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 \( f_i(x_i|\theta_i) > 0 \),

\[
g_i(x_i|y_i, \theta_i) \overset{\text{def}}{=} \ln \frac{f_i(x_i|y_i)}{f_i(x_i|\theta_i)} \quad (i = 1, 2, \ldots, k),
\]

then

\[
L_i(y_i) = \sum_{\gamma=1}^{n_i} g_i(x_{i,\gamma}|y_i, \theta_i) + \sum_{\gamma=1}^{n_i} \ln f_i(x_{i,\gamma}|\theta_i) \quad (i = 1, 2, \ldots, k)
\]

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

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

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

\[
0 < \eta_i \leq \min(\theta_i - c_i, d_i - \theta_i) \quad \text{if } \theta_i \text{ is an innerpoint of } I_i,
\]

\[
0 < \eta_i \leq d_i - c_i \quad \text{if } \theta_i \text{ is a borderpoint of } I_i.
\]

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

\[
|y_i - \theta_i| \leq \eta_i \quad (i = 1, 2, \ldots, k).
\]

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

\[
1. \quad \mathbb{E} \{ g_i(x_i|y_i, \theta_i)|\theta_i \} < 0, \quad \text{for each } y_i \in I_i(\eta_i) \text{ with } y_i \neq \theta_i
\]

\[
2. \quad \frac{\sigma^2 \{ g_i(x_i|y_i, \theta_i)|\theta_i \}}{(\mathbb{E} \{ g_i(x_i|y_i, \theta_i)|\theta_i \})^2} < \infty \quad (i = 1, 2, \ldots, k).
\]
Some of the conditions mentioned in WALD's paper [3] are in our

Lemma III: If condition (4.3) is satisfied, if \( \eta_i \) satisfies (6.4) and if

\[
\begin{align*}
E \{ \ln f_i(x_i | y_i) | \theta_i \} < \infty \text{ for each } y_i \in I_i(\eta_i) \text{ with } y_i \neq \theta_i, \\
\end{align*}
\]

(6.7)

\[
\begin{align*}
- \infty < E \{ \ln f_i(x_i | \theta_i) | \theta_i \} < \infty
\end{align*}
\]

then

\[
E \{ g_i(x_i | y_i, \theta_i) | \theta_i \} < 0 \text{ for each } y_i \in I_i(\eta_i) \text{ with } y_i \neq \theta_i.
\]

(6.8)

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

Now consider the case that \( E \{ \ln f_i(x_i | y_i) | \theta_i \} > - \infty \); then

\[
- \infty < E \{ g_i(x_i | y_i, \theta_i) | \theta_i \} < \infty
\]

(6.9)

and from (6.9) it follows that

\[
E \{ g_i(x_i | y_i, \theta_i) | \theta_i \} \leq \ln E \{ e^{g_i(x_i | y_i, \theta_i)} | \theta_i \} = \ln \int_{I_i(x_i | \theta_i) > 0} f_i(x_i | y_i) dF_i(x_i | \theta_i) \leq \ln 1 = 0.
\]

(6.10)

Further

\[
E \{ g_i(x_i | y_i, \theta_i) | \theta_i \} = \ln E \{ e^{g_i(x_i | y_i, \theta_i)} | \theta_i \}
\]

(6.11)

if and only if a value \( c \) exists such that

\[
P \{ g_i(x_i | y_i, \theta_i) = c | \theta_i \} = 1.
\]

(6.12)

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

\[
P \{ f_i(x_i | y_i) = e^c f_i(x_i | \theta_i) | \theta_i \} = 1.
\]

(6.13)

From

\[
\int dF_i(x_i | y_i) = \int dF_i(x_i | \theta_i) = 1
\]

(6.14)

it then follows that \( c = 0 \).

