Ann. Inst. Statist. Math.
Vol. 45, No. 4, 741–758 (1993)

ON THE EXISTENCE OF MINIMUM CONTRAST ESTIMATES IN BINARY RESPONSE MODEL

TADASHI NAKAMURA[1] AND CHAE-SHIN LEE[2]

[1] *Department of Information Science, Shimane University, 1060 Nishikawatsu-cho, Matsue, Shimane 690, Japan*
[2] *Graduate School of Natural Science and Technology, Okayama University, 1-1-1 Tsushima-Naka, Okayama 700, Japan*

(Received August 8, 1991; revised November 11, 1992)

Abstract. When random samples are drawn from a 3-parameter distribution with a shifted origin and the observations corresponding to each sample are binary, criteria for the existence of minimum contrast estimates are given. These criteria can be drived by a method, called the probability contents boundary analysis. The probabilities of the existence of maximum likelihood estimates and least squares estimates are evaluated, by simulation with 1000 replications, in the case where the underlying distribution is a 3-parameter lognormal distribution or a 3-parameter loglogistic distribution.

Key words and phrases: Binary response data, biological assay, dose response curve model, minimum contrast estimate, maximum likelihood estimate, least squares estimate, three-parameter distribution, slowly varying function.

1. Introduction

Suppose that individuals exposed to a level, x, of a stimulus are observed for the occurrence of a certain response. Such observations are called binary response data. Important fields of application are: (i) Bioassay, where, different level of stimulus may represent different doses of a toxine, and the binary response is death or survival (see Finney (1971)); (ii) Analysis of survival data, where, different level of stimulus may represent different elapsed time, and the binary response is failure or survival (see Elandt-Johnson and Johnson (1980)). Similar situations arise in many other fields (see Mann *et al.* (1974), Cox and Snell (1989)).

It is usual to model the relation between the level x of stimulus and the probability $P(x)$ of a response as $H(P(x)) = \alpha h(x) + \beta$, where the parameters α and β are to be estimated. The stimulus metameter $h(x)$ is usually taken to be a logarithmic function of x. Some of typical choices of the response metameter (or linearizing transformation) $H(z)$ that have been used in the literature are given below:

(i) The logit transformation $H(z) = \log \text{probit}(z)$;

(ii) The probit transformation $H(z) = \Phi^{-1}(z)$, where Φ is the standard normal distribution function (d.f.);

(iii) The log-log transformation $H(z) = -\log(-\log z)$;

(iv) The complementary log-log transformation $H(z) = \log(-\log(1-z))$;

(v) The angular transformation $H(z) = \sin^{-1}\sqrt{z}$.

Similar transformations called the link function are found in the generalized linear model (see McCullagh and Nelder (1989)).

We shall introduce a general relationship (a 3-parameter model) $H(P(x)) = \alpha h(x-\lambda) - \beta$, where the parameters α, β and λ are to be estimated, and the stimulus metameter $h(x)$ and the response metameter $H(z)$ are known and $H(z)$ has the inverse function $H^{-1}(u)$. The parameter λ is called a shifted origin (or threshold parameter). The relation $\Phi^{-1}(P(x)) = \alpha\log(x-\lambda) - \beta$ was found in Finney (1971), in which, however, only the case $\lambda = 0$ was treated. This model implies that $P(x)$ is a 3-parameter log-normal d.f. with the shifted origin λ. Another generalization is found in Prentice (1976). The maximum likelihood estimation is adopted in many literature for estimating the true parameter. Simultaneously, we are faced with the problem whether the maximum likelihood estimate (MLE) based on a binary response data exists or not. As for the maximum likelihood estimation based on the interval-censored data from a 3-parameter distribution with a shifted origin, there are two useful methods for deriving criteria for the existence of MLE's. One is the method of using the Brower fixed point theorem (see Cheng and Amin (1982)). This gives an asymptotic criterion. Another is the method of using the continuous topological extension of the log-likelihood to a compact set including the parameter space (see Nakamura (1991)). This gives a criterion for every finite size of sample. However, criteria given by Cheng and Amin (1982) and by Nakamura (1991) can not be applied to the binary response case.

The purpose of this paper is to give criteria for the existence of minimum contrast estimates based on the binary response data when $P(x) = H^{-1}(\alpha h(x-\lambda) - \beta)$, that is, $P(x)$ is a 3-parameter d.f. with a shifted origin and when $h(s)$ is a slowly varying function at infinity. In Section 2, a minimum contrast estimate (MCE) is defined. The MLE and the least squares estimate (LSE) are special types of the MCE. A family $\mathcal{F}$, in which $P(x)$ is supposed to be, of 3-parameter distributions is considered. This family $\mathcal{F}$ covers the families adopted in Cheng and Amin (1982) and in Nakamura (1991) as a special case. An approach for finding a criterion for the existence of MCE's is discussed. In Section 3, criteria for a 2-parameter subfamily of the 3-parameter family $\mathcal{F}$ and for the 3-parameter family $\mathcal{F}$ are derived. In Section 4, various kinds of practical criteria are given. In Section 5, the simulation study is designed to evaluate the probabilities of the existence of MLE's and LSE's in the case where the family $\mathcal{F}$ is the 3-parameter lognormal family or the 3-parameter loglogistic family.

2. Mathematical formulation and approach

Let $F(x)$ $(\equiv H^{-1}(x))$ be a continuously differentiable distribution function on the real line $\mathcal{R}$ with a positive density function $f(x)$, let $h(s)$ be a strictly increasing and continuously differentiable function on the interval $\mathcal{R}_+ = (0, \infty)$ such that $h(\mathcal{R}_+) = \mathcal{R}$ and define $t(x, \theta)$ $(\theta = (\alpha, \beta, \lambda) \in \mathcal{R}_+ \times \mathcal{R} \times \mathcal{R})$ by $t(x, \theta) = \alpha h(x - \lambda) - \beta$, if $x > \lambda$ and $t(x, \theta) = -\infty$ if $x \leq \lambda$. Throughout this paper, it is assumed that $h(s)$ is a $((\phi_1, h_1), (\phi_2, h_2), \phi)$-slowly varying function at infinity, that is, it is real-valued on $[s_0, \infty)$, for some $s_0 > 0$, and the following conditions are satisfied:

(i) $\phi_1(s)$, $\phi_2(s)$ and $\phi(s)$ are non-constant and positive functions on $[s_0, \infty)$.
(ii) $\phi_2(s) = o(\phi_1(s))$ $(s \to \infty)$ and $\phi(s) = o(\phi_2(s))$ $(s \to \infty)$.
(iii) $h_1(x)$ and $h_2(x)$ are non-constant function on $\mathcal{R}$.
(iv) For each $x \in \mathcal{R}$,

$$h(x + s) = h(s) + h_1(x)\phi_1(s) + h_2(x)\phi_2(s) + O(\phi(s)) \qquad (s \to \infty).$$

For detailed discussion on the slowly varying functions, see Ash *et al.* (1974), Senata (1976) and Hirai *et al.* (1992).

We shall give examples of the slowly varying function $h(s)$.

Example 2.1.

(i) Consider the case $h(s) = \log s$. By Taylor's expansion, we see that $h(s)$ is $((\phi_1, h_1), (\phi_2, h_2), \phi)$-slowly varying function at infinity, where $h_1(x) = x$, $h_2(x) = -x^2/2$, $\phi_1(s) = s^{-1}$, $\phi_2(s) = s^{-2}$ and $\phi(s) = s^{-3}$.

(ii) Consider the case $h(s) = s^r - s^{-r}$, where $0 < r < 1$. By Taylor's expansion, we see that $h(s)$ is $((\phi_1, h_1), (\phi_2, h_2), \phi)$-slowly varying function at infinity, where in case $0 < r < 1/2$,

$$h_1(x) = x, \quad h_2(x) = rx, \quad \phi_1(s) = rs^{r-1}, \quad \phi_2(s) = s^{-r-1} \quad \text{and} \quad \phi(s) = s^{r-2};$$

in case $r = 1/2$,

$$h_1(x) = x, \quad h_2(x) = x/2 - x^2/8,$$
$$\phi_1(s) = s^{-1/2}/2, \quad \phi_2(s) = s^{-2/3} \quad \text{and} \quad \phi(s) = s^{-5/2};$$

in case $1/2 < r < 1$,

$$h_1(x) = x, \quad h_2(x) = r(r-1)x/2,$$
$$\phi_1(s) = rs^{r-1}, \quad \phi_2(s) = s^{r-2} \quad \text{and} \quad \phi(s) = s^{-r-1}.$$

Especially $h(s) = \sqrt{s} - 1/\sqrt{s}$ is called the Birnbaum-Saunders transformation (see Mann *et al.* (1974)).

