Mem. Fac. Sci. Shimane Univ.,
24, pp. 127–138 Dec. 24, 1991

Criteria for Existence of a Minimum Contrast Estimate from a Pooled Grouped Data

Tadashi NAKAMURA, Chae-Shin LEE* and Takashi KAYANO

Department of Information Science, Shimane University, Matsue City, Simane 690, Japan
(Received September 5, 1990)

Criteria for the existence of a minimum contrast estimate from a pooled grouped data are discussed. The minimum contrast estimate from the pooled grouped data covers both the maximum likelihood estimate and the minimum chi-square estimate from the grouped data as a special case. These criteria can be drived by a method called probability contents boundary analysis. This method gives systematically sufficient, or necessary and sufficient conditions for the existence of minimum contrast estimates for a wide class of families of distributions. Resulting criteria do not depend on the form of underlying distribution function.

1. Introduction

A criterion for the existence of a maximum likelihood estimate (MLE) from grouped data was given by Rao (1957) when the size of sample is large. To drive criteria for the existence of an MLE when the parameter space is an interval of the real line R, the method of Kulldorff (1962) and the method of successive approximation of Carter et al. (1971) are known to be efficient. Since these methods require some restriction on the parameter space, it seems to be difficult to generalize any one of them to the case where the sample space is a subset of Euclidean n-space $R^n (n \geq 2)$, or a discrete set. The probability contents boundary (p.c.b.) analysis proposed by Nakamura (1984) overcomes the demerits of these methods. In this analysis no restriction is imposed on the parameter space and on the form of distribution function

On the other hand, a criterion for the existence of a minimum chisquare estimate (MCSE) from a grouped data was also given by Rao (1955) when the sample size is large. But, in the case where the sample size is not large, there seems to be no systematic approach to finding criteria for the existence of MCSE's.

One of our purposes of this paper is to propose a unified approach for driving criteria for the existence of a minimum contrast estimate (MCE) from a pooled grouped data, which will be introduced in Section 2, and to give practical criteria for the existence of an MCE. In order to obtain practival criteria for the existence of

* Graduate School of Natural Science and Technology, Okayama University, Okayama City, Okayama 700, Japan

an MCE, a method called the p.c.b. analysis, is stated in Section 3. A practical criterion for the existence of an MCE is given in Section 4. The k-regularity of the underlying family of distributions is introduced in Section 4. In case the underlying family of distributions is k-regular, the above criteria is expressed by the observed frequencies only. Practical criteria for the 1-regular case (resp. 2-regular case) are given in Section 5. (resp. Section 6). In the case where the pooled grouped data is a binary response data, a practical criterion for the 2-regular case is given in Section 7.

The theoretical background of the p.c.b. analysis is given in the appendix.

2. Problem setting

Let T' and T be given constants with $-\infty \leq T' < T \leq \infty$, X be a random variable with values in the open interval (T', T) and the distribution of X belong to a family $\mathscr{P} = \{P_\theta;\ \theta \in \Theta\}$ of probability measures on (T', T), where the parameter space Θ is an arbitrary nonempty set. Let $(X_{h1},\dots,X_{hn_h})$, $1 \leq h \leq \ell$, be a random sample from the distribution of X and suppose that information available for X_{hi} is only that its value lies in a set $\mathscr{C}_{hi} \in \{[x_{h0}, x_{h1}),\dots,[x_{hr_h}, x_{hr_h+1})\}$, where the positive integers r_h's of observed times and the group limits x_{hi}'s are pre-assigned and independent of the size n_h and of θ, and $x_{h0} = T' < x_{h1} < \cdots < x_{hr_h} < x_{hr_h+1} = T$. The collection $\mathscr{C} = \{C_{hi};\ 0 \leq i \leq n_h,\ 1 \leq h \leq \ell\}$ is called a pooled grouped data of size $n.$, where $n. = \sum_{h=1}^{\ell} n_h$. We assume that

(2.1) *there exists no* $\mathscr{C}_{hi}$ *such that* $P_\theta(\mathscr{C}_{hi}) \equiv 0$ *or* 1 *on* Θ.

Let $S(t)$ be a finite-valued and positive function on $(0, \infty)$ and let $D(z, p)$ be a finite-valued function on $(0, 1] \times [0, 1]$ such that $D(z, p)$, as a function of z, is continuous on $(0, 1]$ for every fixed $p \in [0, 1]$ and that $D(z, 0)$ is independent of $z \in (0, 1]$. Assume that, for every $p \in [0, 1]$, there exists an extended real number $K_D(p)$ with $-\infty < K_D(p) \leq \infty$ such that

$$\lim_{z \to 0} D(z, p) = K_D(p) \quad \text{for all } p \in [0, 1]. \tag{2.2}$$

Consider the following minimizing problem:

$$(P) \qquad \textit{Find } \inf\Big\{\sum_{h=1}^{\ell}\sum_{k=0}^{r_h} S(n_h) D(P_\theta([x_{hk}, x_{hk+1}), n_{hk}/n_h);\ \theta \in \Theta\Big\},$$

where n_{hk}, $1 \leq h \leq \ell;\ 0 \leq k \leq r_h$, is the number of $\mathscr{C}_{hj}$, $1 \leq j \leq n_h$, such that $\mathscr{C}_{hj} = [x_{hk}, x_{hk+1})$. The objective function of Problem (P) is called a contrast function. In computing the contrast function, the following rules are used: $D(0, p) = K_D(p)$ for all $p \in [0, 1]$, $t \cdot (\pm\infty) = \pm\infty$ for all $t > 0$ and $\infty + \infty = \infty$. Following the definition of Grossmann (1982), we say an optimal solution of

Problem (P) a minimum contrast estimate (MCE) of the unknown true parameter of the distribution of X.