Further if

\[
P \{ g_i(x_i | y_i, \theta_i) = 0 | \theta_i \} = 1,
\]

(6.15)

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

\[
P \{ g_i(x_i | y'_i, \theta_i) > 0 | \theta_i \} = 1 \text{ for each } y'_i \text{ between } y_i \text{ and } \theta_i,
\]

(6.16)

i.e.

\[
P \{ f_i(x_i | y'_i) > f_i(x_i | \theta_i) | \theta_i \} = 1 \text{ for each } y'_i \text{ between } y_i \text{ and } \theta_i.
\]

(6.17)
and this is in contradiction with
\[(6.18) \quad \int dF_i(x_i|y_i) = \int dF_i(x_i|\theta_i) = 1.\]
Thus there does not exist a value \(c\) satisfying (6.12).

Now let (cf. section 4) \(M_\nu\) \((\nu = 1, 2, \ldots, N)\) be \(N\) subsets of the numbers 1, 2, \ldots, \(k\) with
\[(6.19) \quad \begin{cases} 1. \quad \bigcup_{\nu=1}^{N} M_\nu = \{1, 2, \ldots, k\}, \\ 2. \quad M_\nu \cap M_{\nu'} \neq \emptyset \text{ for each pair } (\nu, \nu') \text{ with } \nu \neq \nu', \\ 3. \quad \theta_i = \theta_i' \text{ for each pair } i, j \in M_\nu \text{ for any value of } \nu \end{cases}\]
and let \(I_{M_\nu}\) be defined by (4.5); then \(I_{M_\nu} \neq 0 \quad (\nu = 1, 2, \ldots, N)\).
The value of \(\theta_i\) for \(i \in M_\nu\) will be denoted by \(\theta_i' \quad (\nu = 1, 2, \ldots, N)\).
From theorem I it then follows that
\[(6.20) \quad L' (z_1, z_2, \ldots, z_N) - L' (\theta_1', \theta_2', \ldots, \theta_N') = \sum_{\nu=1}^{N} \sum_{i \in M_\nu} \sum_{\nu=1}^{n_\nu} g_i (x_i, | z_\nu, \theta_i') \]
possesses a unique maximum in (cf. (4.6))
\[(6.21) \quad G_N = \prod_{\nu=1}^{N} I_{M_\nu}, \]
say in the point \((z_1^*, z_2^*, \ldots, z_N^*)\). Let further
\[(6.22) \quad \eta_\nu' \overset{\text{def}}{=} \min_{i \in M_\nu} \eta_i \quad (\nu = 1, 2, \ldots, N) \]
and
\[(6.23) \quad n \overset{\text{def}}{=} \sum_{i=1}^{N} n_i. \]

Then the following lemma holds

**Lemma IV:** If
\[(6.24) \quad \lim_{n \to \infty} n_i = \infty \quad \text{for each } i = 1, 2, \ldots, k, \]
then
\[(6.25) \quad \lim_{n \to \infty} P [ | \mathbf{z}^* - \mathbf{\theta}_* | \leq \varepsilon \quad \text{for each } \nu | \theta_1', \theta_2', \ldots, \theta_N' ] = 1 \quad \text{for each } \varepsilon > 0 \]
for each set \(M_1, M_2, \ldots, M_N\) satisfying (6.19) and each \(N\).

**Proof:** Let
\[(6.26) \quad \begin{cases} 1. \quad \beta_i (z_i) \overset{\text{def}}{=} \sigma [ g_i (x_i, | z_\nu, \theta_i') | \theta_i' ] \\ 2. \quad \delta_i (z_i) \overset{\text{def}}{=} \sigma^2 [ g_i (x_i, | z_\nu, \theta_i') | \theta_i' ] \end{cases} \quad i \in M_\nu \quad (\nu = 1, 2, \ldots, N) \]
and let further \(\varepsilon_1\) be a positive number satisfying
\[(6.27) \quad \varepsilon_1 \leq \min \eta_\nu'. \]
Then

\begin{equation}
1. \theta' + e_1 \in I_{M_p} \text{ and } \theta' - e_1 \in I_{M_p} \text{ if } \theta' \text{ is an innerpoint of } I_{M_p},
\end{equation}

\begin{equation}
2. \theta' + e_1 \in I_{M_p} \text{ or } \theta' - e_1 \in I_{M_p} \text{ if } \theta' \text{ is a borderpoint of } I_{M_p}.
\end{equation}

