \quad y_i = m_i, \\ < 0 & \text{if} \quad y_i > m_i. \end{cases} \qquad (i = 1, 2, \ldots, k).$$

From (6.1.4) it follows that $L_i(y_i)$ is a **strictly unimodal** function of y_i in the interval $(-\infty, +\infty)$, thus $L_i(y_i)$ is a strictly unimodal function of y_i in each closed subinterval ϑ_i of the interval $(-\infty, +\infty)$ $(i = 1, 2, \ldots, k)$.
Further if $y_i = y_j$ then $L_i(y_i) + L_j(y_j)$ reduces to one term of the form
(6.1.5) $\quad L_i(y_i) + L_j(y_i) = -\frac{1}{2} n'_i \lg 2\pi - \frac{1}{2} \sum_{\gamma=1}^{n'_i} (x'_{i,\gamma} - y_i)^2$
and analogous relations hold if more than two of the y_i are equal. Thus L satisfies condition (4.3).
From (6.1.5) it follows further that if L attain its maximum for $y_i = y_j$ then the two samples of $\underline{x}_i$ and $\underline{x}_j$ are to be pooled.

The procedure will now be illustrated by means of the following example.
Suppose $k = 4$, $r_0 = 2$, $r_1 = 4$ and
(6.1.6) $\quad \alpha_{1,2} = \alpha_{1,3} = \alpha_{3,4} = 1.$
Let further

(6.1.7)

i	1	2	3	4
$x_{i,\gamma}$	-0,40	1,43	-0,70	0,29
	2,56	1,86	2,61	0
	0,25	0,06	0,79	1,31
	2,87	0,07	0,86	0,15
		1,14	0,14	2,53
		0,29		1,86
		2,57		
		0,85		
		1,21		
$n_i m_i$	5,28	9,48	3,70	6,14
n_i	4	9	5	6
m_i	1,32	1,05	0,74	1,02

and let $\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4$ be the intervals

(6.18)

i	1	2	3	4
ϑ_i	$(-\infty, 1)$	$(-\infty, +\infty)$	$(0,5, \infty)$	$(-\infty, +\infty)$

Then it follows from (6.1.7) and (6.1.8) that the coordinates of the maximum in G are

(6.1.9)

i	1	2	3	4
v_i	1	1,05	0,74	1,02

From (6.1.6) and (6.1.9) it then follows that the pairs $i=3$, $j=2$ and $i=4, j=2$ satisfy (5.7) and (5.8). Thus according to theoremVI L attains its maximum in D for

(6.1.10) $$y_1 \leqq y_3 \leqq y_4 \leqq y_2 .$$

From (6.1.9), (6.1.10) and theorem V then follows

(6.1.11) $$t_1 = t_3 .$$

In this way the problem is reduced to the case of 3 samples with $\varkappa_0' = 0$,

(6.1.12) i	$1(+3)$	4	2
$x'_{i,\gamma}$	-0,40	0,29	1,43
	2,56	0	1,86
	0,25	1,31	0,06
	2,87	0,15	0,07
	-0,70	2,53	1,14
	2,61	1,86	0,29
	0,79		2,57
	0,86		0,85
	0,14		1,21
$n'_i m'_i$	8,98	6,14	9,48
n'_i	9	6	9
m'_i	0,998	1,02	1,05
ϑ'_i	(0,5 , 1)	$(-\infty, +\infty)$	$(-\infty, +\infty)$
v'_i	0,998	1,02	1,05

and

(6.1.13) $$\alpha'_{1,4} = \alpha'_{4,2} = 1.$$

From (6.1.11), (6.1.12) and (6.1.13) then follows

(6.1.14) $$t_1 = t_3 = 0{,}998\ ,\ t_2 = 1{,}05\ ,\ t_4 = 1{,}02.$$

6.2. x_i possesses a normal distribution with known mean and variance θ_i $(i = 1, 2, \ldots, k)$.

