MAXIMUM LIKELIHOOD ESTIMATION OF PARTIALLY OR COMPLETELY ORDERED PARAMETERS. II

BY

CONSTANCE VAN EEDEN

(Communicated by Prof. D. VAN DANTZIG at the meeting of November 24, 1956)

6. *The consistency of the estimates*

In this section the consistency of the estimates will be investigated. The method used stems from a paper by A. WALD [3], but is modified by condition (4.3), which does not occur in his paper.

Let, for $f_i(x_i|\theta_i)>0$,

$$g_i(x_i\,|\,y_i,\theta_i)\overset{\text{def}}{=}\ln\frac{f_i(x_i\,|\,y_i)}{f_i(x_i\,|\,\theta_i)}\qquad(i=1,2,\ldots,k),\tag{6.1}$$

then

$$L_i(y_i)=\sum_{\gamma=1}^{n_i}g_i(x_{i,\gamma}\,|\,y_i,\theta_i)+\sum_{\gamma=1}^{n_i}\ln f_i(x_{i,\gamma}\,|\,\theta_i)\qquad(i=1,2,\ldots,k)\tag{6.2}$$

and 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 $\sum_{i=1}^{k}\sum_{\gamma=1}^{n_i} g_i(x_{i,\gamma}|y_i, \theta_i)$ in the domain D, the last term in (6.2) being constant.

Further

$$g_i(x_i\,|\,\theta_i,\theta_i)=0 \text{ for each } x_i\quad(i=1,2,\ldots,k)\tag{6.3}$$

and from condition (4.3) it follows that $g_i(x_i\,|\,y_i, \theta_i)$ is, for each x_i, a strictly unimodal function of y_i in the interval $I_i(i=1, 2, \ldots, k)$.

Let I_i $(i=1, 2, \ldots, k)$ be the interval $c_i \leqq y_i \leqq d_i$ (with $c_i<d_i$) and let $\eta_1, \eta_2, \ldots, \eta_k$ be k numbers satisfying

$$\begin{cases}0<\eta_i\leqq \min(\theta_i-c_i, d_i-\theta_i) \text{ if } \theta_i \text{ is an innerpoint of } I_i,\\ 0<\eta_i\leqq d_i-c_i \text{ if } \theta_i \text{ is a borderpoint of } I_i.\end{cases}\tag{6.4}$$

Let further $I_i(\eta_i)$ denote the set of all values $y_i \in I_i$ satisfying

$$|y_i-\theta_i|\leqq\eta_i\qquad(i=1,2,\ldots,k).\tag{6.5}$$

In the following it will be supposed that the following condition is satisfied.

(6.6) *Condition: There exist k values $\eta_1, \eta_2, \ldots, \eta_k$ satisfying* (6.4) *such that*

$$\begin{cases}1.\quad \mathscr{E}\{g_i(\mathbf{x}_i\,|\,y_i,\theta_i)\,|\,\theta_i\}<0, & \text{for each } y_i\in I_i(\eta_i) \text{ with } y_i\neq\theta_i\\ 2.\quad \dfrac{\sigma^2\{g_i(\mathbf{x}_i\,|\,y_i,\theta_i)\,|\,\theta_i\}}{[\mathscr{E}\{g_i(\mathbf{x}_i\,|\,y_i,\theta_i)\,|\,\theta_i\}]^2}<\infty & (i=1,2,\ldots,k).\end{cases}$$

Some of the conditions mentioned in WALD's paper [3] are in our case sufficient for (6.6.1) and may therefore be useful for the application of our theorems. These conditions may be stated as follows.

Lemma III: *If condition* (4.3) *is satisfied, if* η_i *satisfies* (6.4) *and if*

$$(6.7)\quad \begin{cases} 1. & \mathscr{E}\{\ln f_i(\mathbf{x}_i \mid y_i) \mid \theta_i\} < \infty \text{ for each } y_i \in I_i(\eta_i) \text{ with } y_i \neq \theta_i, \\ 2. & -\infty < \mathscr{E}\{\ln f_i(\mathbf{x}_i \mid \theta_i) \mid \theta_i\} < \infty \end{cases}$$

then

$$(6.8)\qquad \mathscr{E}\{g_i(\mathbf{x}_i|y_i,\, \theta_i)|\theta_i\} < 0 \text{ for each } y_i \in I_i(\eta_i) \text{ with } y_i \neq \theta_i.$$

Proof: Consider any $y_i \in I_i(\eta_i)$, then $\mathscr{E}\{\ln f_i(\mathbf{x}_i|y_i)|\theta_i\} < \infty$. Clearly $\mathscr{E}\{g_i(\mathbf{x}_i|y_i,\, \theta_i)|\theta_i\} < 0$ if $\mathscr{E}\{\ln f_i(\mathbf{x}_i|y_i)|\theta_i\} = -\infty$.