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

\begin{equation}
1. \theta' + e_1 \in I_{M_p} \text{ for } \nu \in S,
\end{equation}

\begin{equation}
2. \theta' + e_1 \notin I_{M_p} \text{ for } \nu \notin S,
\end{equation}

then

\begin{equation}
z_\nu \leq \theta' \text{ for } \nu \notin S.
\end{equation}

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

\begin{equation}
\mathcal{E} \left\{ \sum_{i \in M_p} \sum_{\nu=1}^{n_i} g_i(x_{i,\nu} | \theta' + e_1, \theta'_0) \right\} = \sum_{i \in M_p} n_i \beta_i (\theta'_0 + e_1) < 0
\end{equation}

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

\begin{equation}
P \left\{ \sum_{i \in M_p} \sum_{\nu=1}^{n_i} g_i(x_{i,\nu} | \theta' + e_1, \theta'_0) \geq 0 | \theta'_0 \right\} \leq \frac{ \sum_{i \in M_p} n_i \beta_i (\theta'_0 + e_1) }{ \left[ \sum_{i \in M_p} n_i \beta_i (\theta'_0 + e_1) \right]^2 } ^{1/2} = \frac{ \sum_{i \in M_p} n_i \beta_i (\theta'_0 + e_1) }{ \sum_{i \in M_p} n_i \beta_i (\theta'_0 + e_1) } ^{1/2}.
\end{equation}

Thus

\begin{equation}
P \left\{ \sum_{i \in M_p} \sum_{\nu=1}^{n_i} g_i(x_{i,\nu} | \theta' + e_1, \theta'_0) < 0 \text{ for each } \nu \in S \right\} \geq 1 - \sum_{\nu \in S} \sum_{i \in M_p} n_i \beta_i (\theta'_0 + e_1) ^{1/2}.
\end{equation}

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

\begin{equation}
\sum_{i \in M_p} \sum_{\nu=1}^{n_i} g_i(x_{i,\nu} | z_\nu, \theta'_0)
\end{equation}

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

\begin{equation}
z_\nu^* \leq \theta' + e_1 \text{ (} \nu = 1, 2, \ldots, N \text{) if}
\end{equation}

\begin{equation}
\sum_{i \in M_p} \sum_{\nu=1}^{n_i} g_i(x_{i,\nu} | \theta' + e_1, \theta'_0) < 0 \text{ for each } \nu \in S.
\end{equation}

Thus (cf. 6.33))

\begin{equation}
P \left\{ z_\nu^* - \theta'_0 \leq e_1 \text{ for each } \nu \right\} \geq 1 - \sum_{\nu \in S} \sum_{i \in M_p} \frac{ \delta_i (\theta'_0 + e_1) }{ \left[ \sum_{i \in M_p} n_i \beta_i (\theta'_0 + e_1) \right] ^{1/2} }.
\end{equation}
From (6.6.2), (6.24) and (6.35) then follows

\[ \lim_{n \to \infty} P [z_i^* - \theta_i' \leq \varepsilon \text{ for each } v | \theta_1', \theta_2', \ldots, \theta_N'] = 1 \text{ for each } \varepsilon > 0. \] (6.36)

In an analogous way it may be proved that

\[ \lim_{n \to \infty} P [z_i^* - \theta_i' \geq -\varepsilon \text{ for each } v | \theta_1', \theta_2', \ldots, \theta_N'] = 1 \text{ for each } \varepsilon > 0. \] (6.37)

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

\[ \lim_{n \to \infty} [\mathbf{v}_i - \mathbf{t}_i] \leq \varepsilon \text{ for each } i | \theta_1, \theta_2, \ldots, \theta_k = 1 \text{ for each } \varepsilon > 0. \] (6.38)

