KONINKL. NEDERL. AKADEMIE VAN WETENSCHAPPEN – AMSTERDAM
Reprinted from Proceedings, Series A, **60**, No. 5 and Indag. Math., **19**, No. 5, 1957

MATHEMATICS

NOTE ON TWO METHODS FOR ESTIMATING ORDERED PARAMETERS OF PROBABILITY DISTRIBUTIONS [1)]

BY

CONSTANCE VAN EEDEN

(Communicated by Prof. D. VAN DANTZIG at the meeting of May 25, 1957)

1. *Introduction and summary*

Methods to find maximum likelihood estimates for partially or completely ordered sets of parameters have recently been developed independently by H. D. BRUNK and the present author. Relevant references are given at the end of this paper. The special case of ordered sets of unknown probabilities was treated in [1] by BRUNK and other authors and in [3], [4] and [5] by the present author.

At first sight the two methods, indicated as method B and A respectively do not look at all alike. Also the conditions imposed are different, those imposed by BRUNK (method B) being the more stringent. It therefore seemed worthwile to give a proof of the identity of the two methods at the same time indicating that method B is also valid under the more general conditions imposed on method A.

In order to achieve this purpose method B is first described in the notation of method A (section 2) and the proof of the identity of the two methods is given in the sections 3 and 4.

2. *Description of method B*

The situation in which method B can be applied is, in our notation, described in [2] as follows. "Let u denote an n-tuple, $u=(u^1, \ldots, u^n)$, of real numbers and let u_i $(i=1, \ldots, k)$ denote one member of a set of k such n-tuples. Let further $\mathbf{x}_1, \ldots, \mathbf{x}_k$ [2)] be k independent random variables and let the distribution function of $\mathbf{x}_i$ be completely specified by the knowledge of a single parameter θ_i $(i=1, \ldots, k)$. These parameters $\theta_1, \ldots, \theta_k$ satisfy the following monotonicity condition: there is a function $\theta(u)$, monotone non decreasing in each of the separate variables u^j $(j=1, \ldots, n)$, such that $\theta_i=\theta(u_i)$ $(i=1, \ldots, k)$. Further the distribution of $\mathbf{x}_i$ belongs to the "exponential family" $(i=1, \ldots, k)$ and the distribution functions of $\mathbf{x}_i$ and $\mathbf{x}_j$ are identical if and only if $\theta_i=\theta_j$ $(i, j=1, \ldots, k)$". No other restrictions are imposed on the parameters $\theta_1, \ldots, \theta_k$.

[1)] Report SP 55 of the Statistical Department of the Mathematical Centre, Amsterdam.

[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.

Now it may be remarked that the monotonicity of the function $\theta(u)$ is equivalent with the (partial or complete) ordering of the parameters $\theta_1, \ldots, \theta_k$ specified by the following set of inequalities. Let $\alpha_{i,j}$ $(i,j = 1,\ldots,k)$ be numbers satisfying

$$(2.1)\qquad \begin{cases} 1. & \alpha_{i,j} = -\alpha_{j,i}, \\ 2. & \alpha_{i,j} = 1 \text{ if no coordinate of } u_i \text{ is greater than the corresponding coordinate of } u_j, \\ 3. & \alpha_{i,j} = 0 \text{ in all other cases.} \end{cases}$$

Then it follows from the fact that $\theta(u)$ is monotone non decreasing in each of the separate variables u^j $(j=1, \ldots, n)$ and that $\theta_i = \theta(u_i)$ $(i = 1,\ldots,k)$ that $\theta_1, \ldots, \theta_k$ satisfy the inequalities

$$(2.2)\qquad \alpha_{i,j}(\theta_i - \theta_j) \leqq 0 \qquad (i, j = 1, \ldots, k)$$

and this is identical with (2.4) in [6], I_i being in this case the set of all values y for which $F_i(x_i|y)$ is a distribution function $(i = 1, \ldots, k)$.

On the other hand every partial or complete ordering of the θ_i can be represented in the abovementioned way by means of a function $\theta(u)$ in a space of a sufficiently large number of dimensions.