We shall give the four examples of the measure $D(z, p)$ (see Berkson (1980)).

Example 2.1 (Minimum chi-square estimation). Let $S(t) = t$ and let $D(z, p) = (z - p)^2/z$ for $(z, p) \in (0, 1]$ and $D(z, 0) = 0$ for $z \in (0, 1]$. In this case, $K_D(p) = \infty$ for all $p \in (0, 1]$ and the MCE is called a minimum chi-square estimate (MCSE).

Example 2.2 (Maximum likelihood estimation). Let $S(t) = 2t$ and let $D(z, p) = p \cdot \log(p/z)$ for $(z, p) \in (0, 1] \times (0, 1]$ and $D(z, 0) = 0$ for $z \in (0, 1]$. In this case, $K_D(p) = \infty$ for all $p \in (0, 1]$ and the MCE is called a maximum likelihood estimate (MLE).

Example 2.3 (Minimum logit chi-square estimation). Let $S(t) = t$ and let $D(z, p) = p(1-p)(\log p(1-p)^{-1} - \log z(1-z)^{-1})^2$ for $(z, p) \in (0, 1) \times (0, 1)$ and $D(z, p) = 0$ for $(z, p) \in (0, 1] \times \{1\} \cup \{1\} \times [0, 1] \cup (0, 1] \times \{0\}$. In this case, $K_D(p) = \infty$ for all $p \in (0, 1)$, $K_D(1) = 0$ and the MCE is called a minimum logit chi-square estimate (MLCSE)

Example 2.4 (Kullback-Leibler estimation). Let $S(t) = 2t$ and let $D(z, p) = z \log(z/p)$ for $(z, p) \in (0, 1] \times (0, 1]$ and $D(z, 0) = 0$ for $z \in (0, 1]$. In this case, $K_D(p) = 0$ for all $p \in (0, 1]$ and the MCE is called Kullback-Leibler estimate (KLE).

3. Probability contents boundary analysis

Our aim is to find criteria for the existence of an MCE. To do this, we shall propose an approach, which is called the probability contents boundary (p.c.b.) analysis. Before stating this approach, we need some notation and definitions. Put $F(x, \theta) = P_\theta([T', x))$. The two points x and x' of $[T', T]$ are said to be equivalent (with respect to the family $\mathscr{P}$) if $F(x', \theta) \equiv F(x, \theta)$ on Θ. Let $\bar{R}$ (resp. $\bar{\varphi}$) denote the set of extended real numbers (resp. the closure of a subset of R^r), and for subsets φ_1 and φ_2 of R^r, $\varphi_1 - \varphi_2$ denote the difference between φ_1 and φ_2. Put $Z_r = \{(z_1, \dots, z_r) \in R^r; 0 < z_1 < \cdots < z_r < 1\}$, and define $\tilde{D}(z, p)$ by $\tilde{D}(z, p) = D(z, p)$ if $(z, p) \in (0, 1] \times [0, 1]$ and by $\tilde{D}(z, p) = K_D(p)$ if $(z, p) \in \{0\} \times [0, 1]$. The p.c.b. analysis consists of the following four steps:

Step 1: *Find a positive integer m and a set $\{x_i\}$ of $(m+2)$ points of $[T', T]$ with the following properties*:

$$x_0 = T' < x_1 < \cdots < x_m < x_{m+1} = T.$$

The points x_i and x_j are not equivalent whenever $i \neq j$.
Each extreme point of $\mathscr{C}_{hk}$, $1 \le h \le \ell$; $1 \le k \le n_h$ is equivalent to some x_i, $0 \le i \le m+1$.

Each x_i, $1 \le i \le m$, *is equivalent to some extreme point of* C_{hk}, $1 \le h \le \ell$; $1 \le k \le n_h$.

Step 2: *Definite nonnegative integers* n_{hij}, $1 \le h \le \ell$; $0 \le i \le j \le m+1$, *a mapping* $\mathbf{F}: \Theta \to \bar{Z}_m$ *and a function* $L: \bar{Z}_m \to \bar{R}$ *by*

n_{hij} = *number of* $\mathscr{C}_{hk}$, $1 \le k \le n_h$ *such that* x_i*(resp.* x_j*) is equivalent to the small (resp. large) one of extreme points of* $\mathscr{C}_{hk}$,

$\mathbf{F}(\theta) = (F(x_1, \theta), \ldots, F(x_m, \theta))$,

$$L(\mathbf{z}) = \sum_{h=1}^{\ell} S(n_h) \sum_{0 \le i < j \le m+1; p_{hij} \ne 0} \tilde{D}(z_j - z_i, p_{hij}), \tag{3.1}$$

where $z_0 = 0$, $z_{m+1} = 1$ *and* $p_{hij} = n_{hij}/n_h$.