We suppose without loss of generality $\mathcal{E}\, x_i = 0\ (i = 1, 2, \ldots, k)$; then

(6.2.1) $$L_i(y_i) = -\tfrac{1}{2} n_i \lg 2\pi - \tfrac{1}{2} n_i \lg y_i - \tfrac{1}{2} \frac{\sum_{\gamma=1}^{n_i} x_{i,\gamma}^2}{y_i} \qquad (i = 1, 2, \ldots, k).$$

From (6.2.1) it follows, if

(6.2.2) $$s_i^2 \overset{\text{def}}{=} \frac{1}{n_i} \sum_{\gamma=1}^{n_i} x_{i,\gamma}^2 \qquad (i = 1, 2, \ldots, k),$$

that

(6.2.3) $$\frac{d L_i(y_i)}{d y_i} \begin{cases} > 0 & \text{if} \quad 0 \leqq y_i < s_i^2, \\ = 0 & \text{if} \quad y_i = s_i^2 \\ < 0 & \text{if} \quad y_i > s_i^2 \end{cases} \qquad (i = 1, 2, \ldots, k);$$

thus $L_i(y_i)$ is a strictly unimodal function of y_i in the interval $(0, \infty)$.

Further if $y_i = y_j$ then $L_i(y_i) + L_j(y_j)$ reduces to

(6.2.4) $$L_i(y_i)+L_j(y_i) = -\tfrac{1}{2}n'_i \lg 2\pi - \tfrac{1}{2}n'_i \lg y_i - \tfrac{1}{2}\frac{\sum_{\gamma=1}^{n'_i} x'^2_{i,\gamma}}{y_i}$$

and analogously for more than two of the y_i equal.
Thus L satisfies condition (4.3) and if L attains its maximum for $y_i = y_j$ then the two samples of $\underline{x}_i$ and $\underline{x}_j$ are to be pooled. Numerically the method is thus precisely the same as in 6.1, with s_i^2 in stead of m_i .

6.3 <u>$\underline{x}_i$ possesses a Poisson distribution with parameter θ_i $(i=1,2,\dots,k)$</u>

In this case we have

(6.3.1) $$L_i(y_i) = -n_i y_i + \sum_{\gamma=1}^{n_i} x_{i,\gamma} \lg y_i - \sum_{\gamma=1}^{n_i} \lg x_{i,\gamma}! \quad (i=1,2,\dots,k).$$

From (6.3.1) it follows that, if

(6.3.2) $$m_i \overset{\text{def}}{=} \frac{1}{n_i}\sum_{\gamma=1}^{n_i} x_{i,\gamma} \qquad (i=1,2,\dots,k),$$

then

(6.3.3) $$\frac{dL_i(y_i)}{dy_i} \begin{cases} >0 & \text{if } 0 \leqq y_i < m_i, \\ =0 & \text{if } y_i = m_i, \\ <0 & \text{if } y_i > m_i; \end{cases} \qquad (i=1,2,\dots,k)$$

thus $L_i(y_i)$ is a strictly unimodal function of y_i the interval $(0,\infty)$ $(i=1,2,\dots,k)$.
Further if $y_i = y_j$ then $L_i(y_i)+L_j(y_j)$ reduces to

(6.3.4) $$L_i(y_i)+L_j(y_i) = -n'_i y_i + \sum_{\gamma=1}^{n'_i} x'_{i,\gamma} \lg y_i - \sum_{\gamma=1}^{n'_i} x'_{i,\gamma}!\ ;$$

thus L satisfies condition (4.3) and if L attains its maximum for $y_i = y_j$ then the two samples of $\underline{x}_i$ and $\underline{x}_j$ are to be pooled.

The theorems of the foregoing sections may e.g. also be applied in the following case.

6.4. <u>$\underline{x}_i$ possesses a normal distribution with mean θ_i and known variance for $i=l_1,l_2,\dots,l_g$ and a Poisson distribution with parameter θ_i for $i \neq l_1,l_2,\dots,l_g$.</u>

Taking $\sigma^2\{\underline{x}_i\}=1$ for $i=l_1,l_2,\dots,l_g$ we have

(6.4.1) $$\begin{cases} L_i(y_i) = -\tfrac{1}{2}n_i \lg 2\pi - \tfrac{1}{2}\sum_{\gamma=1}^{n_i}(x_{i,\gamma}-y_i)^2 & (i=l_1,l_2,\dots,l_g), \\ L_i(y_i) = -n_i y_i + \sum_{\gamma=1}^{n_i} x_{i,\gamma}\lg y_i - \sum_{\gamma=1}^{n_i}\lg x_{i,\gamma}! & (i\neq l_1,l_2,\dots,l_g). \end{cases}$$