Let $(X_{i1}, \ldots, X_{in_i})$ $(1 \leq i \leq N)$ be a random sample from the unknown distribution $P(x) \equiv F(t(x, \theta_0)) \in \mathcal{F} \equiv \{F(t(x, \theta)); \theta \in \mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}\}$ and suppose that information available for each X_{ij} is only that its value lies in a proper subinterval $\mathcal{C}_{ij}$ of $\mathcal{R}$ with nonempty interior. The collection $\mathcal{C} \equiv \{\mathcal{C}_{ij}; 1 \leq i \leq N, 1 \leq j \leq n_i\}$ is called a pooled interval-censored (p.i.c.) data. When $N = 1$, the p.i.c. data $\mathcal{C}$ is simply called an interval-censored (i.c.) data. The p.i.c. data $\mathcal{C}$ is called a grouped data if $N = 1$ and each $\mathcal{C}_{1j}$ belongs to a set of mutually disjoint intervals whose union is equal to $\mathcal{R}$. The p.i.c. data $\mathcal{C}$ is called a binary response data if there exists a strictly increasing N sequence $\{x_i\}$ such that $\mathcal{C}_{ij} \in \{(-\infty, x_i), [x_i, \infty)\}$ for all $j = 1, \ldots, n_i$ and for all $i = 1, \ldots, N$. Note that the strictly increasing N sequence $\{x_i\}$ often corresponds to levels of the stimulus. Cheng and Amin (1982) discussed the existence of MLE's in the case where $h(s) = \log s$, the p.i.c. data $\mathcal{C}$ is a grouped data and $F(x)$ is the standard normal distribution function. Nakamura (1991) discussed the existence of MLE's in the case where $h(s) = \log s$ and the p.i.c. data $\mathcal{C}$ is an i.c. data. In this paper, we shall treat the case where the p.i.c. data $\mathcal{C}$ is a binary response data and the case $N \geq 3$.

Remark 2.1. Consider an another model: $H(P(x)) = ((x-\lambda)/\sigma)^{\gamma}$, where the parameter $(\gamma, \lambda, \sigma) \in \mathcal{R} \times \mathcal{R}_+ \times \mathcal{R}_+$ is to be estimated. This model implies that $P(x) \in \{G(((x-\lambda)/\sigma)^{\gamma}); -\infty < \lambda < \infty, \gamma > 0, \sigma > 0\}$. Here $G(x)$ is a d.f. on $\mathcal{R}_+$. Define $G_1(x) = G(\exp(x))$. Then $G_1(x)$ is a d.f. on $\mathcal{R}$ and $G_1(\gamma \log(x - \lambda) - \gamma \log \sigma) = G(((x-\lambda)/\sigma)^{\gamma})$. Hence this model can be reduced to our model.

Remark 2.2. Here we adopt the rule: $F(-\infty) = 0$ and $F(\infty) = 1$.

For estimating the true parameter $\theta_0 \in \mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$, we adopt a positive function $S(n)$ of positive integer n and a measure $D(z,p) : [0,1] \times [0,1] \to (-\infty, \infty]$ satisfying the following conditions:

(D.1) For every fixed $p \in [0,1]$, $D(z,p)$, as a function of z, is a continuous function from $[0,1]$ into the compact metric space $\bar{\mathcal{R}} = [-\infty, \infty]$ with the usual metric (see Bourbaki (1965)).

(D.2) For every fixed $p \in [0,1]$, $D(z,p)$, as a function of z, is a real-valued and continuously differentiable function on $(0,1)$.

Define $\ell : \mathcal{R}_+ \times \mathcal{R} \times \mathcal{R} \to \bar{\mathcal{R}}$ by

$$\ell(\theta) = \sum_{i=1}^{N} S(n_i)(D(F(t(x_i,\theta)), n_{i0}/n_i) + D(1 - F(t(x_i,\theta)), n_{i1}/n_i)),$$

where n_{i0} (resp. n_{i1}) denotes the number of $\mathcal{C}_{ij}$, $1 \leq j \leq n_i$ with $\mathcal{C}_{ij} = (-\infty, x_i)$ (resp. $\mathcal{C}_{ij} = [x_i, \infty)$). Following the terminology and the definition of Pfanzagl (1969) (or Grossmann (1982)), we say $\ell(\theta)$ a contrast function and define a minimum contrast estimate (MCE) for $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$ is a solution $\hat{\theta} \in \mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$ of the following minimization problem:

$$\ell(\hat{\theta}) = \min_{\theta \in \mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}} \ell(\theta).$$

Some measures $D(z,p)$ used for the estimation of parameters are as follows (see Takeuchi (1975), Berkson (1980) and Cressie and Read (1984)).

Example 2.2.

(i) Maximum likelihood (m.l.): $S(n) = 2n$ and

$$D(z,p) = \begin{cases} -p\log\left(\dfrac{z}{p}\right), & z \neq 0,\, p \neq 0 \\ 0, & p = 0, \\ \infty, & p \neq 0,\, z = 0. \end{cases}$$

(ii) Least squares (l.s.): $S(n) = 1$, $D(z,p) = (z-p)^2$.

(iii) Hellinger distance (H.d.): $S(n) = 1$, $D(z,p) = -\sqrt{zp}$.

(iv) Kullback-Leibler separator (K.L.s.): $S(n) = 2n$ and

$$D(z,p) = \begin{cases} z\log\left(\dfrac{z}{p}\right), & p \neq 0 \text{ and } z \neq 0 \\ 0, & p = 0 \text{ or } z = 0. \end{cases}$$

Remark 2.3. Other interesting examples are the minimum chi-square and the least absolute deviation. Unfortunately, our argument can not cover the least absolute deviation: $S(n) = 1$ and $D(z,p) = |z-p|$, because it is not differentiable.

Our aim in this paper is to give practical criteria for the existence of an MCE for $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$ when $h(s)$ is a $((\phi_1, h_1), (\phi_2, h_2), \phi)$-slowly varying function at infinity. In order to do this, we shall propose an approach called the probability contents boundary (PCB) analysis. Before describing this analysis, we need some notation and definitions. Let $\bar{\mathcal{S}}$ denote the closure of a subset $\mathcal{S}$ of Euclidean N-space $\mathcal{R}^N$ and $\mathcal{S}_1 - \mathcal{S}_2$ denote the difference between two subsets $\mathcal{S}_1$ and $\mathcal{S}_2$ of $\mathcal{R}^N$. Let expand the parameter space $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$ to the set $\mathcal{R}_+ \times \mathcal{R} \times [-\infty, \infty)$ as follows:

$$t(x, (\alpha, \beta, -\infty)) = \alpha h_1(x) - \beta, \qquad (\alpha, \beta) \in \mathcal{R}_+ \times \mathcal{R}.$$

Note that the contrast function $\ell(\theta)$ can be expanded to the set $\mathcal{R}_+ \times \mathcal{R} \times [-\infty, \infty)$. Let $\mathcal{Z} = \{(z_1, \ldots, z_N) \in \mathcal{R}^N; 0 \leq z_1 \leq \cdots \leq z_N \leq 1\}$ and define $\boldsymbol{F} : \mathcal{R}_+ \times \mathcal{R} \times [-\infty, \infty) \to \mathcal{Z}$ by $\boldsymbol{F}(\theta) = (F(t(x_1, \theta)), \ldots, F(t(x_N, \theta)))$. The set $\partial \boldsymbol{F}(\Theta) = \overline{\boldsymbol{F}(\Theta)} - \boldsymbol{F}(\Theta)$ is called the probability contents inner boundary (PCIB) of the family $\mathcal{F}(\Theta) \equiv \{F(t(x, \theta)); \theta \in \Theta\}$ with respect to the mapping $\boldsymbol{F}$, where Θ is an arbitrary nonempty subset of $\mathcal{R}_+ \times \mathcal{R} \times [-\infty, \infty)$. Define $L(\boldsymbol{z})$ $(\boldsymbol{z} \in \mathcal{Z})$ by

$$L(\boldsymbol{z}) = \sum_{i=1}^{N} S(n_i)(D(z_i, n_{i0}/n_i) + D(1 - z_i, n_{i1}/n_i)),$$

where $\boldsymbol{z} = (z_1, \ldots, z_N) \in \mathcal{Z}$. It follows from condition (D.1) that $L(\boldsymbol{z})$ is continuous on the compact set $\mathcal{Z}$ and $L(\boldsymbol{F}(\theta)) = \ell(\theta)$. A parameter θ' called an MCE for Θ if $\theta' \in \Theta$ and $\ell(\theta') = \min_{\theta \in \Theta} \ell(\theta)$.