Step 3: *Determine the probability contents inner boundary (p.c.i.b) of the family* $\mathscr{P}$ *(for the pooled grouped data* $\mathscr{C}$*)*

$$\partial\mathbf{F}(\Theta) = \overline{\mathbf{F}(\Theta)} - \mathbf{F}(\Theta).$$

Step 4: *Find a sufficient condition which implies the following condition*:

(3.2) *There exists* $\mathbf{z}' \in \mathbf{F}(\Theta)$ *such that* $L(\mathbf{z}') \le M_b \equiv \inf\{L(\mathbf{z}); \mathbf{z} \in \partial\mathbf{F}(\Theta)\}$,

where the infimum of a function over the empty set ($\emptyset$) *is defined to be* ∞.

Steps 1 and 2 can be easily perforemed. Determination of the structure of the p.c.i.b. $\partial\mathbf{F}(\Theta)$ is very important in the p.c.b. analysis. For detailed discussions on the structure of $\partial\mathbf{F}(\Theta)$, see Nakamura (1985b, 1985c). After the performance of Step 3, the procedure in Step 4 can be performed without difficulty. A sufficient condition obtained in Step 4 becomes a criterion for the existence of an MCE. This fact is due to the following theorem. The proof is given in the appendix.

THEOREM 3.1. *An MCE exists if and only if condition* (3.2) *is satisfied.*

As an immediate consequence of Theorem 3.1, we have

COROLLARY 3.1. *If* $M_b = \infty$, *then an MCE exists.*

In case $\partial\mathbf{F}(\Theta) = \emptyset$ then $M_b = \infty$ and by Corollary 3.1, an MCE always exists. Hereafter, unless otherwise stated, we assume that $\partial\mathbf{F}(\Theta) \ne \emptyset$ and $m \ge 2$. It should be noted that, because of (2.1), $n_h = \sum_{j=1}^{m+1} \sum_{i=0}^{j-1} n_{hij}$, $1 \le h \le \ell$.

4. Evaluation of the value of M_b

A criterion for the existence of an MCE is obtained by seeking a sufficient condition for the statement (3.2). For this reason, it is important to find the value of M_b (see Step 4 for its definition). For many of the well-known estimations in the statistical inference (see Berkson (1980)), there are, possibly, two cases:

Case I. $K_D(p) = \infty$ for all $p \in (0, 1)$.
Case II. $K_D(p) < \infty$ for all $p \in (0, 1]$.

In case II, the value of M_b does depend on an estimation, while in Case I, that of M_b does not so depend on an estimation as Case II. To apply the p.c.b. analysis systematically, we shall concentrate our attention to Case I. A detailed discussion for Case II are to be given in a forthcoming paper.

To find the value of M_b in terms of the p.c.i.b. $\partial\mathbf{F}(\Theta)$ and the observed relative frequencies p_{hij}'s, express $L(\mathbf{z})$ as

$$L(\mathbf{z}) = \sum_{h=1}^{\ell} S(n_h)\Big(\sum_{(i,j)\in\mathcal{N}(h;\mathbf{z})} D(z_j - z_i, p_{hij}) + \sum_{(i,j)\in\leq\mathcal{N}_1(h;\mathbf{z})} D(1, p_{hij})\Big) + \sum_{h=1}^{\ell} S(n_h) \sum_{(i,j)\in\mathcal{N}_0(h;\mathbf{z})} K_D(p_{hij}), \tag{4.1}$$

where $\mathbf{z} = (z_1, \dots, z_m) \in \bar{Z}_m$ and

$$\mathcal{N}(h; \mathbf{z}) = \{(i, j);\ 0 \leq i < j \leq m+1,\ p_{hij} \neq 0 \textit{ and } 0 < z_j - z_i < 1\}, \tag{4.2}$$

$$\mathcal{N}_1(h; \mathbf{z}) = \{(i, j);\ 0 \leq i < j \leq m+1,\ p_{hij} \neq 0 \textit{ and } z_j - z_i = 1\}, \tag{4.3}$$

$$\mathcal{N}_0(h; \mathbf{z}) = \{(i, j);\ 0 \leq i < j \leq m+1,\ p_{hij} \neq 0 \textit{ and } z_j - z_i = 0\}, \tag{4.4}$$

Here we use the rule: $\infty \cdot 0 = 0$, and the sum over the empty set is defined to be 0.