Now consider the case that $\mathscr{E}\{\ln f_i(\mathbf{x}_i|y_i)|\theta_i\} > -\infty$; then

$$(6.9)\qquad -\infty < \mathscr{E}\{g_i(\mathbf{x}_i|y_i,\, \theta_i)|\theta_i\} < \infty$$

and from (6.9) it follows that

$$(6.10)\qquad \begin{cases} \mathscr{E}\{g_i(\mathbf{x}_i \mid y_i,\, \theta_i) \mid \theta_i\} \leqq \ln \mathscr{E}\{e^{g_i(\mathbf{x}_i \mid y_i,\theta_i)} \mid \theta_i\} = \\ = \ln \displaystyle\int_{f_i(x_i \mid \theta_i) > 0} \frac{f_i(x_i \mid y_i)}{f_i(x_i \mid \theta_i)}\, dF_i(x_i \mid \theta_i) \leqq \ln 1 = 0. \end{cases}$$

Further

$$(6.11)\qquad \mathscr{E}\{g_i(\mathbf{x}_i \mid y_i,\, \theta_i) \mid \theta_i\} = \ln \mathscr{E}\{e^{g_i(\mathbf{x}_i \mid y_i,\theta_i)} \mid \theta_i\}$$

if and only if a value c exists such that

$$(6.12)\qquad \mathrm{P}\,[g_i(\mathbf{x}_i|y_i,\, \theta_i) = c|\theta_i] = 1.$$

Thus lemma III is proved if we show that such a value c does not exist.

This may be proved as follows. Suppose there exists a value c satisfying (6.12), then it follows from (6.9) that $|c| < \infty$ and further we have

$$(6.13)\qquad \mathrm{P}\,[f_i(\mathbf{x}_i|y_i) = e^c f_i(\mathbf{x}_i|\theta_i)|\theta_i] = 1.$$

From

$$(6.14)\qquad \int dF_i(x_i|y_i) = \int dF_i(x_i|\theta_i) = 1$$

it then follows that $c = 0$.

Further if

$$(6.15)\qquad \mathrm{P}\,[g_i(\mathbf{x}_i|y_i,\, \theta_i) = 0|\theta_i] = 1,$$

then it follows from (6.3) and the fact that $g_i(x_i|y_i,\, \theta_i)$ is, for each x_i, a strictly unimodal function of y_i in the interval I_i that

$$(6.16)\qquad \mathrm{P}\,[g_i(\mathbf{x}_i|y_i',\, \theta_i) > 0|\theta_i] = 1 \text{ for each } y_i' \text{ between } y_i \text{ and } \theta_i,$$

i.e.

$$(6.17)\qquad \mathrm{P}\,[f_i(\mathbf{x}_i|y_i') > f_i(\mathbf{x}_i|\theta_i)|\theta_i] = 1 \text{ for each } y_i' \text{ between } y_i \text{ and } \theta_i$$

and this is in contradiction with

$$\int dF_i(x_i|y_i') = \int dF_i(x_i|\theta_i) = 1. \tag{6.18}$$

Thus there does not exist a value c satisfying (6.12).

Now let (cf. section 4) M_ν ($\nu = 1, 2, \ldots, N$) be N subsets of the numbers $1, 2, \ldots, k$ with

$$\begin{cases} 1. & \bigcup_{\nu=1}^{N} M_\nu = \{1, 2, \ldots, k\}, \\ 2. & M_{\nu_1} \cap M_{\nu_2} \neq 0 \text{ for each pair } (\nu_1, \nu_2) \text{ with } \nu_1 \neq \nu_2, \\ 3. & \theta_i = \theta_j \text{ for each pair } i, j \in M_\nu, \text{ for any value of } \nu \end{cases} \tag{6.19}$$

and let I_{M_ν} be defined by (4.5); then $I_{M_\nu} \neq 0$ ($\nu = 1, 2, \ldots, N$). The value of θ_i for $i \in M_\nu$ will be denoted by θ_ν' ($\nu = 1, 2, \ldots, N$). From theorem I it then follows that

$$L'(z_1, z_2, \ldots, z_N) - L'(\theta_1', \theta_2', \ldots, \theta_N') = \sum_{\nu=1}^{N} \sum_{i \in M_\nu} \sum_{\gamma=1}^{n_i} g_i(x_{i,\gamma} \,|\, z_\nu, \theta_\nu') \tag{6.20}$$

possesses a unique maximum in (cf. (4.6))

$$G_N = \prod_{\nu=1}^{N} I_{M_\nu}, \tag{6.21}$$

say in the point $(z_1^*, z_2^*, \ldots, z_N^*)$. Let further

$$\eta_\nu' \stackrel{\text{def}}{=} \min_{i \in M_\nu} \eta_i \qquad (\nu = 1, 2, \ldots, N) \tag{6.22}$$

and

$$n \stackrel{\text{def}}{=} \sum_{i=1}^{k} n_i. \tag{6.23}$$

Then the following lemma holds

Lemma IV: *If*

$$\lim_{n \to \infty} n_i = \infty \quad \text{for each } i = 1, 2, \ldots, k, \tag{6.24}$$

then

$$\lim_{n \to \infty} \mathrm{P}\,[|\mathbf{z}_\nu^* - \theta_\nu'| \leq \varepsilon \text{ for each } \nu \,|\, \theta_1', \theta_2', \ldots, \theta_N'] = 1 \quad \text{for each } \varepsilon > 0 \tag{6.25}$$

for each set $M_1, M_2, \ldots, M_N$ *satisfying* (6.19) *and each* N.