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

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

then

\[ \lim_{n \to \infty} [\mathbf{w}_i^0 - \theta_i'] \leq \varepsilon \text{ for each } i | \theta_1, \theta_2, \ldots, \theta_N' = 1 \text{ for each } \varepsilon > 0. \] (6.40)

**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 \( (w_1^0, w_2^0, \ldots, w_N^0) \).

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

\[ \lim_{n \to \infty} [\mathbf{w}_i^0 - \theta_i'] \leq \varepsilon \text{ for each } v | \theta_1', \theta_2', \ldots, \theta_N' = 1 \text{ for each } \varepsilon > 0 \] (6.41)

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

Now suppose that it has been proved that

\[ \lim_{n \to \infty} [\mathbf{w}_i^{(s)} - \theta_i'] \leq \varepsilon \text{ for each } v | \theta_1', \theta_2', \ldots, \theta_N' = 1 \text{ for each } \varepsilon > 0 \] (6.42)

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

\[ \lim_{n \to \infty} [\mathbf{w}_i^{(s+1)} - \theta_i'] \leq \varepsilon \text{ for each } v | \theta_1', \theta_2', \ldots, \theta_N' = 1 \text{ for each } \varepsilon > 0 \] (6.43)

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: \( \theta_{i_1} \leq \theta_{i_2} \).

Then the following two cases may be distinguished.

1. \( \theta_{i_1} < \theta_{i_2} \); then a positive value \( \varepsilon_1 \) exists satisfying

\[ D_{N,s_0} \cap \prod_{r=1}^{N} I_{M_r} (\varepsilon_1) \subset D_{N,s_0+1}. \] (6.44)

Further we have, for each \( \varepsilon_1 \) satisfying (6.44),

\[ w_{(s_0+1)}^v = w_{(s_0)}^v (v = 1, 2, \ldots, N) \text{ if } |w_{(s_0)}^v - \theta_i'| \leq \varepsilon_1 \text{ for each } v = 1, 2, \ldots, N. \] (6.45)
From (6.42) and (6.45) then follows

\[
\lim_{n \to \infty} P \left[ \left| w_{j_a}^{(s_0 + 1)} - \theta_j' \right| \leq \varepsilon \right] \text{ for each } \nu \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right| = 1
\]

(6.46)

and from (6.46) follows

\[
\lim_{n \to \infty} P \left[ \left| w_{j_a}^{(s_0 + 1)} - \theta_j' \right| \leq \varepsilon \right] \text{ for each } \nu \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right| = 1 \text{ for each } \varepsilon > 0.
\]

(6.47)

2. \( \theta_j = \theta_j' \); then we have for each \( \varepsilon > 0 \)

\[
P \left[ \left| w_{j_a}^{(s_0 + 1)} - \theta_j' \right| \leq \varepsilon \right] \text{ for each } \nu \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right|
\]

\[
= P \left[ w_{j_a}^{(s_0)} < w_{j_a}^{(s_0)} \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right| \right]
\]

\[
+ P \left[ w_{j_a}^{(s_0)} \geq w_{j_a}^{(s_0)} \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right| \right]
\]

\[
\cdot \left| w_{j_a}^{(s_0 + 1)} - \theta_j' \right| \leq \varepsilon \text{ for each } \nu \left| w_{j_a}^{(s_0 + 1)} = w_{j_a}^{(s_0 + 1)} ; \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right|
\]

(6.48)

because if \( w_{j_a}^{(s_0)} < w_{j_a}^{(s_0)} \) then the maximum under \( s_0 \) restrictions coincides with the maximum under \( s_0 + 1 \) restrictions and if \( w_{j_a}^{(s_0)} \geq w_{j_a}^{(s_0)} \) then (according to theorem II) \( w_{j_a}^{(s_0 + 1)} = w_{j_a}^{(s_0 + 1)} \).