We have

THEOREM 4.1. *Let $K_D(p) = \infty$ for all $p \in (0, 1]$. Then $M_b = \infty$ if and only if*

$$\sum_{h=1}^{\ell} \sum_{(i,j)\in\mathcal{N}_0(h;\mathbf{z})} p_{hij} > 0 \qquad \textit{for all } \mathbf{z} \in \partial\mathbf{F}(\Theta). \tag{4.5}$$

PROOF. If (4.5) is satisfied, then for each $\mathbf{z} \in \partial\mathbf{F}(\Theta)$, there exists a triple (h, i, j) such that $(i, j) \in \mathcal{N}_0(h; \mathbf{z})$. Hence, $K_D(p_{hij}) = \infty$ and the last term of the right-hand side of (4.1) is equal to ∞. This, together with (4.1), implies $M_b = \infty$. To prove the converse, assume that (4.5) is not satisfied. Then for every h, $1 \leq h \leq \ell$, and for every $\mathbf{z} \in \partial\mathbf{F}(\Theta)$, $\sum_{h=1}^{\ell} \sum_{(i,j)\in\mathcal{N}_0(h;\mathbf{z})} p_{hij} = 0$. This implies that $\mathcal{N}_0(h; \mathbf{z}) = \emptyset$ for all h, $1 \leq h \leq \ell$ and $L(\mathbf{z}) < \infty$ for all $\mathbf{z} \in \partial\mathbf{F}(\Theta)$. Hence $M_b < \infty$. This completes the proof.

Similarly we have

THEOREM 4.2. *Let $K_D(p) = \infty$ for all $p \in (0, 1)$ and $K_D(1) < \infty$. Then $M_b = \infty$ if and only if*

$$\sum_{h=1}^{\ell} \sum_{(i,j)\in\mathcal{N}_{00}(h;\mathbf{z})} p_{hij} > 0 \qquad \textit{for all } \mathbf{z} \in \partial\mathbf{F}(\Theta), \tag{4.6}$$

where $\mathcal{N}_{00}(h; \mathbf{z}) = \{(i, j) \in \mathcal{N}_0(h; \mathbf{z}); p_{hij} \neq 1\}$.

Both condition (4.5) and condition (4.6) are sufficient conditions for the existence of an MCE (see Corollary 3.1). But they are not practical themselves unless the structure of the p.c.i.b. $\partial\mathrm{F}(\Theta)$ is expressed in a precise form. A good many families of distributions used in the statistical analysis have, commonly, some kind of structure called the k-regurarity (see Nakamura (1985b, 1985c)), which will be defined below. The family $\mathscr{P}$ is said to be k-regular if, for any $(z_1,\dots,z_m) \in \partial\mathrm{F}(\Theta)$, the number of distinct z_j's ($i = 1,\dots,m$) values such that $0 < z_j < 1$ is at most equal to k-1. In the case where the family $\mathscr{P}$ is k-regular, the left-hand sides of (4.5) and (4.6) can be expressed by the observed relative frequencies only, that is, those conditions turn out to be a practical criterion for the existence of an MCE. Henceforth, to proceed a systematic argument, we assume that the family $\mathscr{P}$ is 1-regular or 2-regular and that $K_D(p) = \infty$ for all $p \in (0, 1]$ (the argument in the case where $K_D(p) = \infty$ for all $p \in (0, 1)$ and $K_D(1) < \infty$ can be carried out by the same argument as the case just stated and by Theorem 4.2). It should be noted that the p.c.b. analysis also derives practical criteria for the existence of an MCE in non-regular cases (see Nakamura (1985a, 1990)).

5. Practical criteria for the 1-regular case

In this section we shall give practical criteria for the existence of an MCE in the case where $K_D(p) = \infty$ for all $p \in (0, 1]$, and the family $\mathscr{P}$ is 1-regular, i.e.,

$$\partial\mathrm{F}(\Theta) \subset \{\mathbf{a}_0,\dots,\mathbf{a}_m\}, \tag{5.1}$$

where $\mathbf{a}_i = (\overbrace{0,\dots,0}^{i}, 1,\dots,1)$. Note that $\mathbf{0} \equiv (0,\dots,0) = \mathbf{a}_m$ and $\mathbf{1} \equiv (1,\dots,1) = \mathbf{a}_0$. Define $p_{hi\cdot}$ and $p_{\cdot i\cdot}$ by $p_{hi\cdot} = \sum_{j=i+1}^{m+1} p_{hij}$ and by $p_{\cdot i\cdot} = \sum_{h=1}^{\ell} p_{hi\cdot}$ respectively. The notation $p_{h\cdot j}$ and $p_{\cdot\cdot j}$ can be defined analogously. It follows from (4.2)–(4.4) that for each k, $0 \le k \le m$ and for each h, $1 \le h \le \ell$,

$$\mathcal{N}(h; \mathbf{a}_k) = \emptyset \tag{5.2}$$

$$\mathcal{N}_1(h; \mathbf{a}_k) = \{(i, j); 0 \le i \le k < j \le m + 1 \text{ and } p_{hij} \neq 0\}, \tag{5.3}$$

$$\mathcal{N}_0(h; \mathbf{a}_k) = \{(i, j); 0 \le i < j \le k \text{ or } k < i < j \le m + 1, \text{ and } p_{hij} \neq 0\}. \tag{5.4}$$

From this and (4.1)

$$L(\mathbf{a}_k) = \sum_h S(n_h) \sum_{(i,j) \in \mathcal{N}_1(h;\mathbf{a}_k)} D(1, p_{hij}) + \sum_h S(n_h) \sum_{(i,j) \in \mathcal{N}_0(h;\mathbf{a}_k)} K_D(p_{hij}), \tag{5.5}$$

where the symbol $\sum_h$ denotes the summation from $h = 1$ to ℓ. Note that the last term of the right-hand side of (5.5) is equal to 0 if $\mathcal{N}_0(h; \mathbf{a}_k) = \emptyset$ for all $h = 1,\dots,\ell$.