Proof: Let

$$\begin{cases} 1. & \beta_i(z_\nu) \stackrel{\text{def}}{=} \mathscr{E}\{g_i(\mathbf{x}_i \,|\, z_\nu, \theta_\nu') \,|\, \theta_\nu'\} \\ 2. & \delta_i(z_\nu) \stackrel{\text{def}}{=} \sigma^2\{g_i(\mathbf{x}_i \,|\, z_\nu, \theta_\nu') \,|\, \theta_\nu'\} \end{cases} \qquad i \in M_\nu \; (\nu = 1, 2, \ldots, N) \tag{6.26}$$

and let further ε_1 be a positive number satisfying

$$\varepsilon_1 \leq \min \eta_\nu'. \tag{6.27}$$

Then

$$(6.28)\quad \begin{cases} 1.\ \ \theta'_\nu+\varepsilon_1 \in I_{M_\nu} \text{ and } \theta'_\nu-\varepsilon_1 \in I_{M_\nu} \text{ if } \theta'_\nu \text{ is an innerpoint of } I_{M_\nu}, \\ 2.\ \ \theta'_\nu+\varepsilon_1 \in I_{M_\nu} \text{ or } \theta'_\nu-\varepsilon_1 \in I_{M_\nu} \text{ if } \theta'_\nu \text{ is a borderpoint of } I_{M_\nu}. \end{cases}$$

Now let S be a subset of the numbers $1, 2, \ldots, N$ such that

$$(6.29)\quad \begin{cases} 1.\ \ \theta'_\nu+\varepsilon_1 \in I_{M_\nu} \text{ for } \nu \in S, \\ 2.\ \ \theta'_\nu+\varepsilon_1 \notin I_{M_\nu} \text{ for } \nu \notin S, \end{cases}$$

then

$$(6.30)\quad z_\nu \leqq \theta'_\nu \text{ for } \nu \notin S.$$

Further it follows from (6.6.1), for $\nu \in S$, that

$$(6.31)\quad \mathscr{E}\{\sum_{i\in M_\nu}\sum_{\gamma=1}^{n_i} g_i(\mathbf{x}_{i,\gamma}\,|\,\theta'_\nu+\varepsilon_1, \theta'_\nu)\,|\,\theta'_\nu\} = \sum_{i\in M_\nu} n_i\,\beta_i(\theta'_\nu+\varepsilon_1) < 0$$

and from (6.31) and Bienaymé's inequality then follows

$$(6.32)\quad \begin{cases} \mathrm{P}\,[\sum_{i\in M_\nu}\sum_{\gamma=1}^{n_i} g_i(\mathbf{x}_{i,\gamma}\,|\,\theta'_\nu+\varepsilon_1, \theta'_\nu) \geqq 0\,|\,\theta'_\nu] \leqq \dfrac{\sum_{i\in M_\nu} n_i\,\delta_i(\theta'_\nu+\varepsilon_1)}{[\sum_{i\in M_\nu} n_i\,\beta_i(\theta'_\nu+\varepsilon_1)]^2} = \\ = \sum_{i\in M_\nu} \dfrac{n_i\,\delta_i(\theta'_\nu+\varepsilon_1)}{[\sum_{j\in M_\nu} n_j\,\beta_j(\theta'_\nu+\varepsilon_1)]^2} \leqq \sum_{i\in M_\nu} \dfrac{n_i\,\delta_i(\theta'_\nu+\varepsilon_1)}{n_i^2[\beta_i(\theta'_\nu+\varepsilon_1)]^2} = \sum_{i\in M_\nu} \dfrac{\delta_i(\theta'_\nu+\varepsilon_1)}{n_i[\beta_i(\theta'_\nu+\varepsilon_1)]^2}. \end{cases}$$

Thus

$$(6.33)\quad \begin{cases} \mathrm{P}\,[\sum_{i\in M_\nu}\sum_{\gamma=1}^{n_i} g_i(\mathbf{x}_{i,\gamma}\,|\,\theta'_\nu+\varepsilon_1, \theta'_\nu) < 0 \text{ for each } \nu \in S\,|\,\theta'_1, \theta'_2, \ldots, \theta'_N] \geqq \\ \qquad\qquad \geqq 1 - \sum_{\nu\in S}\sum_{i\in M_\nu} \dfrac{\delta_i(\theta'_\nu+\varepsilon_1)}{n_i[\beta_i(\theta'_\nu+\varepsilon_1)]^2}. \end{cases}$$