Further, \( w_{j_a}^{(s_0 + 1)} ; w_{j_2}^{(s_0 + 1)} ; \ldots ; w_{j_N}^{(s_0 + 1)} \) are, under the condition \( w_{j_a}^{(s_0 + 1)} = w_{j_a}^{(s_0 + 1)} \), the values of \( z_1, z_2, \ldots, z_N \) which maximize \( L'(z_1, z_2, \ldots, z_N) - L'(\theta_j', \theta_{j_2}', \ldots, \theta_{j_N}') \) in a domain \( D_{N', s_0} \) where \( N' = N - 1 \) and \( s_0 \leq s_0 - 1 \). Thus from (6.42) it follows that

\[
\lim_{n \to \infty} P \left[ \left| w_{j_a}^{(s_0 + 1)} - \theta_j' \right| \leq \varepsilon \right] \text{ for each } \nu \left| w_{j_a}^{(s_0 + 1)} = w_{j_a}^{(s_0 + 1)} ; \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right] = 1
\]

(6.49)

Thus if

\[
P_n \overset{\text{def}}{=} P \left[ \left| w_{j_a}^{(s_0 + 1)} - \theta_j' \right| \leq \varepsilon \right] \text{ for each } \nu \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right]
\]

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

\[
\left| w_{j_a}^{(s_0)} - \theta_j' \right| \leq \varepsilon \text{ for each } \nu
\]

\[
w_{j_a}^{(s_0)} < w_{j_a}^{(s_0)}
\]

and

\[
w_{j_a}^{(s_0)} \geq w_{j_a}^{(s_0)}
\]

respectively then it follows from (6.42)

\[
\lim_{n \to \infty} P \left[ A_n \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right] = 1
\]

(6.51)

and from (6.48) and (6.49)

\[
1 \geq \lim_{n \to \infty} P_n = \lim_{n \to \infty} \left\{ P \left[ B_n \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right] \right] \right\}
\]

\[
\left\{ \cdot P \left[ A_n \left| B_n \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right] \right] + P \left[ \bar{B}_n \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right] \right] \right\}
\]

(6.52)

\[
= \lim_{n \to \infty} \left\{ P \left[ A_n \right. \text{ and } B_n \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right] \right. + P \left[ \bar{B}_n \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right] \right) \}
\]

\[
\geq \lim_{n \to \infty} P \left[ A_n \left| \theta_j', \theta_{j_2}', \ldots, \theta_{j_N}' \right] = 1.
\]
Thus

\[ (6.53) \quad \lim_{n \to \infty} P_n = 1. \]

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 \( \theta_i \) and known variance \( \sigma_i^2 \) \( (i = 1, 2, \ldots, k) \). Then

\[ (7.1) \quad L_i(y_i) = -\frac{1}{2} n_i \ln 2\pi \sigma_i^2 - \frac{\sum_{y=1}^{n_i} (x_{i,y} - y_i)^2}{\sigma_i^2} \quad (i = 1, 2, \ldots, k). \]

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

\[ (7.2) \quad y_i = m_i = \frac{1}{n_i} \sum_{y=1}^{n_i} x_{i,y} \quad (i = 1, 2, \ldots, k). \]

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

\[ (7.3) \quad y_i = \begin{cases} m_i & \text{if } c_i \leq m_i \leq d_i, \\ c_i & \text{if } m_i < c_i, \\ d_i & \text{if } m_i > d_i. \end{cases} \]

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

\[ (7.4) \quad L_M(z) = -\frac{1}{2} \sum_{i \in M} \left( n_i \ln 2\pi \sigma_i^2 + \frac{\sum_{y=1}^{n_i} (x_{i,y} - z)^2}{\sigma_i^2} \right) \]

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

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

Now let \( M \) consist of the numbers \( h_1, h_2, \ldots, h_m \), then if \( \sigma_i^2 = \sigma^2 \) for each \( i \in M \)

\[ (7.6) \quad L_M(z) = -\frac{1}{2} n_M \ln 2\pi \sigma^2 - \frac{\sum_{y=1}^{n_M} (x_{M,y} - z)^2}{\sigma^2}, \]

where

\[ (7.7) \quad n_M \overset{\text{def}}{=} \sum_{i \in M} n_i \]