We have

THEOREM 5.1. *An MCE exists if the following condition is satisfied*:

$$\mathbf{a}_k \notin \partial \mathrm{F}(\Theta) \ or \ \sum_{j=1}^{k} p_{\cdot\cdot j} + \sum_{i=k+1}^{m} p_{\cdot 1 \cdot} \neq 0, \ k = 0,\dots,m. \tag{5.6}$$

PROOF. The expression (5.4) derives that $\sum_{(i,j)\in\mathcal{N}_0(h;\mathbf{a}_k)} p_{hij} = \sum_{j=1}^{k} p_{h\cdot j} + \sum_{i=k+1}^{m} p_{hi\cdot}$ and $\sum_h \sum_{(i,j)\in\mathcal{N}_0(h;\mathbf{a}_k)} p_{hij} = \sum_{j=1}^{k} p_{\cdot\cdot j} + \sum_{i=1k+1}^{m} p_{\cdot i\cdot}$. This, together with (5.1) and(5.6), yields that condition (4.5) is satisfied. Because of Theorem 4.1 and Corollary 3.1, an MCE exists.

Remark 5.1. Condition (5.6) does not depend on the form of $F(x, \theta)$ but on the p.c.i.b. $\partial \mathrm{F}(\Theta)$.

The following theorem shows that under some restrictions, condition (5.6) is necessary for the existence of an MCE.

THEOREM 5.2. *Assume that*:

$$0 < z_1 \le z_m < 1 \qquad \textit{for all } (z_1,\dots,z_m)\in \mathrm{F}(\Theta). \tag{5.7}$$

$$\tilde{D}(1, 1) < \tilde{D}(z, 1) \qquad \textit{for all } z\in[0, 1). \tag{5.8}$$

Then an MCE exists if and only if condition (5.6) *is satisfied.*

PROOF. Because of Theorem 5.1, it suffices to show the "only if" part of the theorem. Suppose that $\mathbf{a}_k \in \partial \mathrm{F}(\Theta)$ and $\sum_{j=1}^{k} p_{\cdot\cdot 1} + \sum_{i=k+1}^{m} p_{\cdot i\cdot} = 0$ for some k, $0 \le k \le m$. Then $\sum_{j=1}^{k} p_{h\cdot j} + \sum_{i=k+1}^{m} p_{hi\cdot} = 0$ for all $h = 1,\dots,\ell$. This implies that, for each h, $1 \le h \le \ell$, there exists a unique pair (i_h, j_h) of integers with $0 \le i_h \le k < j_h \le m+1$ and $p_{hi_hj_h} = 1$. From (5.2)–(5.4), $\mathcal{N}(h; \mathbf{a}_k) = \mathcal{N}_0(h; \mathbf{a}_k) = \emptyset$ and $\mathcal{N}_1(h; \mathbf{a}_k) = \{(i_h, j_h)\}$ for all $h = 1,\dots,\ell$. This, together with (4.1), (5.5), (5.7) and (5.8), yields that for each $\mathbf{z} = (z_1,\dots,z_m)\in \mathrm{F}(\Theta)$,

$$\begin{aligned} L(\mathbf{z}) &= \textstyle\sum_h S(n_h)\tilde{D}(z_{j_h} - z_{i_h}, 1) \\ &= \textstyle\sum_h S(n_h)\sum_{(i,j)\in\mathcal{N}_1(h;\mathbf{a}_k)} \tilde{D}(z_j - z_i, 1) \\ &> \textstyle\sum_h S(n_h)\sum_{(i,j)\in\mathcal{N}_1(h;\mathbf{a}_k)} D(1, 1) = L(\mathbf{a}_k). \end{aligned}$$

Hence $M_b < L(\mathbf{z})$ for all $\mathbf{z}\in \mathrm{F}(\Theta)$, and by Theorem 3.1, an MCE does not exist. This completes the proof.

6. Practical criteria for the 2-regular case (part 1)

In this section we shall give practical criteria for the existence of an MCE in the case where $K_D(p) = \infty$ for all $p\in(0, 1]$, and the family $\mathscr{P}$ is 2-regular, i.e.,

$$\partial \mathrm{F}(\Theta) \subset \{\mathbf{a}_0,\dots,\mathbf{a}_m\} \cup \mathscr{A} \cup (\textstyle\bigcup_{i=1}^{m} \mathscr{A}_i), \tag{6.1}$$

where $\mathscr{A}=\{z\mathbb{1}\,;\,0<z<1\}$ and $\mathscr{A}_i=\{\mathbf{a}_i(z)\,;\,0<z<1\}$ with $\mathbf{a}_i(z)=(\overbrace{0,\dots,0}^{i-1}, z, \overbrace{1,\dots,1}^{m-i})$. Note that, if the family $\mathscr{P}$ is 1-regular, then the family $\mathscr{P}$ is 2-regular. It is esay to see that:

$$\mathscr{N}_0(h;\mathbf{z})=\{(i,j);\,0\le i<j<k \text{ or } k<i<j\le m+1, \text{ and } p_{hij}\neq 0\} \text{ if } \mathbf{z}\in\mathscr{A}_k. \tag{6.2}$$

$$\mathscr{N}_0(h;\mathbf{z})=\{(i,j);\,1\le i<j\le m \text{ and } p_{hij}\neq 0\} \text{ if } \mathbf{z}\in\mathscr{A}. \tag{6.3}$$

Put $p_{\cdot ij}=\sum_h p_{hij}$.