Further it follows from (6.3), (6.30) and the fact that

$$\sum_{i\in M_\nu}\sum_{\gamma=1}^{n_i} g_i(x_{i,\gamma}\,|\,z_\nu, \theta'_\nu)$$

is a strictly unimodal function of z_ν in the interval I_{M_ν} ($\nu = 1, 2, \ldots, N$) (cf. condition (4.3)) that

$$(6.34)\quad \begin{cases} z^*_\nu \leqq \theta'_\nu+\varepsilon_1\,(\nu = 1, 2, \ldots, N) \text{ if} \\ \sum_{i\in M_\nu}\sum_{\gamma=1}^{n_i} g_i(x_{i,\gamma}\,|\,\theta'_\nu+\varepsilon_1, \theta'_\nu) < 0 \text{ for each } \nu \in S. \end{cases}$$

Thus (cf. 6.33))

$$(6.35)\quad \mathrm{P}\,[\mathbf{z}^*_\nu - \theta'_\nu \leqq \varepsilon_1 \text{ for each } \nu\,|\,\theta'_1, \theta'_2, \ldots, \theta'_N] \geqq 1 - \sum_{\nu\in S}\sum_{i\in M_\nu} \frac{\delta_i(\theta'_\nu+\varepsilon_1)}{n_i[\beta_i(\theta'_\nu+\varepsilon_1)]^2}.$$

From (6.6.2), (6.24) and (6.35) then follows

(6.36) $\lim_{n\to\infty} \mathrm{P}\,[\mathbf{z}^*_\nu - \theta'_\nu \leqq \varepsilon$ for each $\nu \mid \theta'_1, \theta'_2, \ldots, \theta'_N] = 1$ for each $\varepsilon > 0$.

In an analogous way it may be proved that

(6.37) $\lim_{n\to\infty} \mathrm{P}\,[\mathbf{z}^*_\nu - \theta'_\nu \geqq -\varepsilon$ for each $\nu \mid \theta'_1, \theta'_2, \ldots, \theta'_N] = 1$ for each $\varepsilon > 0$.

If we take $N = k$ in lemma IV then (6.25) reduces to (cf. remark 2 section 4)

(6.38) $\lim_{n\to\infty} \mathrm{P}\,[|\mathbf{v}_i - \theta_i| \leqq \varepsilon$ for each $i \mid \theta_1, \theta_2, \ldots, \theta_k] = 1$ for each $\varepsilon > 0$.

Theorem VIII: *If t_i is the maximum likelihood estimate of θ_i $(i = 1, 2, \ldots, k)$ under the restrictions $R_1, R_2, \ldots, R_s$ and if*

(6.39) $$\lim_{n\to\infty} n_i = \infty \text{ for each } i = 1, 2, \ldots, k,$$

then

(6.40) $\lim_{n\to\infty} \mathrm{P}\,[|\mathbf{t}_i - \theta_i| \leqq \varepsilon$ for each $i \mid \theta_1, \theta_2, \ldots, \theta_k] = 1$ for each $\varepsilon > 0$.

Proof: This theorem will be proved by induction.

Consider the function $L'(z_1, z_2, \ldots, z_N) - L'(\theta'_1, \theta'_2, \ldots, \theta'_N)$ (cf. (6.20)). From theorem I it follows that this function possesses a unique maximum in $D_{N,s}$ (cf. (4.9)), say in the point $(w_1^{(s)}, w_2^{(s)}, \ldots, w_N^{(s)})$.

From lemma IV then follows (for $s = 0$)

(6.41) $\lim_{n\to\infty} \mathrm{P}\,[|\mathbf{w}_\nu^{(0)} - \theta'_\nu| \leqq \varepsilon$ for each $\nu \mid \theta'_1, \theta'_2, \ldots, \theta'_N] = 1$ for each $\varepsilon > 0$

for each set $M_1, M_2, \ldots, M_N$ satisfying (6.19) and each N.

Now suppose that it has been proved that

(6.42) $\lim_{n\to\infty} \mathrm{P}\,[|\mathbf{w}_\nu^{(s)} - \theta'_\nu| \leqq \varepsilon$ for each $\nu \mid \theta'_1, \theta'_2, \ldots, \theta'_N] = 1$ for each $\varepsilon > 0$

for each $s \leqq s_0$, each set $M_1, M_2, \ldots, M_N$ satisfying (6.19) and each N. Then it will be proved that

(6.43) $\lim_{n\to\infty} \mathrm{P}\,[|\mathbf{w}_\nu^{(s_0+1)} - \theta'_\nu| \leqq \varepsilon$ for each $\nu \mid \theta'_1, \theta'_2, \ldots, \theta'_N] = 1$ for each $\varepsilon > 0$

for each set $M_1, M_2, \ldots, M_N$, satisfying (6.19) and each N.

Consider, for a given set $M_1, M_2, \ldots, M_N$ satisfying (6.19), 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: $\theta_{i_\lambda} \leqq \theta_{j_\lambda}$. Then the following two cases may be distinguished.