The following result is fundamental to our analysis.

THEOREM 2.1. *Let Θ be a subset of $\mathcal{R}_+ \times \mathcal{R} \times [-\infty, \infty)$. An MCE for Θ exists if and only if*

(2.1) *there exists $\boldsymbol{z}^* \in \boldsymbol{F}(\Theta)$ such that $L(\boldsymbol{z}^*) \leq M_b(\Theta)$,* *where* $M_b(\Theta) = \inf\{L(\boldsymbol{z}); \boldsymbol{z} \in \partial \boldsymbol{F}(\Theta)\}$, *which is defined to be* ∞ *if* $\partial \boldsymbol{F}(\Theta) = \emptyset$ *(empty set).*

PROOF. Put $M_c = \inf\{L(\boldsymbol{z}); \boldsymbol{z} \in \overline{\boldsymbol{F}(\Theta)}\}$ and $M = \inf\{L(\boldsymbol{z}); \boldsymbol{z} \in \boldsymbol{F}(\Theta)\}$. The continuity of $L(\boldsymbol{z})$ yields the relation $M = M_c$. It is obvious that $M_c \le M_b(\Theta)$. Hence $M \le M_b(\Theta)$. The existence of an MCE, together with this inequality, means that (2.1) is satisfied. Conversely assume that (2.1) is satisfied. Let $\boldsymbol{z}^*$ be that of (2.1). If $L(\boldsymbol{z}^*) = M$, then an MCE exists. Let $M < L(\boldsymbol{z}^*)$. Choose $\hat{\boldsymbol{z}} \in \overline{\boldsymbol{F}(\Theta)}$ so that $M_c = L(\hat{\boldsymbol{z}})$. Then $L(\hat{\boldsymbol{z}}) = M < L(\boldsymbol{z}^*) \le M_b(\Theta)$. This means that $\hat{\boldsymbol{z}} \in \boldsymbol{F}(\Theta)$. Hence an MCE for Θ exists. This completes the proof. □

In order to make condition (2.1) available, we have to investigate the structure of $\partial \boldsymbol{F}(\Theta)$. Put $\mathbf{1} = (\overbrace{1, \ldots, 1}^{N})$, $\boldsymbol{a}_i(z) = (\overbrace{0, \ldots, 0}^{i-1}, z, \overbrace{1, \ldots, 1}^{N-i})$, $1 \le i \le N$; $0 \le z \le 1$ and $\boldsymbol{b}_i(z, z') = (\overbrace{0, \ldots, 0}^{i-1}, z, \overbrace{z', \ldots, z'}^{N-i})$, $1 \le i \le N-1$; $0 \le z \le z' \le 1$.

The following result is due to Nakamura (1984) and is useful for deriving a practical criterion for the existence of an MCE for $\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}$.

THEOREM 2.2. *The PCIB* $\partial \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\})$ *can be expressed as*

$$\partial \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}) = \{z\mathbf{1}; 0 < z < 1\} \cup \left(\bigcup_{i=1}^{N} \{\boldsymbol{a}_i(z); 0 \le z \le 1\} \right).$$

In order to find the structure of $\partial \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R})$, we make the following conditions.

(H.1) The function $h(s)$ is not convex linear, i.e.,

$$h(ux + (1-u)y) \ne uh(x) + (1-u)h(y)$$

for each $x, y \in \mathcal{R}_+$ and for each $u \in (0, 1)$.

(H.2) $h_1(x)$ is continuous on $\mathcal{R}$.

(H.3) $\lim_{s \to \infty} \phi_1(x+s)/\phi_1(s) = 1$ for all $x \in \mathcal{R}$.

Note that condition (H.2) implies that $h_1(x) = Kx$, where K is a positive constant (see Hirai *et al.* (1992)). Without loss of generality, we may assume that $h_1(x) = x$ by replacing $\phi_1(s)$ by $K\phi_1(s)$.

The following result is useful for deriving a practical criterion for the existence of an MCE for $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$. We do not give the proof here but details are available from the author if required (see Nakamura and Yokoyama (1992), Hirai *et al.* (1992)).

THEOREM 2.3. *Let conditions* (H.1)–(H.3) *be satisfied. Then the PCIB*

$\partial \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R})$ can be expressed as

$$\begin{aligned}\partial \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}) = \{z\mathbf{1}; 0 < z < 1\} &\cup \left(\bigcup_{i=1}^{N-1} \{\boldsymbol{a}_i(z); 0 \le z \le 1\}\right) \\ &\cup \left(\bigcup_{i=1}^{N-2} \{\boldsymbol{b}_i(z, z'); 0 \le z < z' < 1\}\right) \\ &\cup \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}).\end{aligned}$$

Now, we make some remarks about the practical meaning of the PCIB $\partial \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R})$. The element $z\mathbf{1}$ of the PCIB implies that: (i) In the interval $[x_1, x_N)$ of the stimulus, we always observe no response; (ii) In the interval $(-\infty, x_1)$ of the stimulus, we observe the response with probability z; (iii) In the interval (x_N, ∞) of the stimulus, we observe the response with probability $1 - z$. The element $\boldsymbol{a}_i(z)$ of the PCIB implies that: (i) In the interval $(-\infty, x_{i-1})$ of the stimulus, we always observe no response; (ii) In the interval $[x_{i-1}, x_i)$ of the stimulus, we observe the response with probability z; (iii) In the interval $[x_i, x_{i+1})$ of the stimulus, we observe the response with probability $1 - z$; (iv) In the interval $[x_{i+1}, \infty)$ of the stimulus, we observe no response. A similar interpretation to above can be obtained for the element $\boldsymbol{b}_i(z, z')$ of the PCIB. The relation $\boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}) \subset \partial \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R})$ implies that the two-parameter family $\{F(\alpha x - \beta); (\alpha, \beta) \in \mathcal{R}_+ \times \mathcal{R}\}$ competes with the underlying three-parameter family $\mathcal{F}$. Hence, the PCIB can be interpreted as a competing model to the three-parameter model.

3. Criteria

In this section, we shall give criteria for the existence of MCE's for $\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}$ and for $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$. For convenience sake, put $p_i = n_{i0}/n_i$ and $q_i = 1 - p_i$, $1 \le i \le N$. By the definition of $L(\boldsymbol{z})$,

$$L(z\mathbf{1}) = \sum_{i=1}^{N} S(n_i)(D(z, p_i) + D(1 - z, q_i)), \tag{3.1}$$

$$\begin{aligned}L(\boldsymbol{a}_i(z)) = &\sum_{j<i} S(n_j)(D(0, p_j) + D(1, q_j)) \\ &+ S(n_i)(D(z, p_i) + D(1 - z, q_i)) \\ &+ \sum_{j>i} S(n_j)(D(1, p_j) + D(0, q_j)), \qquad 1 \le i \le N.\end{aligned} \tag{3.2}$$

Choose $\hat{z}_i$ (resp. $\hat{z}_0$) $\in [0, 1]$ so that $L(\boldsymbol{a}_i(\hat{z}_i)) = \min_{0 \le z \le 1} L(\boldsymbol{a}_i(z))$ (resp. $L(\hat{z}_0\mathbf{1}) = \min_{0 \le z \le 1} L(z\mathbf{1})$), $1 \le i \le N$. Put $D_z(z, p) = \partial D(z, p)/\partial z$, $D_z(+0, p) = \lim_{z \to +0} D_z(z, p)$ and $D_z(1 - 0, p) = \lim_{z \to 1-0} D_z(z, p)$. Hereafter, we further assume that the measure $D(z, p)$ satisfies the following condition:

(D.3) $-\infty \le D_z(+0, p) < \infty$ and $-\infty < D_z(1 - 0, p) \le \infty$ for all $p \in [0, 1]$.

Next theorem gives a criterion for the existence of an MCE $(\tilde{\alpha}, \tilde{\beta}, -\infty)$ for $\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}$.