We have

THEOREM 6.1. *An MCE exists if condition* (5.6) *and the following conditions are satisfied*:

$$\partial\mathbf{F}(\Theta)\cap\mathscr{A}=\emptyset \text{ or } \sum_{1\le i<j\le m} p_{\cdot ij}\neq 0. \tag{6.4}$$

$$\partial\mathbf{F}(\Theta)\cap\mathscr{A}_k=\emptyset \text{ or } \sum_{j=1}^{k-1} p_{\cdot\cdot j}+\sum_{i=k+1}^{m} p_{\cdot i\cdot}\neq 0,\ k=1,\dots,m. \tag{6.5}$$

PROOF. From (6.2) and (6.3),

$$\sum_h\sum_{(i,j)\in\mathscr{N}_0(h;\mathbf{z})} p_{hij}=\begin{cases}\sum_{1\le i\le j\le m} p_{\cdot ij} & \text{if } \mathbf{z}\in\mathscr{A},\\ \sum_{j=1}^{k-1} p_{\cdot\cdot j}+\sum_{i=k+1}^{m} p_{\cdot i\cdot} & \text{if } \mathbf{z}\in\mathscr{A}_k.\end{cases}$$

This, together with (6.1), (6.4) and (6.5), shows that condition (4.5) is satisfied. The existence of an MCE follows from Theorem 4.1 and Corollary 3.1.

Remark 6.1. Conditions (6.4)–(6.5) do not depends on the form of $F(x,\theta)$.

Under some restriction, we can give a necessary and sufficient condition for the existence of an MCE.

THEOREM 6.2. *Let* $\ell=1$. *Assume that condition* (5.7) *and the following conditions are satisfied*:

(6.6) $\mathscr{A}_k\subset\partial\mathbf{F}(\Theta)$ *whenever* $\partial\mathbf{F}(\Theta)\cap\mathscr{A}_k\neq\emptyset$.

(6.7) $\mathscr{A}\subset\partial\mathbf{F}(\Theta)$ *whenever* $\partial\mathbf{F}(\Theta)\cap\mathscr{A}\neq\emptyset$.

(6.8) *For every fixed* $p\in(0,1]$, *the inequality*

$$\tilde{D}(p,p)+\tilde{D}(1-p,1-p)<\tilde{D}(z,p)+\tilde{D}(z',1-p)$$

holds for all $(z,z')\in[0,1]\times[0,1]$ *with* $z+z'<1$.

Then an MCE exists if and only if conditions (5.6), (6.4) *and* (6.5) *are satisfied.*

PROOF. Because of Theorem 6.1, it suffices to show the "only if" part of the theorem. Suppose that (5.6) is not satisfied. Notice that (6.8) with $p=1$ is equivalent to (5.8). From this and (5.7), we can show, by an argument similar to the proof of Theorem 5.2, that $M_b < L(\mathbf{z})$ for all $\mathbf{z} \in \mathrm{F}(\Theta)$. Hence, by Theorem 3.1, an MCE does not exist. Suppose that (6.5) is not satisfied, i.e., $\partial\mathrm{F}(\Theta) \cap \mathscr{A}_k \neq \emptyset$ and $\sum_{j=1}^{k-1} p_{1j-1j} + \sum_{i=k+1}^{m} p_{1ii+1} = 0$ for some k, $1 \leq k \leq m$. Hence $p_{1k-1k} + p_{1kk+1} = 1$. From (3.1), with $\mathbf{z} = (z_1, \ldots, z_m) \in \mathrm{F}(\Theta)$ and $p = p_{1k-1k}$,

$$L(\mathbf{z}) = \begin{cases} S(n_1)\tilde{D}(z_k - z_{k-1}, 1) & \text{if } p = 1, \\ S(n_1)\tilde{D}(z_k - z_{k-1}, p) + \tilde{D}(z_{k+1} - z_k, 1-p) & \text{if } 0 < p < 1, \\ S(n_1)\tilde{D}(z_{k+1} - z_k, 1) & \text{if } p = 0. \end{cases}$$

Consider the case $0 < p < 1$. By (5.7), $(z_k - z_{k-1}) + (z_{k+1} - z_k) < 1$. Hence, by (6.8),

$$\begin{aligned} L(\mathbf{z}) &= S(n_1)(\tilde{D}(z_k - z_{k-1}, p) + \tilde{D}(z_{k+1} - z_k, 1-p)) \\ &> S(n_1)(D(p, p) + D(1-p, 1-p)) \\ &= L(\mathbf{a}_k(p)). \end{aligned}$$