1. $\theta_{i_\lambda} < \theta_{j_\lambda}$; then a positive value ε_1 exists satisfying

(6.44) $$D_{N,s_0} \cap \prod_{\nu=1}^{N} I_{M_\nu}(\varepsilon_1) \subset D_{N,s_0+1}.$$

Further we have, for each ε_1 satisfying (6.44),

(6.45) $w_\nu^{(s_0+1)} = w_\nu^{(s_0)}$ $(\nu = 1, 2, \ldots, N)$ if $|w_\nu^{(s_0)} - \theta'_\nu| \leqq \varepsilon_1$ for each $\nu = 1, 2, \ldots, N$.

From (6.42) and (6.45) then follows

$$(6.46)\qquad \begin{cases} \lim_{n\to\infty} \mathrm{P}\,[|\mathbf{w}_\nu^{(s_0+1)}-\theta'_\nu| \leqq \varepsilon_1 \text{ for each } \nu \,|\, \theta'_1, \theta'_2, \dots, \theta'_N] = 1 \\ \qquad\qquad\qquad\qquad \text{for each } \varepsilon_1 \text{ satisfying (6.44)} \end{cases}$$

and from (6.46) follows

$$(6.47)\quad \lim_{n\to\infty} \mathrm{P}\,[|\mathbf{w}_\nu^{(s_0+1)}-\theta'_\nu| \leqq \varepsilon \text{ for each } \nu \,|\, \theta'_1, \theta'_2, \dots, \theta'_N] = 1 \text{ for each } \varepsilon > 0.$$

2. $\theta_{i_\lambda} = \theta_{j_\lambda}$; then we have for each $\varepsilon > 0$

$$(6.48)\qquad \begin{cases} \mathrm{P}\,[|\mathbf{w}_\nu^{(s_0+1)}-\theta'_\nu| \leqq \varepsilon \text{ for each } \nu \,|\, \theta'_1, \theta'_2, \dots, \theta'_N] = \\ = \mathrm{P}\,[\mathbf{w}_{i_\lambda}^{(s_0)} < \mathbf{w}_{j_\lambda}^{(s_0)} \,|\, \theta'_1, \theta'_2, \dots, \theta'_N]\cdot \\ \qquad \cdot \mathrm{P}\,[|\mathbf{w}_\nu^{(s_0)}-\theta'_\nu| \leqq \varepsilon \text{ for each } \nu \,|\, \mathbf{w}_{i_\lambda}^{(s_0)} < \mathbf{w}_{j_\lambda}^{(s_0)}; \theta'_1, \theta'_2, \dots, \theta'_N] \\ + \mathrm{P}\,[\mathbf{w}_{i_\lambda}^{(s_0)} \geqq \mathbf{w}_{j_\lambda}^{(s_0)} \,|\, \theta'_1, \theta'_2, \dots, \theta'_N]\cdot \\ \qquad \cdot \mathrm{P}\,[|\mathbf{w}_\nu^{(s_0+1)}-\theta'_\nu| \leqq \varepsilon \text{ for each } \nu \,|\, \mathbf{w}_{i_\lambda}^{(s_0+1)} = \mathbf{w}_{j_\lambda}^{(s_0+1)}; \theta'_1, \theta'_2, \dots, \theta'_N], \end{cases}$$

because if $w_{i_\lambda}^{(s_0)} < w_{j_\lambda}^{(s_0)}$ then the maximum under s_0 restrictions coincides with the maximum under s_0+1 restrictions and if $w_{i_\lambda}^{(s_0)} \geqq w_{j_\lambda}^{(s_0)}$ then (according to theorem II) $w_{i_\lambda}^{(s_0+1)} = w_{j_\lambda}^{(s_0+1)}$.

Further $w_1^{(s_0+1)}, w_2^{(s_0+1)}, \dots, w_N^{(s_0+1)}$ are, under the condition $w_{i_\lambda}^{(s_0+1)} = w_{j_\lambda}^{(s_0+1)}$, the values of $z_1, z_2, \dots, z_N$ which maximize $L'(z_1, z_2, \dots, z_N) - L'(\theta'_1, \theta'_2, \dots, \theta'_N)$ in a domain $D_{N', s_0'}$ where $N' = N-1$ and $s'_0 \leqq s_0 - 1$. Thus from (6.42) it follows that

$$(6.49)\qquad \begin{cases} \lim_{n\to\infty} \mathrm{P}\,[|\mathbf{w}_\nu^{(s_0+1)}-\theta'_\nu| \leqq \varepsilon \text{ for each } \nu \,|\, \mathbf{w}_{i_\lambda}^{(s_0+1)} = \mathbf{w}_{j_\lambda}^{(s_0+1)}; \theta'_1, \theta'_2, \dots, \theta'_N] = 1 \\ \qquad\qquad\qquad\qquad \text{for each } \varepsilon > 0. \end{cases}$$