From the sections 6.1 and 6.3 it follows that $L_i(y_i)$ is a strictly unimodal function of y_i in the interval $(-\infty,+\infty)$ for $i=\ell_1,\ell_2,\dots,\ell_g$ and in the interval $(0,\infty)$ for $i\neq\ell_1,\ell_2,\dots,\ell_g$. Further, if $y_i=y_j$, where $\underline{x}_i$ possesses a normal and $\underline{x}_j$ a Poisson distribution then $L_i(y_i)+L_j(y_j)$ reduces to

$$(6.4.3)\qquad L_i(y_i)+L_j(y_i)=-\tfrac{1}{2}n_i\lg 2\pi-\tfrac{1}{2}\sum_{\gamma=1}^{n_i}(x_{i,\gamma}-y_i)^2-n_jy_i+\\+\sum_{\gamma=1}^{n_j}x_{j,\gamma}\lg y_i-\sum_{\gamma=1}^{n_j}\lg x_{j,\gamma}!\,.$$

It may be proved as follows that $L_{i,j}(y_i)\overset{\text{def}}{=}L_i(y_i)+L_j(y_i)$ is a strictly unimodal function of y_i in the interval $(0,\infty)$. We have

$$(6.4.4)\qquad \frac{dL_{i,j}(y_i)}{dy_i}=n_i(m_i-y_i)-n_j+\frac{n_jm_j}{y_i}\,.$$

Thus if $m_i-\frac{n_j}{n_i}\leq 0$ and $m_j=0$ then

$$(6.4.5)\qquad \frac{dL_{i,j}(y_i)}{dy_i}<0 \text{ for each } y_i>0$$

and in all other cases

$$(6.4.6)\qquad \frac{dL_{i,j}(y_i)}{dy_i}\begin{cases}>0 \text{ if } 0\leq y_i<m_i'\overset{\text{def}}{=}\frac{1}{2}\left\{m_i-\frac{n_j}{n_i}+\sqrt{\left(m_i-\frac{n_j}{n_i}\right)^2+4\frac{n_jm_j}{n_i}}\right\},\\ =0 \text{ if } y_i=m_i',\\ <0 \text{ if } y_i>m_i'.\end{cases}$$

Analogous relations hold if more than two of the y_i are equal. Thus L satisfies condition (4.3).
This case will be illustrated by means of the following example.
Suppose $k=4$, $r_0=r_1=3$,

$$(6.4.7)\qquad \alpha_{1,2}=\alpha_{1,4}=\alpha_{3,4}=1$$

and $\ell_1=1$, $\ell_2=2$, $g=2$. Further

(6.4.8)

i	1	2	3	4
$x_{i,\gamma}$	5,38	4,84	4	2
	3,88	3,56	5	7
	4,14	4,40	3	5
	5,36	4,77	3	4
	5,48		4	
$n_i m_i$	24,24	17,57	19	18
n_i	5	4	5	4
m_i	4,85	4,39	3,8	4,5
$\mathcal{Y}_i$	$(-\infty, 5)$	$(-\infty, +\infty)$	$(0, \infty)$	$(0, 4)$
v_i	4,85	4,39	3,8	4

Then the pairs $i=3$, $j=2$; $i=4, j=2$ and $i=3, j=1$ satisfy (5.7) and (5.8). Thus the problem is reduced to the case of the 4 samples (6.4.8) with $\alpha_0' = 0$ and

(6.4.9) $$\alpha'_{3,1} = \alpha'_{1,4} = \alpha'_{4,2} = 1.$$

From (6.4.3), (6.4.9) and theorem V then follows

(6.4.10) $$t_1 = t_4.$$

In this way the problem is reduced to the problem of maximizing the function

(6.4.11) $$L'(y_1, y_2, y_3) \overset{\text{def}}{=} L(y_1, y_2, y_3, y_1)$$

in the domain

(6.4.12) $$D' \begin{cases} 0 \leqq y_3 \leqq y_1 \leqq y_2, \\ y_1 \leqq 4. \end{cases}$$

From (6.4.5) and (6.4.6) it follows that

(6.4.13)

i	3	1	2
m_i'	3,8	4,8	4,39
$\mathcal{Y}_i'$	$(0, \infty)$	$(0, 4)$	$(-\infty, +\infty)$
v_i'	3,8	4	4,39

Thus

(6.4.14) $$t_1 = t_4 = 4\ ,\quad t_2 = 4{,}39\ ,\quad t_3 = 3{,}8.$$

References

[1] BRINK, H.D., Maximum likelihood estimates of monotone parameters, Ann. Math. Stat. 26 (1955), 607-615.

[2] van Eeden, Constance, Maximum likelihood estimation of ordered probabilities, Proc. Kon. Ned. Akad. v. Wet. A 59 (1956), Indagationes Mathematicae 18 (1956)