In deriving the maximum likelihood estimates BRUNK does not specify n and the function $\theta(u)$ but only uses the abovementioned monotonicity-property. His solution may be formulated as follows.

If M is a subset of the numbers $1, \ldots, k$ with (cf. (4.1) in [6]) $I_M \neq 0$, then v_M is defined as the value of z which maximizes (cf. (4.2) in [6]) $L_M(z)$ for $z \in I_M$. The existence of v_M follows immediately from the fact that the distributions of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ belong to the "exponential family" (cf. [2], p. 611). Let further

$$(2.3)\qquad \begin{cases} S_i \overset{\text{def}}{=} i \cup \mathscr{Ens}\{j \mid \alpha_{j,i} = 1\}, \\ T_i \overset{\text{def}}{=} i \cup \mathscr{Ens}\{j \mid \alpha_{i,j} = 1\}. \end{cases} \qquad (i = 1, \ldots, k).$$

In [2] S_i (respectively T_i) is called a lower (respectively an upper) interval. Further if M is a subset of the numbers $1, \ldots, k$ then

$$(2.4)\qquad \begin{cases} 1. & S \overset{\text{def}}{=} \bigcup_{i \in M} S_i, \\ 2. & T \overset{\text{def}}{=} \bigcup_{i \in M} T_i \end{cases}$$

and in [2] S (respectively T) is called a lower (respectively an upper) layer. The complement $\bar{S}$ of a lower layer S is an upper layer with respect to an other M and vice versa. Theorem I in [2] then states that

$$(2.5)\qquad t_i = \max_{\substack{T \\ i \in T \cap \bar{S}}} \min_{S} v_{T \cap \bar{S}} \qquad (i = 1, \ldots, k).$$

In the special case of estimating completely ordered probabilities with $I_i \equiv (0, 1)$ $(i=1, \ldots, k)$ (2.5) reduces to (cf. [1], p. 644)

$$t_i = \max_{1 \leqq r \leqq i} \min_{i \leqq s \leqq k} \frac{a_r + \ldots + a_s}{n_r + \ldots + n_s} \qquad (i=1, \ldots, k). \tag{2.6}$$

It may easily be seen (cf. [2], p. 611) that condition (4.3) of [6] is satisfied if the distributions of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ belong to the exponential family. Thus method A may be applied if the conditions of method B are satisfied.

On the other hand the conditions for method B as mentioned in [2] need not be satisfied if the conditions for method A are satisfied; we have e.g.

1. if $\mathbf{x}_i$ possesses a rectangular distribution between 0 and $\theta_i (i=1, \ldots, k)$ then condition (4.3) of [6] is satisfied, but the distributions of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ do not belong to the "exponential family",

2. if $\mathbf{x}_i$ possesses a normal distribution with mean θ_i and variance 1 for $i=l_1, \ldots, l_g$ $(1 \leqq g \leqq k-1)$ and a Poisson distribution with parameter θ_i for $i \neq l_1, \ldots, l_g$ then condition (4.3) of [6] is satisfied. There exists however at least one pair of values (i, j) such that, for $\theta_i = \theta_j$, $\mathbf{x}_i$ and $\mathbf{x}_j$ do not possess the same probability distribution,

3. for method A it is not necessary that I_i is the set of all values of y for which $F_i(x_i|y)$ is a distribution function.

In section 4 it will be proved that (2.5) also holds if the conditions for method A are satisfied. For that purpose we need some lemmas which will be proved in section 3.

3. *Lemmas*

In this and the following sections we suppose that the conditions for method A are satisfied and unless explicitely stated otherwise the function $L(y_1, \ldots, y_k)$ will only be considered in the domain D. Further the set $\{1, \ldots, k\}$ will be denoted by E and the complement of a subset M of E by $\bar{M}$, i.e.

$$\begin{cases} M \cup \bar{M} = E, \\ M \cap \bar{M} = 0. \end{cases} \tag{3.1}$$

Lemma I. *If for any pair of values (i, j)*

$$\begin{cases} \alpha_{i,j} = 0, \\ t_i \leqq t_j \end{cases} \tag{3.2}$$

then the estimates $t_1, \ldots, t_k$ may also be found by maximizing L in the sub-domain D' of D where $y_i \leqq y_j$.