Thus if

$$(6.50)\qquad P_n \stackrel{\text{def}}{=} \mathrm{P}\,[|\mathbf{w}_\nu^{(s_0+1)}-\theta'_\nu| \leqq \varepsilon \text{ for each } \nu \,|\, \theta'_1, \theta'_2, \dots, \theta'_N]$$

and if A_n, B_n and $\bar{B}_n$ respectively denote the events

$$|w_\nu^{(s_0)} - \theta'_\nu| \leqq \varepsilon \text{ for each } \nu$$

$$w_{i_\lambda}^{(s_0)} < w_{j_\lambda}^{(s_0)}$$

and

$$w_{i_\lambda}^{(s_0)} \geqq w_{j_\lambda}^{(s_0)}$$

respectively then it follows from (6.42)

$$(6.51)\qquad \lim_{n\to\infty} \mathrm{P}\,[A_n \,|\, \theta'_1, \theta'_2, \dots, \theta'_N] = 1$$

and from (6.48) and (6.49)

$$(6.52)\qquad \begin{cases} 1 \geqq \lim_{n\to\infty} P_n = \lim_{n\to\infty} \{\mathrm{P}\,[B_n \,|\, \theta'_1, \theta'_2, \dots, \theta'_N]\cdot \\ \qquad\qquad \cdot \mathrm{P}\,[A_n \,|\, B_n; \theta'_1, \theta'_2, \dots, \theta'_N] + \mathrm{P}\,[\bar{B}_n \,|\, \theta'_1, \theta'_2, \dots, \theta'_N]\} = \\ = \lim_{n\to\infty} \{\mathrm{P}\,[A_n \text{ and } B_n \,|\, \theta'_1, \theta'_2, \dots, \theta'_N] + \mathrm{P}\,[\bar{B}_n \,|\, \theta'_1, \theta'_2, \dots, \theta'_N]\} \geqq \\ \geqq \lim_{n\to\infty} \mathrm{P}\,[A_n \,|\, \theta'_1, \theta'_2, \dots, \theta'_N] = 1. \end{cases}$$

Thus

$$\lim_{n\to\infty} P_n = 1. \tag{6.53}$$

7. *Examples*

In this section some examples will be given where the conditions (4.3) and (6.6) are satisfied.

Example 1

Let $\mathbf{x}_i$ possess a normal distribution with mean θ_i and known variance σ_i^2 $(i=1, 2, \ldots, k)$. Then

$$L_i(y_i) = -\tfrac{1}{2} n_i \ln 2\pi\sigma_i^2 - \tfrac{1}{2} \frac{\sum_{\gamma=1}^{n_i} (x_{i,\gamma} - y_i)^2}{\sigma_i^2} \qquad (i=1, 2, \ldots, k). \tag{7.1}$$

From (7.1) it follows that $L_i(y_i)$ is a strictly unimodal function of y_i in the interval $(-\infty, +\infty)$ and attains its maximum in this interval for

$$y_i = m_i \overset{\text{def}}{=} \frac{1}{n_i} \sum_{\gamma=1}^{n_i} x_{i,\gamma} \quad (i=1, 2, \ldots, k). \tag{7.2}$$

Thus $L_i(y_i)$ is a strictly unimodal function of y_i in each closed subinterval I_i of the interval $(-\infty, +\infty)$ and if I_i is the interval (c_i, d_i) then $L_i(y_i)$ attains its maximum in I_i for

$$y_i = \begin{cases} m_i \text{ if } c_i \leqq m_i \leqq d_i, \\ c_i \text{ if } m_i < c_i, \\ d_i \text{ if } m_i > d_i. \end{cases} \qquad (i=1, 2, \ldots, k) \tag{7.3}$$

Further if M is a subset of the numbers $1, 2, \ldots, k$ then (cf. (4.2))

$$L_M(z) = -\tfrac{1}{2} \sum_{i\in M} \left\{ n_i \ln 2\pi\sigma_i^2 + \frac{\sum_{\gamma=1}^{n_i} (x_{i,\gamma} - z)^2}{\sigma_i^2} \right\} \tag{7.4}$$

and from (7.4) it follows easily that $L_M(z)$ is a strictly unimodal function of z in the interval $(-\infty, +\infty)$. Thus L satisfies condition (4.3).

Further $L_M(z)$ attains its maximum in the interval $(-\infty, +\infty)$ for

$$z = m_M \overset{\text{def}}{=} \left(\sum_{i\in M} \frac{n_i}{\sigma_i^2} \right)^{-1} \sum_{i\in M} \frac{n_i m_i}{\sigma_i^2}. \tag{7.5}$$

Now let M consist of the numbers $h_1, h_2, \ldots, h_\mu$, then if $\sigma_i^2 = \sigma^2$ for each $i \in M$

$$L_M(z) = -\tfrac{1}{2} n_M \ln 2\pi\sigma^2 - \tfrac{1}{2} \frac{\sum_{\gamma=1}^{n_M} (x_{M,\gamma} - z)^2}{\sigma^2}, \tag{7.6}$$

where

$$n_M \overset{\text{def}}{=} \sum_{i\in M} n_i \tag{7.7}$$

and where $x_{M,\gamma}$ $(\gamma=1, 2, \ldots, n_M)$ denote the pooled samples of $\mathbf{x}_{h_1}, \mathbf{x}_{h_2}, \ldots, \mathbf{x}_{h_\mu}$. Thus if L attains its maximum for $y_{h_1}=y_{h_2}=\ldots=y_{h_\mu}$ then the samples of $\mathbf{x}_{h_1}, \mathbf{x}_{h_1}, \ldots, \mathbf{x}_{h_\mu}$ are to be pooled if $\sigma_i^2=\sigma^2$ for each $i \in M$.