THEOREM 3.1. *An MCE* $(\tilde{\alpha}, \tilde{\beta}, -\infty)$ *for* $\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}$ *exists if conditions* (3.3)–(3.5) *or conditions* (3.6)–(3.8) *are satisfied.*

$$L(\hat{z}_0 \mathbf{1}) \leq \min_{1 \leq i \leq N} L(\boldsymbol{a}_i(\hat{z}_i)), \tag{3.3}$$

$$\sum_{i=1}^{N} S(n_i) D_z(+0, p_i) < \sum_{i=1}^{N} S(n_i) D_z(1-0, q_i) \quad \text{and} \quad \sum_{i=1}^{N} S(n_i) D_z(+0, q_i) < \sum_{i=1}^{N} S(n_i) D_z(1-0, p_i), \tag{3.4}$$

$$\sum_{i=1}^{N} S(n_i) D_z(\hat{z}_0, p_i) x_i < \sum_{i=1}^{N} S(n_i) D_z(1-\hat{z}_0, q_i) x_i; \tag{3.5}$$

$$L(\boldsymbol{a}_k(\hat{z}_k)) = \min_{1 \leq i \leq N} L(\boldsymbol{a}_i(\hat{z}_i)) \leq L(\hat{z}_0 \mathbf{1}), \tag{3.6}$$

$$D_z(+0, p_i) < D_z(1-0, q_i) \quad \text{for all } i = 1, \ldots, k, \tag{3.7}$$

$$D_z(+0, q_i) < D_z(1-0, p_i) \quad \text{for all } i = k, \ldots, N. \tag{3.8}$$

PROOF. Put $\Theta = \mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}$ and assume that (3.3)–(3.5) are satisfied. It follows from Theorem 2.2 that $M_b(\Theta) = \min(L(\hat{z}_0 \mathbf{1}), \min_{1 \leq i \leq N} L(\boldsymbol{a}_i(\hat{z}_i)))$. By virtue of (3.3), $M_b(\Theta) = L(\hat{z}_0 \mathbf{1})$. On the other hand, (3.1) and (3.4) yield that $\hat{z}_0 \in (0,1)$. Hence

$$\sum_{i=1}^{N} S(n_i) D_z(\hat{z}_0, p_i) = \sum_{i=1}^{N} S(n_i) D_z(1-\hat{z}_0, q_i). \tag{3.9}$$

Put $u = F^{-1}(\hat{z}_0)$, where $F^{-1}(z)$ is the inverse function of $F(x)$. Define a path $\theta(s)$ $(s > 0)$ by $\theta(s) = (s, -u-cs, -\infty)$, where c is an arbitrary real number. Since $L(\boldsymbol{F}(\theta(s))) = \sum_{i=1}^{N} S(n_i)(D(F((x_i+c)s+u), p_i) + D(1-F((x_i+c)s+u), q_i))$,

$$L'(\boldsymbol{F}(\theta(s))) = \sum_{i=1}^{N} S(n_i)(x_i+c) f((x_i+c)s+u) \cdot (D_z(F((x_i+c)s+u), p_i) - D_z(1-F((x_i+c)s+u), q_i)).$$

Hence, by (3.9),

$$\begin{aligned}\lim_{s \to +0} L'(\boldsymbol{F}(\theta(s))) &= f(u) \sum_{i=1}^{N} S(n_i)(x_i+c)(D_z(\hat{z}_0, p_i) - D_z(1-\hat{z}_0, q_i)) \\ &= f(u) \sum_{i=1}^{N} S(n_i) x_i (D_z(\hat{z}_0, p_i) - D_z(1-\hat{z}_0, q_i)).\end{aligned}$$

With the aid of (3.5), $L'(\boldsymbol{F}(\theta(s))) < 0$ for sufficiently small $s > 0$. Noting that $\lim_{s\to+0} L(\boldsymbol{F}(\theta(s))) = L(\hat{z}_0\mathbf{1})$, we see that there exists $\boldsymbol{z}^* \in \boldsymbol{F}(\Theta)$ such that $L(\boldsymbol{z}^*) < L(\hat{z}_0\mathbf{1}) = M_b(\Theta)$. This, together with Theorem 2.1, yields that an MCE for Θ exists.

Assume that (3.6)–(3.8) are satisfied. Theorem 2.2 and (3.6) yield $M_b(\Theta) = L(\boldsymbol{a}_k(\hat{z}_k))$. The inequalities $D_z(+0,p_k) < D_z(1-0,q_k)$ and $D_z(+0,q_k) < D_z(1-0,p_k)$ imply $0 < \hat{z}_k < 1$. Define $\theta(s)$ $(s > 0)$ by $\theta(s) = (s, sx_k - u_k, -\infty)$, where $u_k = F^{-1}(\hat{z}_k)$. Noting that $L(\boldsymbol{F}(\theta(s))) = \sum_{i=1}^N S(n_i)(D(F((x_i - x_k)s + u_k), p_i) + D(1 - F((x_i - x_k)s + u_k), q_i))$, we have

$$
\begin{aligned}
&L'(\boldsymbol{F}(\theta(s)))\\
&\quad= \sum_{1\le i<k} S(n_i)(x_i - x_k)f(a_i(s))(D_z(F(a_i(s)),p_i) - D_z(1 - F(a_i(s)),q_i))\\
&\qquad+ \sum_{k<i\le N} S(n_i)(x_i - x_k)f(a_i(s))(D_z(F(a_i(s)),p_i) - D_z(1 - F(a_i(s)),q_i)),
\end{aligned}
$$

where $a_i(s) = (x_i - x_k)s + u_k$, $1 \le i \le N$. Hence, by (3.7) and (3.8), $L'(\boldsymbol{F}(\theta(s))) > 0$ for sufficiently large s. This, together with the fact $\lim_{s\to+\infty} L(\boldsymbol{F}(\theta(s))) = L(\boldsymbol{a}_k(\hat{z}_k))$, yields that there exists $\boldsymbol{z}^* \in \boldsymbol{F}(\Theta)$ such that $L(\boldsymbol{z}^*) < L(\boldsymbol{a}_k(\hat{z}_k)) = M_b(\Theta)$. Thus, by Theorem 2.1, an MCE exists. This completes the proof. □

In order to give a criterion for the existence of an MCE $(\hat{\alpha}, \hat{\beta}, \hat{\lambda})$ for $\mathcal{R}_+\times\mathcal{R}\times\mathcal{R}$, we prepare some notation and definitions. By the definition of $L(\boldsymbol{z})$,

$$
\begin{aligned}
L(\boldsymbol{b}_i(u,v)) &= S(n_i)(D(u,p_i) + D(1-u,q_i))\\
&\quad+ \sum_{j\ge i+1} S(n_j)(D(v,p_j) + D(1-v,q_j))\\
&\quad+ \sum_{j\le i-1} S(n_j)(D(0,p_j) + D(1,q_j)), \qquad 1 \le i \le N-2.
\end{aligned}
$$

Denote by $(\hat{u}_i, \hat{v}_i) \in [0,1]\times[0,1]$, $1 \le i \le N-2$, an optimal solution of the following minimization problem: $L(\boldsymbol{b}_i(\hat{u}_i,\hat{v}_i)) = \min_{0\le u\le v\le 1} L(\boldsymbol{b}_i(u,v))$. Choose an integer m so that $L(\boldsymbol{b}_m(\hat{u}_m,\hat{v}_m)) = \min_{1\le i\le N-2} L(\boldsymbol{b}_i(\hat{u}_i,\hat{v}_i))$. Put

$$R(x,s) = (h(x+s) - h(s))/\phi_1(s) - h_1(x) - h_2(x)\phi_2(s)/\phi_1(s).$$

Note that $R(x,s) = o(1)$ $(s\to\infty)$ and

$$h(x+s) = h(s) + h_1(x)\phi(s) + h_2(x)\phi_2(s) + \phi_1(s)R(x,s). \tag{3.10}$$

Let us make the following conditions.

(H.4) $(\partial/\partial s)R(x,s) = o((d/ds)(\phi_2(s)/\phi_1(s)))$ $(s\to\infty)$ and $(d/ds)(\phi_2(s)/\phi_1(s)) < 0$ for sufficiently large $s > 0$.

(H.5) $h'(s) > 0$ on $\mathcal{R}_+$ and $h(s) = o(h'(s))$ $(s\to+0)$.

Now we prove the main result which gives a criterion for the existence of an MCE $(\hat{\alpha}, \hat{\beta}, \hat{\lambda})$ for $\mathcal{R}_+\times\mathcal{R}\times\mathcal{R}$.