On the other hand, (6.6) implies that $\mathbf{a}_k(p) \in \partial\mathrm{F}(\Theta)$. Thus $M_b < L(\mathbf{z})$ for all $z \in \mathrm{F}(\Theta)$, and an MCE does not exist. Consider the case $p=1$. By (5.7), $z_k - z_{k-1} < 1$. This and (6.8) with $p=1$ yield that $L(\mathbf{z}) > S(n_1)D(1,1) = L(\mathbf{a}_k(1))$. From (5.7), it follows that $\mathbf{a}_k(1) \notin \mathrm{F}(\Theta)$. Noting that $\mathbf{a}_k(1) \in \mathscr{A}_k \in \overline{\partial\mathrm{F}(\Theta)}$, we see that $\mathbf{a}_k(1) \in \partial\mathrm{F}(\Theta)$. Thus $M_b < L(\mathbf{z})$ for all $\mathbf{z} \in \mathrm{F}(\Theta)$, and an MCE does not exist. By the same way as before, the non-existence of an MCE for the case $p=0$ follows. Suppose that (6.4) is not satisfied. From (3.1), with $\mathbf{z} = (z_1, \ldots, z_m) \in \mathrm{F}(\Theta)$ and $p = p_{101}$,

$$L(\mathbf{z}) = \begin{cases} S(n_1)\tilde{D}(z_1, 1) & \text{if } p = 1, \\ S(n_1)\tilde{D}(z_1, p) + \tilde{D}(z_m, 1-p) & \text{if } 0 < p < 1. \\ S(n_1)\tilde{D}(1 - z_m, 1) & \text{if } p = 0. \end{cases}$$

By the same reasoning as above, we see that an MCE does not exist.

For $\ell \geq 2$, we have

THEOREM 6.3. *Let* $\ell \geq 2$, $\sum_{1 \leq i < j \leq m} p_{\cdot ij} \neq 0$ *and* $\inf\{L(\mathbf{z}); \mathbf{z} \in \mathrm{F}(\Theta)\} < \infty$. *Assume that conditions* (5.7) *and* (6.6) *and the following condition are satisfied*:

(6.9) *For every fixed* $p \in (0, 1]$, $D(z, p)$ *is strictly decreasing in* z *on* $(0, 1]$.

Then an MCE exists if and only if conditions (5.6) *and* (6.5) *are satisfied.*

PROOF. The "if" part of the theorem immediately follows from Theorem 6.1. We can show, by the same argument as in the proof of Theorem 5.2, that an

MCE does not exist if (5.6) is not satisfied. Suppose that (6.5) is not satisfied, i.e., $\partial F(\Theta) \cap \mathscr{A}_k \neq \emptyset$ and $\sum_{j=1}^{k-1} p_{\cdot\cdot j} + \sum_{i=k+1}^{m} p_{\cdot i \cdot} = 0$ for some k, $1 \leq k \leq m$. Choose $\mathbf{z} = (z_1, \ldots, z_m) \in F(\Theta)$ so that $L(\mathbf{z}) < \infty$. Then, by (4.3) and (5.7), $\mathscr{N}_1(h; \mathbf{z}) = \emptyset$ for all $h = 1, \ldots, \ell$, and by (4.1) and (4.4), $\mathscr{N}_0(h; \mathbf{z}) = \emptyset$ for all $h = 1, \ldots, \ell$. This, together with the equality $\sum_{j=1}^{k-1} p_{\cdot\cdot j} + \sum_{i=k+1}^{m} p_{\cdot i \cdot} = 0$, yields that $\mathscr{N}(h; \mathbf{z}) = \{(i, j);\ 0 \leq i < k < j \leq m+1 \text{ and } p_{hij} = 1\} \cup \{(i, k);\ 0 \leq i < k \text{ and } p_{hik} \neq 0\} \cup \{(k, j);\ k < j \leq m+1 \text{ and } p_{hkj} \neq 0\} \neq \emptyset$ for all $h = 1, \ldots, \ell$. By (4.1) and (6.9),

$$
\begin{aligned}
L(\mathbf{z}) = & \sum_h S(n_h) \Big(\sum_{0 \leq i < k < j \leq m+1;\, p_{hij} \neq 0} D(z_j - z_i, 1) \\
& + \sum_{k < j \leq m+1;\, p_{hij} \neq 0} D(z_j - z_k, p_{hkj}) \\
& + \sum_{0 \leq i < k;\, p_{hik} \neq 0} D(z_k - z_i, p_{hik}) \Big) \\
> & \sum_h S(n_h) \Big(\sum_{0 \leq i < k < j \leq m+1;\, p_{hij} \neq 0} D(1, p_{hij}) \\
& + \sum_{k < j \leq m+1;\, p_{hij} \neq 0} D(1 - z_k, p_{hkj}) \\
& + \sum_{0 \leq i < k;\, p_{hik} \neq 0} D(z_k, p_{hik}) \Big) \\
= & L(\mathbf{a}_k(z_k)).
\end{aligned}
$$

Because of (5.7) and (6.6), $\mathbf{a}_k(z_k) \in \partial F(\Theta)$ and the inequality $M_b < L(\mathbf{z})$ holds. Hence, by Theorem 3.1, an MCE does not exist.