Proof: This lemma follows immediately from the fact that $D' \subset D$ and that $(t_1, \ldots, t_k) \in D'$.

Lemma II. *If for any value of λ*

$$t_{i_\lambda} < t_{j_\lambda} \tag{3.3}$$

then the estimates $t_1, \ldots, t_k$ *may also be found by maximizing* L *in the domain* D'' *which is obtained from* D *by omitting the restriction* R_λ.

Proof: If $t'_1, \ldots, t'_k$ are the values of $y_1, \ldots, y_k$ which maximize L in D'' then it follows from theorem II of [6] that $t_{i_\lambda}=t_{j_\lambda}$ if and only if $t'_{i_\lambda}\geqq t'_{j_\lambda}$. From (3.3) then follows $t'_{i_\lambda}<t'_{j_\lambda}$; thus $t_1=t'_1, \ldots, t_k=t'_k$.

Lemma III. *If the parameters* $\theta_1, \ldots, \theta_k$ *are completely ordered, if there exists a value* $i\geqq 0$ *and a value* $h\geqq 2$ *such that*

$$t_{i+1}= \ldots =t_{i+h} \tag{3.4}$$

and if $l_1, \ldots, l_h$ *is a permutation of the numbers* $i+1, \ldots, i+h$ *with*

$$v_{l_1}\geqq \ldots \geqq v_{l_h}, \tag{3.5}$$

then the estimates $t_1, \ldots, t_k$ *may also be found by maximizing* L *under the restrictions*

$$y_1\leqq \ldots \leqq y_i \leqq y_{l_1}\leqq \ldots \leqq y_{l_h}\leqq y_{i+h+1}\leqq \ldots \leqq y_k. \tag{3.6}$$

Proof: If D' is the subdomain of D where $y_{i+1}= \ldots =y_{i+h}$ then it follows from (3.4) that $(t_1, \ldots, t_k)\in D'$. Further it follows from theorem V of [6] that the point where L attains its maximum under the restrictions (3.6) also lies in D'. The lemma then follows from the uniqueness of the solution.

Lemma IV. *If the parameters* $\theta_1, \ldots, \theta_k$ *are completely ordered, if*

$$t_1= \ldots =t_k \tag{3.7}$$

and if M_r *consists of the numbers* $1, \ldots, r$ $(1\leqq r\leqq k-1)$ *then*

$$v_{M_r}\geqq v_{\bar{M}_r}. \tag{3.8}$$

Proof: If $l_1, \ldots, l_r$ is a permutation of the numbers $1, \ldots, r$ with $v_{l_1}\geqq \ldots \geqq v_{l_r}$ and if $l_{r+1}, \ldots, l_k$ is a permutation of the numbers $r+1, \ldots, k$ with $v_{l_{r+1}}\geqq \ldots \geqq v_{l_k}$ then it follows from lemma III and from the relation $y_{l_r}\leqq y_{l_{r+1}}$ (derived from the complete ordering), that the maximum of L in D coincides with the maximum of L under the restrictions

$$y_{l_1}\leqq \ldots \leqq y_{l_k}. \tag{3.9}$$

From theorem II of [6] with $i_\lambda=l_r$ and $j_\lambda=l_{r+1}$ it follows that $t_{l_r}=t_{l_{r+1}}$ if and only if $t'_{l_r}\geqq t'_{l_{r+1}}$. Further it follows from theorem V of [6] that

$$t'_{l_r}=v_{M_r}\ ,\ t'_{l_{r+1}}=v_{\bar{M}_r}. \tag{3.10}$$

The lemma then follows from (3.7) and (3.10).