THEOREM 3.2. *Assume that conditions* (H.1)–(H.5) *are satisfied. Let conditions* (3.3)–(3.5) *or* (3.6)–(3.8) *be satisfied and let* $\tilde{\theta} = (\tilde{\alpha}, \tilde{\beta}, -\infty)$ *be an MCE for* $\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}$. *Then an MCE* $(\hat{\alpha}, \hat{\beta}, \hat{\lambda})$ *for* $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$ *exists if any one of the following conditions is satisfied:*

(i) $\ell(\tilde{\theta}) \le L(\boldsymbol{b}_m(\hat{u}_m, \hat{v}_m))$ *and*

$$\sum_{i=1}^{N} S(n_i) h_2(x_i) f(\tilde{\alpha} x_i - \tilde{\beta}) D_z(F(\tilde{\alpha} x_i - \tilde{\beta}), p_i) < \sum_{i=1}^{N} S(n_i) h_2(x_i) f(\tilde{\alpha} x_i - \tilde{\beta}) D_z(1 - F(\tilde{\alpha} x_i - \tilde{\beta}), q_i). \tag{3.11}$$

(ii) $L(\boldsymbol{b}_m(\hat{u}_m, \hat{v}_m)) \le \ell(\tilde{\theta})$, $0 < \hat{u}_m < \hat{v}_m < 1$ *and*

$$\sum_{i \ge m+1} S(n_i) h(x_i - x_m) D_z(\hat{v}_m, p_i) < \sum_{i \ge m+1} S(n_i) h(x_i - x_m) D_z(1 - \hat{v}_m, q_i). \tag{3.12}$$

(iii) $L(\boldsymbol{b}_m(\hat{u}_m, \hat{v}_m)) \le \ell(\tilde{\theta})$, $\hat{u}_m = 0 < \hat{v}_m < 1$ *and for some* $x \in (x_m, x_{m+1})$,

$$\sum_{i \ge m+1} S(n_i) h(x_i - x) D_z(\hat{v}_m, p_i) < \sum_{i \ge m+1} S(n_i) h(x_i - x) D_z(1 - \hat{v}_m, q_i). \tag{3.13}$$

(iv) $L(\boldsymbol{b}_m(\hat{u}_m, \hat{v}_m)) \le \ell(\tilde{\theta})$, $0 < \hat{u}_m = \hat{v}_m < 1$ *and for some* $x \in (x_{m-1}, x_m)$,

$$\sum_{i \ge m+1} S(n_i) h(x_i - x) D_z(\hat{v}_m, p_i) < \sum_{i \ge m+1} S(n_i) h(x_i - x) D_z(1 - \hat{v}_m, q_i). \tag{3.14}$$

PROOF. Put $M_b = M_b(\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R})$. By Theorem 3.1 and by (3.3)–(3.5) or (3.6)–(3.8), an MCE $\tilde{\theta}$ for $\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}$ exists. Theorem 2.3 gives $M_b = \min(L(\hat{z}_0 \mathbf{1}), \min_{1 \le i \le N-1} L(\boldsymbol{a}_i(\hat{z}_i)), L(\boldsymbol{b}_m(\hat{u}_m, \hat{v}_m)), \ell(\tilde{\theta}))$. From Theorems 2.2 and 3.1 it follows that $M_b(\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\}) = \min(L(\hat{z}_0 \mathbf{1}), \min_{1 \le i \le N} L(\boldsymbol{a}_i(\hat{z}_i))) \ge \ell(\tilde{\theta})$. Hence, $M_b = \min(\ell(\tilde{\theta}), L(\boldsymbol{b}_m(\hat{u}_m, \hat{v}_m)))$.

Assume that (i) is satisfied. With the aid of the above discussion, this condition implies that $M_b = \ell(\tilde{\theta}) = L(\boldsymbol{F}(\tilde{\theta}))$. Define a path $\theta(s) = (\alpha(s), \beta(s), \lambda(s))$ by

$$\alpha(s) = \frac{\hat{\alpha}}{\phi_1(s)}; \qquad \beta(s) = \hat{\beta} + \alpha(s) h(s); \qquad \lambda(s) = -s.$$

Note that $\phi_1(s) > 0$ for sufficiently large $s > 0$, since $h(s)$ is a $((\phi_1, h_1), (\phi_2, h_2), \phi)$-slowly varying function at infinity. We show that

$$\lim_{s \to \infty} t(x, \theta(s)) = t(x, \tilde{\theta}) \qquad \text{for each } x \in \mathcal{R}. \tag{3.15}$$

The expression (3.10) and $t(x, \theta(s)) = \alpha(s) h(x + s) - \beta(s)$ give

$$t(x, \theta(s)) = \hat{\alpha} x - \hat{\beta} + \hat{\alpha} h_2(x) \frac{\phi_2(s)}{\phi_1(s)} + \hat{\alpha} R(x, s).$$

This, together with $\phi_2(s) = o(\phi_1(s))$ $(s \to \infty)$ and $R(x,s) = o(1)$ $(s \to \infty)$, yields (3.15). From (3.15) and $L(\boldsymbol{F}(\theta(s))) = \sum_{i=1}^N S(n_i)(D(F(t(x_i,\theta(s))),p_i) + D(1 - F(t(x_i,\theta(s))),q_i))$, it follows that $\lim_{s\to\infty} L(\boldsymbol{F}(\theta(s))) = L(\boldsymbol{F}(\tilde{\theta})) = M_b(\mathcal{R}_+ \times \mathcal{R} \times \{-\infty\})$. Note that

$$\begin{aligned} &L'(\boldsymbol{F}(\theta(s))) \\ &\quad = \sum_{i=1}^N S(n_i)\frac{\partial}{\partial s}t(x_i,\theta(s))f(t(x_i,\theta(s))) \\ &\qquad \cdot (D_z(F(t(x_i,\theta(s))),p_i) - D_z(1 - F(t(x_i,\theta(s))),q_i)) \end{aligned}$$

and

$$\frac{\partial}{\partial s}t(x,\theta(s)) = \hat{\alpha}h_2(x)\frac{d}{ds}\frac{\phi_2(s)}{\phi_1(s)} + \hat{\alpha}\frac{\partial}{\partial s}R(x,s).$$

By (H.4), we have

$$\begin{aligned} &\lim_{s\to\infty}\left(\hat{\alpha}\frac{d}{ds}\frac{\phi_2(s)}{\phi_1(s)}\right)^{-1} L'(\boldsymbol{F}(\theta(s))) \\ &\quad = \sum_{i=1}^N S(n_i)f(\hat{\alpha}x_i - \hat{\beta}) \\ &\qquad \cdot (D_z(F(\hat{\alpha}x_i - \hat{\beta}),p_i) - D_z(1 - F(\hat{\alpha}x_i - \hat{\beta}),q_i))h_2(x_i). \end{aligned}$$

The inequality (3.11) and (H.4) yield that $L'(\boldsymbol{F}(\theta(s))) > 0$ for sufficiently large $s > 0$. Hence there exists $\boldsymbol{z}^* \in \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R})$ such that $L(\boldsymbol{z}^*) < M_b$. By Theorem 2.3, an MCE for $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$ exists.

Assume that (ii) is satisfied. By the same reasoning as in the previous case, we can see that $M_b = L(\boldsymbol{b}_m(\hat{u}_m,\hat{v}_m))$. Put $t_1 = F^{-1}(\hat{u}_m)$, $t_2 = F^{-1}(\hat{v}_m)$, $y_1 = x_m$ and $y_2 = x_{m+1}$. Define $\theta(s) = (\alpha(s),\beta(s),\lambda(s))$ $(s > 0)$ by

$$\alpha(s) = \frac{t_2 - t_1}{h(s + y_2 - y_1) - h(s)}; \qquad \beta(s) = \alpha(s)h(s) - t_1; \qquad \lambda(s) = y_1 - s.$$

Then $t(y_i,\theta(s)) = t_i$ for all $s > 0$ $(i = 1,2)$. It can be easily seen that for each $x \geq y_2$,

$$\begin{aligned} &t(x,\theta(s)) = \frac{h(s) - h(s + x - y_1)}{h(s) - h(s + y_2 - y_1)}(t_2 - t_1) + t_1, \qquad x \geq y_1, \\ &\frac{h^2(s)}{h'(s)}\frac{\partial}{\partial s}t(x,\theta(s)) \\ &\quad = (t_2 - t_1)\left(\frac{h(s)}{h(s) - h(s + y_2 - y_1)}\right)^2 (h(s + x - y_1) - h(s + y_2 - y_1)) \\ &\qquad + \frac{h^2(s)}{h'(s)}O\left(\frac{1}{h(s)}\right) \qquad (s \to +0). \end{aligned}$$