Further

$$(7.8) \qquad g_i(x_i \mid y_i, \theta_i)=\frac{(y_i-\theta_i)(2x_i-y_i-\theta_i)}{2\sigma_i^2} \quad (i=1,2,\ldots,k).$$

Thus

$$(7.9) \qquad \begin{cases} \mathscr{E}\{g_i(\mathbf{x}_i \mid y_i, \theta_i) \mid \theta_i\}= -\dfrac{(y_i-\theta_i)^2}{2\sigma_i^2} \\ \sigma^2\{g_i(\mathbf{x}_i \mid y_i, \theta_i) \mid \theta_i\}= \dfrac{(y_i-\theta_i)^2}{\sigma_i^2} \end{cases} \quad (i=1,2,\ldots,k)$$

and

$$(7.10) \qquad \frac{\sigma^2\{g_i(\mathbf{x}_i \mid y_i,\theta_i) \mid \theta_i\}}{[\mathscr{E}\{g_i(\mathbf{x}_i \mid y_i,\theta_i) \mid \theta_i\}]^2}=\frac{4\sigma_i^2}{(y_i-\theta_i)^2} \quad (i=1,2,\ldots,k).$$

From (7.9) and (7.10) it follows that condition (6.6) is satisfied if

$$(7.11) \qquad \sigma_i^2<\infty \qquad (i=1, 2, \ldots, k).$$

Remark 4: From (7.4) and (7.5) it follows that the estimates of $\theta_1, \theta_2, \ldots, \theta_k$ may also be found by means of the method described above if the σ_i^2 are unknown and σ_i^2/σ_j^2 is known for each pair of values $i, j=1, 2, \ldots, k$. Then if

$$(7.12) \qquad K_i \overset{\text{def}}{=} \frac{\sigma_i^2}{\sigma_1^2} \quad (i=1,2,\ldots,k)$$

the maximum likelihood estimate of σ_i^2 is

$$(7.13) \qquad s_i^2 \overset{\text{def}}{=} \frac{K_i}{n} \sum_{j=1}^{k} \sum_{\gamma=1}^{n_j} \frac{(x_{j,\gamma}-t_j)^2}{K_j} \quad (i=1,2,\ldots,k).$$

The procedure will now be illustrated by means of the following example.

Two preparations A and B, known to stimulate the growth of hogs, are added in two concentrations each to the food of four groups of hogs. Let these four additions be denoted by A_1, A_2, B_1 and B_2. It is known that B_1 is at least as good as A_1 (notation $A_1 \leqq B_1$) and that in the same sense $A_1 \leqq A_2$ and $B_1 \leqq B_2$. No decisive knowledge however is available concerning the ordering of A_2 and B_2. The growths of the hogs during a certain period are then the four samples.

The fictitious numerical example given below concerns this partial ordering, but has been made a little more complicated by the introduction of unequal variances and of restrictions on the possible values of each θ_i separately:

Let

(7.14)

	A_1	A_2	B_1	B_2
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		
$\sum_{\gamma=1}^{n_i} x_{i,\gamma}$	5,28	9,48	3,70	6,14
n_i	4	9	5	6
m_i	1,32	1,05	0,74	1,02
σ_i^2	2	4	5	1
I_i	$(-\infty, 1)$	$(-\infty, +\infty)$	$(\frac{1}{2}, +\infty)$	$(-\infty, +\infty)$
v_i	1	1,05	0,74	1,02

and (cf.(2.8))

(7.15) $$\begin{cases} 1. \quad r_0=2,\ r_1=4, \\ 2. \quad \alpha_{1,2}=\alpha_{1,3}=\alpha_{3,4}=1. \end{cases}$$

From (7.14) and (7.15) it follows that the pairs $i=3$, $j=2$ and $i=4$, $j=2$ satisfy (5.7) and (5.8). Thus according to theorem VI L attains its maximum in D for

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

From (7.14), (7.16) and theorem V then follows

(7.17) $$t_1 = t_3,$$

i.e. L attains its maximum in D for

(7.18) $$y_1 = y_3 \leqq y_4 \leqq y_2.$$

From (7.14), (7.18) and (7.5) then follows

(7.19)

i	1	3	4	2
$\sum_{\gamma=1}^{n_i} x_{i,\gamma}$	5,28	3,70	6,14	9,48
n_i	4	5	6	9
m_{M_ν}	1,13	1,13	1,02	1,05
σ_i^2	2	5	1	4
I_{M_ν}	$(\frac{1}{2}, 1)$	$(\frac{1}{2}, 1)$	$(-\infty, +\infty)$	$(-\infty, +\infty)$
v_{M_ν}	1	1	1,02	1,05