Lemma V. *If*

$$t_1= \ldots =t_k \tag{3.11}$$

and if, for given M, S (cf. (2.4.11) and $\bar{S}$ (cf. (3.1)) satisfy

(3.12) $$S \neq 0, \ \bar{S} \neq 0,$$

then

(3.13) $$v_S \geqq v_{\bar{S}}.$$

Proof: For the case of completely ordered parameters this lemma is identical with lemma IV.

From lemma I and (3.11) it follows that for each pair of values (i, j) with $\alpha_{i,j}=0$ the restriction $y_i \leqq y_j$ or $y_i \geqq y_j$ may be added. Such a restriction is added for each pair of values $(i, j) \in S$ with $\alpha_{i,j}=0$ and for each pair of values $(i, j) \in \bar{S}$ with $\alpha_{i,j}=0$ in such a way that within S and within $\bar{S}$ a complete ordering is obtained. This new ordering of the parameters will be denoted by $\alpha'_{i,j}(\theta_i-\theta_j) \leqq 0$ $(i, j \in S)$. Then there exists a value $l_1 \in S$ with $\alpha'_{j,l_1}=1$ for each $j \in S$ and a value $l_2 \in \bar{S}$ with $\alpha'_{l_2,j}=1$ for each $j \in \bar{S}$. Further it follows from the definition of S that $\alpha_{i,j} \geqq 0$ for each pair of values (i, j) with $i \in S$, $j \in \bar{S}$, thus $\alpha'_{l_1,l_2}=\alpha_{l_1,l_2} \geqq 0$. If $\alpha'_{l_1,l_2}=1$ we have obtained a complete ordering; if $\alpha'_{l_1,l_2}=0$ we add (cf. lemma I) the restriction $y_{l_1} \leqq y_{l_2}$. The lemma then follows from lemma IV.

Remark. The lemma holds analogously for any T and $\bar{T}$ with $T \neq 0, \bar{T} \neq 0$

(3.14) $$v_T \leqq v_{\bar{T}}.$$

Lemma VI. *If M_1 and M_2 are two subsets of E with*

(3.15) $$\begin{cases} 1. \ M_1 \neq 0, M_2 \neq 0, \\ 2. \ M_1 \cap M_2 = 0, \\ 3. \ v_{M_1} \leqq v_{M_2}, \end{cases}$$

then

(3.16) $$v_{M_1} \leqq v_{M_1 \cup M_2} \leqq v_{M_2}.$$

Proof: This lemma follows easily from condition (4.3) of [6].

Lemma VII. *If*

(3.17) $$t_1 = \ldots = t_k (=t),$$

then

(3.18) $$t = \max_T v_T = \min_S v_S.$$

Proof: From lemma V and VI it follows that for any $S \neq 0$

(3.19) $$v_S \geqq v_{S \cup \bar{S}} \geqq v_{\bar{S}}$$

and from (3.17) follows

(3.20) $$t = v_E = v_{S \cup \bar{S}}.$$

Thus

(3.21) $$v_S \geqq t \geqq v_{\bar{S}}.$$

From (3.20) and the first inequality of (3.21) then follows

$$t = \min_S v_S. \tag{3.22}$$

The other part of (3.18) follows analogously.

4. *Proof of formula* (2.5)

We first consider the case that $t_1 = \ldots = t_k$. Then it follows from lemma VII that it is sufficient to prove that for every i

$$\max_T \min_{\substack{S \\ i \in T \cap S}} v_{T \cap S} = \min_S v_S = \max_T v_T. \tag{4.1}$$

The following relation always holds

$$\max_T \min_{\substack{S \\ i \in T \cap S}} v_{T \cap S} \geqq \min_{\substack{S \\ i \in T_0 \cap S}} v_{T_0 \cap S} \quad \text{for any } T_0. \tag{4.2}$$

Thus taking $T_0 \equiv E$, we have for given i

$$\max_T \min_{\substack{S \\ i \in T \cap S}} v_{T \cap S} \geqq \min_{\substack{S \\ i \in S}} v_S \geqq \min_S v_S. \tag{4.3}$$

In an analogous way it may be proved that

$$\max_T \min_{\substack{S \\ i \in T \cap S}} v_{T \cap S} \leqq \max_T v_T \tag{4.4}$$

and (4.1) then follows from (3.18), (4.3) and (4.4).