The first equality and $h(\mathcal{R}_+) = \mathcal{R}$ yield that $\lim_{s\to+0} \boldsymbol{F}(\theta(s)) = (\overbrace{0,\dots,0}^{m-1}, \hat{u}_m, \hat{v}_m, \dots, \hat{v}_m) = \boldsymbol{b}_m(\hat{u}_m, \hat{v}_m)$. The second equality, together with $h(\mathcal{R}_+) = \mathcal{R}$ and (H.5), yields

$$\lim_{s\to+0} \frac{h^2(s)}{h'(s)} \frac{\partial}{\partial s} t(x, \theta(s)) = (t_2 - t_1)(h(x - y_1) - h(y_2 - y_1)). \tag{3.16}$$

It can be easily seen that

$$\begin{aligned} L(\boldsymbol{F}(\theta(s))) = & \sum_{i<m} S(n_i)(D(0, p_i) + D(1, q_i)) \\ & + S(n_m)(D(\hat{u}_m, p_m) + D(1 - \hat{u}_m, q_m)) \\ & + S(n_{m+1})(D(\hat{v}_m, p_{m+1}) + D(1 - \hat{v}_m, q_{m+1})) \\ & + \sum_{i\geq m+2} S(n_i)(D(F(t(x_i, \theta(s))), p_i) + D(1 - F(t(x_i, \theta(s))), q_i)). \end{aligned}$$

This and (3.16) give

$$\begin{aligned} & \lim_{s\to+0} \frac{h^2(s)}{h'(s)} L'(\boldsymbol{F}(\theta(s))) \\ & \quad = \sum_{i\geq m+2} S(n_i) f(t_2)(D_z(\hat{v}_m, p_i) - D_z(1 - \hat{v}_m, q_i)) \\ & \qquad \cdot (t_2 - t_1)(h(x_i - y_1) - h(y_2 - y_1)) \\ & \quad = (t_2 - t_1) f(t_2) \sum_{i\geq m+2} S(n_i) h(x_i - y_1)(D_z(\hat{v}_m, p_i) - D_z(1 - \hat{v}_m, q_i)) \\ & \qquad - (t_2 - t_1) f(t_2) h(y_2 - y_1) \sum_{i\geq m+2} S(n_i)(D_z(\hat{v}_m, p_i) - D_z(1 - \hat{v}_m, q_i)). \end{aligned}$$

Since $(\hat{u}_m, \hat{v}_m)$ is a minimizing point of $L(\boldsymbol{b}_m(u, v))$ over the set $\{(u, v);\ 0 \leq u \leq v \leq 1\}$ and since $0 < \hat{u}_m < \hat{v}_m < 1$,

$$\sum_{i\geq m+1} S(n_i)(D_z(\hat{v}_m, p_i) - D_z(1 - \hat{v}_m, q_i)) = 0.$$

Hence

$$\begin{aligned} & \lim_{s\to+0} \frac{h^2(s)}{h'(s)} L'(\boldsymbol{F}(\theta(s))) \\ & \quad = (t_2 - t_1) f(t_2) \sum_{i\geq m+1} S(n_i) h(x_i - y_1)(D_z(\hat{v}_m, p_i) - D_z(1 - \hat{v}_m, q_i)). \end{aligned}$$

We see, by (H.5) and (3.12), that $L'(\boldsymbol{F}(\theta(s))) < 0$ for sufficiently small $s > 0$. Hence there exists $\boldsymbol{z}^* \in \boldsymbol{F}(\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R})$ such that $L(\boldsymbol{z}^*) < M_b$. Hence, by Theorem 2.3, an MCE for $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$ exists.

Assume that (iii) is satisfied. By the same reasoning as in the proof of the previous case (ii), we have $M_b = L(\boldsymbol{b}_m(0, \hat{v}_m))$. Choose $x \in (x_m, x_{m+1})$ and $u \in (0, \hat{v}_m)$. Put $y_1 = x$, $y_2 = x_{m+1}$, $t_1 = F^{-1}(u)$ and $t_2 = F^{-1}(\hat{v}_m)$. The rest of the proof can be carried out by the same way as in the above case (ii).

Assume that (iv) is satisfied. By the same reasoning as in the proof of the previous case (ii), we obtain $M_b = L(\boldsymbol{b}_{m-1}(0, \hat{v}_m))$. Hence our assertion follows from the same argument as in the previous case (iii). This completes the proof. □

Remark 3.1. Theorem 3.2 gives a criterion for the existence of an MCE for $\mathcal{R}_+ \times \mathcal{R} \times \mathcal{R}$. This criterion contains optimal solutions $\tilde{\alpha}$, $\tilde{\beta}$, $\hat{z}_0$, $\hat{z}_i$ $(1 \le i \le N)$, $(\hat{u}_i, \hat{v}_i)$ $(1 \le i \le N-2)$. In general, these optimal solutions do not have explicit expression by the x_i's, the p_i's and the q_i's. Hence some iterative technique is needed to make this criterion available. The iterative techniques and simulational results will be discussed in the next section.

4. Examples

In Section 3, criteria were given for the existence of an MLE. These criteria, themselves, are not practical because the estimation is not specified. In this section we shall give practical form of these criteria for two kinds of estimation. Define $n. = \sum_{i=1}^N n_i$, $\overline{p} = n.^{-1} \sum_{i=1}^N n_i p_i$ and $\overline{q} = 1 - \overline{p}$

(i) Maximum likelihood estimation. Let $S(n) = 2n$ and $D(z,p) = -p\log(z/p)$ $((z,p) \in (0,1] \times [0,1])$, $D(0,p) = \infty$ $(p \in (0,1])$ and $D(z,0) = 0$ $(z \in [0,1])$. Conditions (3.3)–(3.5) become:

$\sum_{i=1}^{j-1} n_i p_i \ne 0$ or $\sum_{i=j+1}^{N} n_i q_i \ne 0$ or $n_j(p_j \log p_j + q_j \log q_j) \le n.(\overline{p}\log\overline{p} + \overline{q}\log\overline{q})$, $1 \le j \le N$;

$$0 < \overline{p} < 1;$$

$$\overline{p}\sum_{i=1}^N n_i q_i x_i < \overline{q}\sum_{i=1}^N n_i p_i x_i.$$

In this case, conditions (3.6)–(3.8) do not hold simultaneously. Recall that

$$\begin{aligned} 2^{-1}\ell(\alpha,\beta,-\infty) = &-\sum_{i=1}^N n_i(p_i \log F(\alpha x_i - \beta) + q_i \log(1 - F(\alpha x_i - \beta))) \\ &+ \sum_{i=1}^N n_i(p_i \log p_i + q_i \log q_i), \qquad (\alpha,\beta) \in \mathcal{R}_+ \times \mathcal{R}; \\ 2^{-1}L(\boldsymbol{b}_i(u,v)) = &-\left(\sum_{j<i-1} n_j p_j\right) \log 0 - n_i(p_i \log u + q_i \log(1-u)) \\ &- \left(\sum_{j>i} n_j p_j\right) \log v - \left(\sum_{j>i} n_j p_j\right) \log(1-v) \\ &+ \sum_{i=1}^N n_i(p_i \log p_i + q_i \log q_i), \end{aligned}$$

$$0 \le u \le v \le 1; \quad 1 \le i \le N-2.$$

As was stated in Section 3, a computational procedure is needed to evaluate the values of optimal solutions $\hat{\alpha}$ and $\hat{\beta}$. This procedure will be discussed in the next section. On the other hand, optimal solutions $(\hat{u}_i, \hat{v}_i)$ $(1 \le i \le N-2)$ can be represented explicitly by the p's. Put $v_{i1} = \min\{p_i, (N-i+1)^{-1}\sum_{j\ge i} p_j\}$ and $v_{i2} = \max\{p_i, (N-i)^{-1}\sum_{j\ge i+1} p_j\}$. Then

$$\hat{v}_i = \begin{cases} v_{i1}, & v_{i1}^2 + (N-i)(v_{i1}^2 - v_{i2}^2) \le 2(v_{i1} - v_{i2})\sum_{j\ge i+1} p_j, \\ v_{i2}, & \text{otherwise;} \end{cases} \tag{4.1}$$

$$\hat{u}_i = \begin{cases} \hat{v}_i, & 0 \le \hat{v}_i \le p_i, \\ p_i, & p_i < \hat{v}_i \le 1. \end{cases} \tag{4.2}$$