Theorem 1. If \( n \to \infty \) and \( M \to \Omega \), then

\[ (7.8) \quad P_M = \lim_{n \to \infty} P_M. \]
and where \( x_{M, \gamma} (\gamma = 1, 2, \ldots, n_M) \) denote the pooled samples of \( x_{11}, x_{21}, \ldots, x_{n_M} \). Thus if \( L \) attains its maximum for \( y_{h_1} = y_{h_2} = \ldots = y_{h_\mu} \) then the samples of \( x_{h_1}, x_{h_1}, \ldots, x_{h_\mu} \) are to be pooled if \( \sigma_i^2 = \sigma^2 \) for each \( i \in M \).

Further

\[
(7.8) \quad g_i (x_i | y_i, \theta_i) = \frac{(y_i - \theta_i)(2x_i - y_i - \theta_i)}{2\sigma_i^2} \quad (i = 1, 2, \ldots, k).
\]

Thus

\[
(7.9) \quad \begin{cases}
\mathcal{E} \{g_i (x_i | y_i, \theta_i) | \theta_i \} = -\frac{(y_i - \theta_i)^2}{2\sigma_i^2} \\
\sigma^2 \{g_i (x_i | y_i, \theta_i) | \theta_i \} = \frac{(y_i - \theta_i)^2}{\sigma_i^2}
\end{cases} \quad (i = 1, 2, \ldots, k)
\]

and

\[
(7.10) \quad \frac{\sigma^2 \{g_i (x_i | y_i, \theta_i) | \theta_i \}}{[\mathcal{E} \{g_i (x_i | y_i, \theta_i) | \theta_i \}]^2} = \frac{4\sigma_i^2}{(y_i - \theta_i)^2} \quad (i = 1, 2, \ldots, k).
\]

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

\[
(7.11) \quad \sigma_i^2 < \infty \quad (i = 1, 2, \ldots, k).
\]

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 \( \sigma_i^2 \) are unknown and \( \sigma_k^2 / \sigma_j^2 \) is known for each pair of values \( i, j = 1, 2, \ldots, k \). Then if

\[
(7.12) \quad K_i \overset{\text{def}}{=} \frac{\sigma_i^2}{\sigma_j^2} \quad (i = 1, 2, \ldots, k)
\]

the maximum likelihood estimate of \( \sigma_i^2 \) is

\[
(7.13) \quad \hat{\sigma}_i^2 = \frac{K_i}{n} \sum_{j=1}^{k} \sum_{\gamma=1}^{n_j} \frac{(x_{i, \gamma} - \theta_i)^2}{K_j} \quad (i = 1, 2, \ldots, k).
\]

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 \( B_2 \) is at least as good as \( A_1 \) (notation \( A_1 \preceq B_2 \)) and that in the same sense \( A_1 \preceq A_2 \) and \( B_1 \preceq B_2 \). 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 \( \theta_i \) separately:
Let

\[
\begin{array}{|c|c|c|c|c|}
\hline
i & A_1 & A_2 & B_1 & B_2 \\
\hline
1 & -0,40 & 1,43 & -0,70 & 0,29 \\
2 & 2,56 & 1,86 & 2,61 & 0 \\
3 & 0,25 & 0,06 & 0,79 & 1,31 \\
4 & 2,87 & 0,07 & 0,86 & 0,15 \\
\hline
\end{array}
\]

\[x_{i,\gamma} \]

\[
\begin{array}{c}
\begin{array}{c}
\sum_{\gamma=1}^{n_i} x_{i,\gamma} \\
n_i \\
m_i \\
\sigma_i^2 \\
I_i \\
v_i \\
\end{array} \\
\begin{array}{c}
5,28 \\
4 \\
1,32 \\
2 \\
(-\infty, 1) \\
1 \\
\end{array} \\
\begin{array}{c}
9,48 \\
9 \\
1,05 \\
4 \\
(-\infty, +\infty) \\
1,05 \\
\end{array} \\
\begin{array}{c}
3,70 \\
5 \\
0,74 \\
5 \\
(\frac{1}{2}, +\infty) \\
0,74 \\
\end{array} \\
\begin{array}{c}
6,14 \\
6 \\
1,02 \\
1 \\
(-\infty, +\infty) \\
1,02 \\
\end{array} \\
\end{array}
\]