We now consider the case that there exists at least one pair of values (i, j) with $t_i \neq t_j$.

Let M_ν $(\nu = 1, \ldots, N)$ be N subsets of E with

$$\begin{cases} 1. \ \bigcup_{\nu=1}^{N} M_\nu = E, \\ 2. \ t_i < t_j \text{ for each pair of values } (i, j) \text{ with } i \in M_{\nu_1}, j \in M_{\nu_2} \\ \qquad\qquad (\nu_1 < \nu_2;\ \nu_1, \nu_2 = 1, \ldots, N), \\ 3. \ t_i = t_j \text{ for each pair of values } (i, j) \in M_\nu \ (\nu = 1, \ldots, N). \end{cases} \tag{4.5}$$

Denoting the value of t_i for $i \in M_\nu$ by t'_ν $(\nu = 1, \ldots, N)$ it follows from theorem IV of [6] and the lemmas II and VII that

$$t'_\nu = \max_T v_{T \cap M_\nu} = \min_S v_{S \cap M_\nu} \qquad (\nu = 1, \ldots, N). \tag{4.6}$$

From (4.6) it follows that

$$v_{S \cap M_\nu} \geqq t'_\nu \qquad \text{for each } S \text{ with } S \cap M_\nu \neq 0 \ (\nu = 1, \ldots, N), \tag{4.7}$$

thus if $M'_\nu \overset{\text{def}}{\equiv} \bigcup_{\mu=\nu}^{N} M_\mu$ $(\nu = 1, \ldots, N)$ then (cf. lemma VI)

$$v_{S \cap M'_\nu} = v_{\bigcup_{\mu=\nu}^{N} (M_\mu \cap S)} \geqq \min_{\nu \leqq \mu \leqq N} v_{M_\mu \cap S} \geqq \min_{\nu \leqq \mu \leqq N} t'_\mu = t'_\nu. \tag{4.8}$$

Further it follows from (4.2) for $i \in M_\nu$ with $T_0 \equiv M'_\nu$

(4.9) $$\max_{\substack{T \\ i \in T \cap S}} \min_{S} v_{T \cap S} \geqq \min_{\substack{S \\ i \in S \cap M'_\nu}} v_{S \cap M'_\nu} \geqq t'_\nu = \min_{S} v_{S \cap M_\nu}.$$

In an analogous way it may be proved that for $i \in M_\nu$

(4.10) $$\max_{\substack{T \\ i \in T \cap S}} \min_{S} v_{T \cap S} \leqq \max_{T} v_{T \cap M_\nu}.$$

Formula (2.5) then follows from (4.6), (4.9) and (4.10).

In a later paper the interesting inequality in [1] (p. 644) will be generalized and interpreted geometrically.

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**, 641–647 (1955).
2. Brunk, H. D., Maximum likelihood estimates of monotone parameters, Ann. Math. Stat. **26**, 607–616 (1955).
3. Eeden, Constance, van, Maximum likelihood estimation of ordered probabilities I, Report S 188 (VP 5) of the Statistical Department of the Mathematical Centre, Amsterdam, January 1956.
4. ————, Maximum likelihood estimation of ordered probabilities II, Report S 196 (VP 7) of the Statistical Department of the Mathematical Centre, Amsterdam, March 1956.
5. ————, Maximum likelihood estimation of ordered probabilities, Proc. Kon. Ned. Akad. v. Wet. **A 59**, (1956), Indagationes Mathematicae **18**, 444–455 (1956).
5. ————, Maximum likelihood estimation of partially or completely ordered parameters, I and II, Proc. Kon. Ned. Akad. v. Wet. **A 60**, (1957), Indagationes Mathematicae **19**, 128–136 and 201–211 (1957).