Hence the integer m and the value $L(\boldsymbol{b}_m(\hat{u}_m, \hat{v}_m))$ can be calculated exactly with ease computation. The inequality (3.11) becomes

$$\sum_{i=1}^{N} \frac{n_i q_i h_2(x_i) f(\hat{\alpha}x_i - \hat{\beta})}{1 - F(\hat{\alpha}x_i - \hat{\beta})} < \sum_{i=1}^{N} \frac{n_i p_i h_2(x_i) f(\hat{\alpha}x_i - \hat{\beta})}{F(\hat{\alpha}x_i - \hat{\beta})}.$$

The inequality (3.12) becomes

$$\sum_{i\ge m+1} \frac{n_i q_i h(x_i - x_m)}{1 - \hat{v}_m} < \sum_{i\ge m+1} \frac{n_i p_i h(x_i - x_m)}{\hat{v}_m}.$$

The inequalities (3.13) and (3.14) become

$$\sum_{i\ge m+1} \frac{n_i q_i h(x_i - x)}{1 - \hat{v}_m} < \sum_{i\ge m+1} \frac{n_i p_i h(x_i - x)}{\hat{v}_m}.$$

(ii) Least squares estimation. Let $S(n) = n$ and $D(z,p) = (z-p)^2$. Conditions (3.3)–(3.5) become:

$$\sum_{i=1}^{N} n_i p_i^2 - n.\overline{p}^2 \le \min_{1\le j\le N} \left\{ \sum_{i<j} n_i p_i^2 - \sum_{i>j} n_i q_i^2 \right\};$$

$$0 < \overline{p} < 1;$$

$$\overline{p}\sum_{i=1}^{N} n_i q_i x_i < \overline{q}\sum_{i=1}^{N} n_i p_i x_i.$$

Conditions (3.6)–(3.8) become:

$$\sum_{i<k} n_i p_i^2 - \sum_{i>k} n_i q_i^2 = \max_{1\le j\le N} \left\{ \sum_{i<j} n_i p_i^2 - \sum_{i>j} n_i q_i^2 \right\} \le \sum_{i=1}^{N} n_i p_i^2 - n.\overline{p}^2;$$

$$p_i > 0 \quad \text{for all } i = 1, \ldots, k;$$

$$q_i > 1 \quad \text{for all } i = k, \ldots, N.$$

Recall that

$$2^{-1}\ell(\alpha,\beta,-\infty)=\sum_{i=1}^{N}n_i(F(\alpha x_i-\beta)-p_i)^2,\qquad(\alpha,\beta)\in\mathcal{R}_+\times\mathcal{R};$$

$$2^{-1}L(\boldsymbol{b}_i(u,v))=u^2-2p_iu+(N-i)v^2-2\left(\sum_{j\geq i+1}p_j\right)v+\sum_{i=1}^{N}p_i^2,$$
$$0\leq u\leq v\leq 1;\quad 1\leq i\leq N-2.$$

Let v_{i1} and v_{i2} be the same as in the maximum likelihood case. Then $(\hat{u}_i,\hat{v}_i)$ can be given by (4.1) and (4.2). Hence the integer m and the value $L(\boldsymbol{b}_m(\hat{u}_m,\hat{v}_m))$ can be calculated exactly with ease computation. The inequality (3.11) becomes

$$\sum_{i=1}^{N}n_ih_2(x_i)f(\hat{\alpha}x_i-\hat{\beta})F(\hat{\alpha}x_i-\hat{\beta})<\sum_{i=1}^{N}n_ip_ih_2(x_i).$$

The inequality (3.12) becomes

$$\sum_{i=1}^{N}n_i\hat{v}_mh(x_i-x_m)<\sum_{i=1}^{N}n_ip_ih(x_i-x_m).$$

The inequalities (3.13) and (3.14) become

$$\sum_{i=1}^{N}n_i\hat{v}_mh(x_i-x)<\sum_{i=1}^{N}n_ip_ih(x_i-x).$$

By the same reasoning as in the previous case (i), a computational procedure is needed to evaluate values of optimal solutions $\hat{\alpha}$ and $\hat{\beta}$. This procedure will be discussed in the next section.

5. Simulation experiments

In this section we shall evaluate probabilities of the existence of MLE's and LSE's for $\mathcal{R}_+\times\mathcal{R}\times\mathcal{R}$, by our criterion (see Theorem 3.2), in the case where $h(x)=\log x$ and d.f. $F(x)$ is the logistic d.f. or the standard normal d.f. As was shown in Example 2.1, $h_1(x)=x$ and $h_2(x)=-x^2/2$ for the case $h(x)=\log x$. To make criterion obtained in Theorem 3.2 available, we have to evaluate values of $\hat{z}_0$, $\hat{z}_i$, $\hat{\alpha}$, $\hat{\beta}$, $(\hat{u}_i,\hat{v}_i)$ $(1\leq i\leq N-2)$. Recall that these optimal values are defined by the following minimization problems:

$$L(\hat{z}_0\mathbf{1})=\min_{0\leq z\leq 1}L(z\mathbf{1}),$$
$$L(\boldsymbol{a}_i(\hat{z}_i))=\min_{0\leq z\leq 1}L(\boldsymbol{a}_i(z)),\qquad 1\leq i\leq N,$$
$$L(\boldsymbol{b}_i(\hat{u}_i,\hat{v}_i))=\min_{0\leq u\leq v\leq 1}L(\boldsymbol{b}_i(u,v)),\qquad 1\leq i\leq N-2,$$
$$\ell(\hat{\alpha},\hat{\beta},-\infty)=\min_{(\alpha,\beta)\in\mathcal{R}_+\times\mathcal{R}}\ell(\alpha,\beta,-\infty).$$

Functions $L(z\mathbf{1})$ and $L(\boldsymbol{a}_i(z))$ are those of one variable. Hence, it is easy to calculate values of $\hat{z}_0$, $\hat{z}_i$ $(1 \leq i \leq N)$ when $D(z,p)$ is convex on $[0,1]$ or $D(z,p)$ is expressed as a simple form. Since $\partial^2 L(\boldsymbol{b}_i(u,v))/\partial u^2 = S(n_i)(D_{zz}(u,p_i) + D_{zz}(1-u,q_i))$, $\partial^2 L(\boldsymbol{b}_i(u,v))/\partial u \partial v = 0$, $\partial^2 L(\boldsymbol{b}_i(u,v))/\partial v^2 = \sum_{j>i+1} S(n_j)(D_{zz}(v,p_j) + D_{zz}(1-v,q_j))$, the function $L(\boldsymbol{b}_i(u,v))$ is strictly convex on the convex set $\{(u,v);\ 0 \leq u \leq v \leq 1\}$ if $D(z,p)$ is strictly convex on $[0,1]$ and has continuous second partial derivative $D_{zz}(z,p)$ with respect to z. Hence, it is easy to calculate values $(\hat{u}_i, \hat{v}_i)$ $(1 \leq i \leq N-2)$. It is difficult to show that $\ell(\alpha,\beta,-\infty)$ is convex on the convex set $\mathcal{R}_+ \times \mathcal{R}$ even if $D(z,p)$ is strictly convex on $[0,1]$. It is shown for the maximum likelihood estimation that $\ell(\alpha,\beta,-\infty)$ is strictly convex on $\mathcal{R}_+ \times \mathcal{R}$ when the density function $f(x)$ is strictly log-concave on $\mathcal{R}$ (see Burridge (1981)). Note that $f(x)$ is strictly log-concave if $F(x)$ is the logistic d.f. or the standard normal d.f. Hence, the function $\ell(\alpha,\beta,-\infty)$ for the maximum likelihood estimation is strictly convex on $\mathcal{R}_+ \times \mathcal{R}$ in our case. However, it is difficult to prove the strict convexity of $\ell(\alpha,\beta,-\infty)$ for the least squares estimation. Therefore, we adopt Davidon method as an iterative method for finding $(\hat{\alpha},\hat{\beta})$. Davidon method does not need the strict convexity of $\ell(\alpha,\beta,-\infty)$ and the twice differentiability of $\ell(\alpha,\beta,-\infty)$.

In this simulation study, we used $\log x$ as a stimulus metameter $h(x)$ and, as distributions, the loglogistic distribution and the lognormal distribution.