(7.14)

and (cf.(2.8))

\[
\begin{cases}
1. & r_0 = 2, r_1 = 4, \\
2. & \alpha_{1,2} = \alpha_{1,3} = \alpha_{3,4} = 1.
\end{cases}
\]

(7.15)

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

\[
y_1 \leq y_3 \leq y_4 \leq y_2.
\]

(7.16)

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

\[
t_1 = t_3,
\]

i.e. \(L\) attains its maximum in \(D\) for

\[
y_1 = y_3 \leq y_4 \leq y_2.
\]

(7.18)

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

\[
\begin{array}{|c|c|c|c|c|}
\hline
i & 1 & 3 & 4 & 2 \\
\hline
\sum_{\gamma=1}^{n_i} x_{i,\gamma} & 5,28 & 3,70 & 6,14 & 9,48 \\
n_i & 4 & 5 & 6 & 9 \\
m_{M_i} & 1,13 & 1,13 & 1,02 & 1,05 \\
\sigma_i^2 & 2 & 5 & 1 & 4 \\
I_{M_i} & (\frac{1}{2}, 1) & (\frac{1}{2}, 1) & (-\infty, +\infty) & (-\infty, +\infty) \\
v_{M_i} & 1 & 1 & 1,02 & 1,05 \\
\end{array}
\]

(7.19)
From (7.19) and theorem III then follows

(7.20) \[ t_1 = t_3 = 1, \ t_2 = 1.05, \ t_4 = 1.02. \]

**Example 2.** Let \( x_i \) possess a Poisson distribution with parameter \( \theta_i \) \((0 < \theta_i < \infty; \ i = 1, 2, \ldots, k)\).

Then

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

thus

(7.22) \[
\frac{dL_i(y_i)}{dy_i} \begin{cases} > 0 & \text{for } 0 \leq y_i < m_i \overset{\text{def}}{=} \frac{1}{n_i} \sum_{\gamma=1}^{n_i} x_{i,\gamma}, \\ = 0 & \text{for } y_i = m_i, \\ < 0 & \text{for } y_i > m_i. 
\end{cases}
\]

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

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

where \( n_M \) is defined by (7.7) and where \( x_{M,\gamma} \) \((\gamma = 1, 2, \ldots, n_M)\) denote the pooled samples of \( x_{h_1}, x_{h_2}, \ldots, x_{h_\mu} \). 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 \( x_{h_1}, x_{h_2}, \ldots, x_{h_\mu} \) are to be pooled.

Further

(7.24) \[ g_i(x_i | y_i, \theta_i) = \theta_i - y_i - x_i \ln \frac{\theta_i}{y_i} \quad (i = 1, 2, \ldots, k), \]

thus

(7.25) \[
\begin{align*}
\sigma^2 \{ g_i(x_i | y_i, \theta_i) | \theta_i \} &= \theta_i - y_i - \theta_i \ln \frac{\theta_i}{y_i} \\
\theta_i \left( \ln \frac{\theta_i}{y_i} \right)^2 &\quad (i = 1, 2, \ldots, k)
\end{align*}
\]

and

(7.26) \[
\frac{\sigma^2 \{ g_i(x_i | y_i, \theta_i) | \theta_i \}}{[\sigma^2 \{ g_i(x_i | y_i, \theta_i) | \theta_i \}]^2} = \frac{\theta_i \left( \ln \frac{\theta_i}{y_i} \right)^2}{\theta_i - y_i - \theta_i \ln \frac{\theta_i}{y_i}} \quad (i = 1, 2, \ldots, k).
\]

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 \( x_i \) possesses

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