From (7.19) and theorem III then follows

$$t_1 = t_3 = 1, \; t_2 = 1{,}05, \; t_4 = 1{,}02. \tag{7.20}$$

Example 2. Let $\mathbf{x}_i$ possess a Poisson distribution with parameter θ_i $(0 < \theta_i < \infty;\ i = 1, 2, \ldots, k)$.
Then

$$L_i(y_i) = -n_i y_i + \sum_{\gamma=1}^{n_i} x_{i.\gamma} \ln y_i - \sum_{\gamma=1}^{n_i} \ln x_{i.\gamma}! \quad (i = 1, 2, \ldots, k); \tag{7.21}$$

thus

$$\frac{dL_i(y_i)}{dy_i} \begin{cases} > 0 \text{ for } 0 \leqq y_i < m_i \overset{\text{def}}{=} \dfrac{1}{n_i} \sum\limits_{\gamma=1}^{n_i} x_{i.\gamma}, \\ = 0 \text{ for } y_i = m_i, \\ < 0 \text{ for } y_i > m_i. \end{cases} \tag{7.22}$$

From (7.22) it follows that $L_i(y_i)$ is a strictly unimodal function of y_i in the interval $(0, \infty)$ $(i = 1, 2, \ldots, k)$.

Further if M consists of the numbers $h_1, h_2, \ldots, h_\mu$ then

$$L_M(z) = -n_M z + \sum_{\gamma=1}^{n_M} x_{M.\gamma} \ln z - \sum_{\gamma=1}^{n_M} \ln x_{M.\gamma}!, \tag{7.23}$$

where n_M is defined by (7.7) and where $x_{M.\gamma}$ $(\gamma = 1, 2, \ldots, n_M)$ denote the pooled samples of $\mathbf{x}_{h_1}, \mathbf{x}_{h_2}, \ldots, \mathbf{x}_{h_\mu}$. Thus L satisfies condition (4.3) and if L attains its maximum for $y_{h_1} = y_{h_2} = \ldots = y_{h_\mu}$ then the samples of $\mathbf{x}_{h_1}, \mathbf{x}_{h_2}, \ldots, \mathbf{x}_{h_\mu}$ are to be pooled.

Further

$$g_i(x_i \mid y_i, \theta_i) = \theta_i - y_i - x_i \ln \frac{\theta_i}{y_i} \quad (i = 1, 2, \ldots, k), \tag{7.24}$$

thus

$$\begin{cases} \mathscr{E}\{g_i(\mathbf{x}_i \mid y_i, \theta_i) \mid \theta_i\} = \theta_i - y_i - \theta_i \ln \dfrac{\theta_i}{y_i} \\ \sigma^2\{g_i(\mathbf{x}_i \mid y_i, \theta_i) \mid \theta_i\} = \theta_i \left(\ln \dfrac{\theta_i}{y_i}\right)^2 \end{cases} \quad (i = 1, 2, \ldots, k) \tag{7.25}$$

and

$$\frac{\sigma^2\{g_i(\mathbf{x}_i \mid y_i, \theta_i) \mid \theta_i\}}{[\mathscr{E}\{g_i(\mathbf{x}_i \mid y_i, \theta_i) \mid \theta_i\}]^2} = \frac{\theta_i \left(\ln \dfrac{\theta_i}{y_i}\right)^2}{\left[\theta_i - y_i - \theta_i \ln \dfrac{\theta_i}{y_i}\right]^2} \quad (i = 1, 2, \ldots, k). \tag{7.26}$$

From (7.25) and (7.26) it may easily be proved that condition (6.6) is satisfied.

A practical situation of ordered parameters of Poisson distributions might occur if several toxicants are to be investigated as to their killing power for certain kinds of bacteria. If the toxicants are added in different concentrations to cultures of bacteria, knowledge may be available leading to a partial or complete ordering of the expected values of the number of survivors in the different experiments.

It may easily be verified that the conditions (4.3) and (6.6) are e.g. also satisfied if $\boldsymbol{x}_i$ possesses

1. a normal distribution with known mean μ_i and variance θ_i $(i=1, 2, \ldots, k)$,
2. an exponential distribution with parameter θ_i $(i=1, 2, \ldots, k)$,
3. a rectangular distribution between 0 and θ_i $(i=1, 2, \ldots, k)$,
4. a normal distribution with mean θ_i and known variance for $i=l_1, l_2, \ldots, l_g$ and a Poisson distribution with parameter θ_i for $i \neq l_1, l_2, \ldots, l_g$.

Acknowledgement

The author's thanks are due to Prof. Dr. J. HEMELRIJK for his stimulating help during the investigation and to Prof. Dr. D. VAN DANTZIG for reading the paper and for his constructive criticism.

(Mathematical Centre, Amsterdam)

REFERENCE

3. WALD, A., Note on the consistency of the maximum likelihood estimate, Ann. Math. Stat. 20, 595–601 (1949).