The simulation experiments were carried out on NEC-PC9801RA (personal computer, 32-bit word). The method of Box and Muller was used to generate normal random variables from uniformly distributed deviates which were produced by the method of Fishman and Moore (1982). The logistic variables were generated from uniform random variables by using the inverse function of distribution function. Random variables having the loglogistic distribution (resp. the lognormal distribution) were obtained from random variable having the logistic distribution (resp. normal distribution).

Consider the loglogistic distribution $F(x) = x(1+x)^{-1}$ and the case $N = 4$. Let $n_1 = n_2 = n_3 = n_4 = n$ and put $x_1 = t$, $x_2 = t+1.5$, $x_3 = t+3.0$, $x_4 = t+4.5$. Various selections of sample size n were taken (n = 10, 30, 50, 100, 500) and the values of t used were 0.111 (about equal to the lower 10% point of loglogistic distribution $F(x)$), 0.25 (lower 20% point), 0.428 (lower 30% point), 0.666 (lower 40% point), 1.0 (lower 50% point), 1.5 (lower 60% point), 2.333 (lower 70% point), 4.0 (lower 80% point). Simulated results based on 1000 replicates are displayed in Table 1.

Consider the lognormal distribution $F(x) = (1/\sqrt{2\pi}) \int_{-\infty}^{\log x} e^{-t^2/2} dt$ and the case $N = 4$. Let $n_1 = n_2 = n_3 = n_4 = n$ and put $x_1 = t$, $x_2 = t+0.5$, $x_3 = t+1.0$, $x_4 = t+1.5$. For each of sample size n (n = 10, 30, 50, 100, 500), the values of t used were 0.28 (about equal to the lower 10% point of lognormal distribution $F(x)$), 0.43 (lower 20% point), 0.59 (lower 30% point), 0.78 (lower 40% point), 1.0 (lower 50% point), 1.3 (lower 60% point), 1.7 (lower 70% point), 2.3 (lower 80% point). Simulated results based on 1000 replicates displayed in Table 2.

On a careful analysis of the Tables 1 and 2, we find that the percentage of samples, assuring the existence of an MLE or an LSE, of 1000 samples tends to increase as the size n increases and tends to decrease with the increase of the

Table 1. Percentages of samples satisfying our criterion (see Theorem 3.2) for the existence of an MLE and LSE in brackets, of 1000 simulations of sample size n.

	t							
n	0.111	0.25	0.428	0.666	1.0	1.5	2.333	4.0
10	91.5	79.4	71.7	59.7	46.6	32.8	27.3	21.3
	(56.8)	(55.2)	(49.4)	(46.7)	(45.8)	(43.8)	(38.1)	(34.1)
30	99.2	96.1	88.7	82.1	66.1	52.5	39.1	28.4
	(82.2)	(80.3)	(72.7)	(65.3)	(61.4)	(55.2)	(51.4)	(46.9)
50	99.9	98.3	94.5	87.9	75.7	61.3	47.4	34.9
	(90.4)	(89.4)	(81.3)	(73.1)	(66.0)	(60.3)	(54.8)	(47.4)
100	100.0	98.7	98.7	96.1	89.1	75.9	61.5	44.8
	(97.1)	(95.5)	(90.4)	(82.6)	(72.6)	(65.9)	(59.0)	(54.0)
500	100.0	99.8	99.6	99.4	98.7	95.8	85.9	68.9
	(100.0)	(100.0)	(99.9)	(98.5)	(93.5)	(83.6)	(75.1)	(65.7)

Table 2. Percentages of samples satisfying our criterion (see Theorem 3.2) for the existence of an MLE and LSE in brackets, of 1000 simulations of sample size n.

	t							
n	0.28	0.43	0.59	0.78	1.0	1.3	1.7	2.3
10	76.7	67.7	57.7	49.5	43.7	36.6	30.0	23.2
	(71.8)	(63.7)	(58.6)	(54.4)	(50.6)	(46.9)	(43.1)	(41.3)
30	90.2	81.5	71.8	65.4	58.4	48.6	42.2	33.7
	(93.4)	(94.0)	(85.8)	(79.8)	(67.4)	(62.0)	(56.7)	(50.2)
50	96.3	88.4	81.5	73.3	66.3	54.6	47.1	36.1
	(98.5)	(97.6)	(95.2)	(86.6)	(81.3)	(71.3)	(62.6)	(53.9)
100	99.0	93.0	89.3	82.1	73.7	63.5	56.5	51.7
	(100.0)	(99.9)	(99.1)	(95.9)	(90.2)	(81.3)	(71.6)	(56.8)
500	99.9	98.9	99.7	99.3	92.3	85.2	72.8	64.0
	(100.0)	(100.0)	(100.0)	(100.0)	(99.8)	(98.5)	(84.8)	(68.2)

percent point of the distribution. As is evident from these tables, the probability of the existence of an MLE (or an LSE) is close to one when n is sufficiently large and t lies in between 10% and 50% points of both distributions. These tables also send out a warning that the probability for the existence of the desired estimate is low when the sample size is small (≤ 30) or when the sampling plan is bad. We conclude that an MLE (or an LSE) does exists if the sample size is large and if the sampling plan is well.

References

Ash, J. M., Erdös, P. and Lubel, L. A. (1974). Very slowly varying functions, *Aequationes Math.*, **10**, 1–9.

Berkson, J. (1980). Minimum chi-square, not maximum likelihood!, *Ann. Statist.*, **8**, 457–487.

Bourbaki, N. (1965). *Topologie Generale*, Chap. 1 et 2, Hermann, Paris.

Burridge, J. (1981). A note on maximum likelihood estimation for regression models using grouped data, *J. Roy. Statist. Soc. Ser. B*, **43**, 41–45.

Cheng, R. C. H. and Amin, N. A. K. (1982). Estimating parameters in continuous univariate distributions with a shifted origin, Mathematics Report 82-1, University of Wales Institute of Science and Technology, Cardiff.

Cox, D. R. and Snell, E. J. (1989). *The Analysis of Binary Data*, Chapman and Hall, London.

Cressie, N. and Read, T. R. C. (1984). Multinomial goodness-of-fit tests, *J. Roy. Statist. Soc. Ser. B*, **46**, 440–464.

Elandt-Johnson, R. C. and Johnson, N. L. (1980). *Survival Models and Data Analysis*, Wiley, New York.

Finney, D. J. (1971). *Probit Analysis*, 3rd ed., Cambridge University Press, Cambridge.

Fishman, G. E. and Moore, L. R. (1982). A statistical evaluation of multiplicative congruential random number generators with $2^{31} - 1$, *J. Amer. Statist. Assoc.*, **77**, 129–136.

Grossmann, W. (1982). On the asymptotic properties of minimum contrast estimates, *Probability and Statistical Inference*, Proc. of the 2nd Pannonian Symposium on Math. Statist., Bad Tatzmannsdorff, Austria (eds. Grossmann, *et al.*), 115–127, Reidel, Dordrecht.

Hirai, Y., Nakamura, T. and Yokoyama, T. (1992). Structure of the probability contents inner boundary of some family of three-parameter distributions II: slowly varying case, *The Bulletin of Faculty of Education, Okayama University*, **91**, 19–29.

Mann, N. R., Schafer, R. E. and Singpurwalla, N. D. (1974). *Methods for Statistical Analysis of Reliability and Life Data*, Wiley, New York.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, Chapman and Hall, London.

Nakamura, T. (1984). Existence theorems of a maximum likelihood estimate from a generalized censored data sample, *Ann. Inst. Statist. Math.*, **36**, 375–393.

Nakamura, T. (1991). Existence of maximum likelihood estimates for interval-censored data from some three-parameter models with shifted origin, *J. Roy. Statist. Soc. Ser. B*, **53**, 211–220.

Nakamura, T. and Yokoyama, T. (1992). Structure of the probability contents inner boundary of some three-parameter distributions, *Hiroshima Math. J.*, **23** (in press).

Pfanzagl, J. (1969). On the measurability and consistency of minimum contrast estimates, *Metrika*, **14**, 249–272.

Prentice, R. L. (1976). A generalization of the probit and logit methods for dose response curves, *Biometrics*, **32**, 761–768.

Senata, E. (1976). Regularly varying functions, *Lectue Notes in Mathematics*, **508**, Springer, Berlin.

Takeuchi, K. (1975). *Kakuritsu-bunpu to Toukei-kaiseki*, Japanese Standard Association, Tokyo (in Japanese).