7. Practical criteria for the 2-regular case (part 11)

In this section we shall give sufficient conditions for the existence of an MCE under the assumption that $K_D(p) = \infty$ for all $p \in (0, 1]$, the family $\mathscr{P}$ is 2-regular, $\partial F(\Theta) \cap \mathscr{A} \neq \emptyset$ and the pooled grouped sample $\mathscr{C}$ is a binary response data sample, i.e., $r_h = 1$ for all $h = 1, \ldots, \ell$. In this case, $\sum_{1 \leq i < j \leq m} p_{\cdot ij} = 0$. Hence, for each h, $1 \leq h \leq \ell$, there exists a unique integer i_h such that $1 \leq i_h \leq m$ and $p_{h0i_h} + p_{hi_hm+1} = 1$. Throughout this section, we put $p_h = p_{h0i_h}$ and $q_h = 1 - p_h$. From (3.1),

$$
(7.1) \qquad L(\mathbf{z}) = \sum_{1 \leq h \leq \ell;\, p_h \neq 0} S(n_h) D(z_{i_h}, p_h) + \sum_{1 \leq h \leq \ell;\, q_h \neq 0} S(n_h) D(1 - z_{i_h}, q_h).
$$

For simplicity we put $L(z) = L(\mathbf{z})\,(\mathbf{z} = (z, \ldots, z) \in \partial F(\Theta) \cap \mathscr{A})$.

To find a sufficient condition for the existence of an MCE, consider the following condition:

(A) For each $z \in (0, 1)$ with $z\mathbf{1} \in \partial F(\Theta) \cap \mathscr{A}$, there exist a positive number t_0, a mapping $\rho(t)$ from $(0, t_0)$ into Θ and a positive function $w(t)$ defined on $(0, t_0)$ such that:

$$
(7.2) \qquad \lim_{t \to 0} F(\rho(t)) = z\mathbf{1}.
$$

(7.3) *For each* i, $1 \leq i \leq m$, $F(x_i, \rho(t))$ *is differentiable on* $(0, t_0)$, *and* $W(x_i; z) = \lim_{t\to 0} w(t)\, dF(x_i, \rho(t))/dt$ *exists and is finite.*

We prove

THEOREM 7.1. *Let condition* (*A*) *be satisfied. Assume that*:

(7.4) *There exists* $z^*\mathbf{1} \in \partial\mathbf{F}(\Theta) \cap \mathscr{A}$ *such that* $L(z^*\mathbf{1}) = \min\{L(\mathbf{z}); \mathbf{z} \in \partial\mathbf{F}(\Theta) \cap \mathscr{A}\}$.

(7.5) *For every fixed* $p \in (0, 1]$, $D(z, p)$ *is continuously differentiable on* $(0, 1)$.

Then an MCE exists if conditions (5.6) *and* (6.5), *and the following condition are satisfied*:

(7.6) $$\sum_{1 \leq h \leq \ell;\, p_h \neq 0} S(n_h) D'(z^*, p_h) W(x_{i_h}; z^*) < \sum_{1 \leq h \leq \ell;\, q_h \neq 0} S(n_h) D'(1 - z^*, q_h) W(x_{i_h}; z^*)$$

where $D'(z, p) = dD(z, p)/dz$.

PROOF. Let z^* be that of (7.4). Because of Theorems 3.1, 5.1 and 6.1, it suffices to show that there exist $z\mathbf{1} \in \mathbf{F}(\Theta)$ with $L(z) \leq L(z^*)$. Let t_0 and $\rho(t)$ be those of (A) with z replaced by z^*. From (7.1) and (7.5),

$$\lim_{t\to 0} w(t) dL(F(\rho(t)))/dt = \sum_{1 \leq h \leq \ell;\, p_h \neq 0} S(n_h) D'(p_h, z^*) W(x_{i_h}; z^*)$$
$$- \sum_{1 \leq h \leq \ell;\, q_h \neq 0} S(n_h) D'(q_h, 1 - z^*) W(x_{i_h}; z^*).$$

Hence (7.6) implies that $dL(\mathbf{F}(\rho(t)))/dt < 0$ for sufficiently small $t > 0$. Thus there exists $z\mathbf{1} \in \mathbf{F}(\Theta)$ with $L(z) \leq L(z^*)$. This completes the proof.

Appendix

We shall prove Theorem 3.1. To do this, we prepare a theoretical background. To relate the function $L(\mathbf{z})$ (see (3.1) for its definition) to the contrast function of Problem (*P*) (see Section 2), we regard $\bar{R}$ as a compact metric space with the distance

$$\text{dist}(t, t') = |\arctan t - \arctan t'|, \quad t, t' \in \bar{R},$$

where $\arctan(\pm\infty) = \pm\pi/2$. Let $p \in (0, 1]$ be fixed. Because of (2.2) and of the continuity of $D(z, p)$, $D(z, p)$ can be extended to a continuous function $f_p(z)$ on $[0, 1]$ to $\bar{R}$ by $f_p(z) = \lim_{n\to\infty} D(z_n, p)$, where $\{z_n\}$ is a sequence in $(0, 1]$ with the limit $z \in [0, 1]$ (cf. Bourbaki (1965, chap. 1)). From the definition of $\tilde{D}(z, p)$ (cf. Section 3), we see that $\tilde{D}(z, p) = f_p(z)$ on $[0, 1]$. It is easily verified that $L(\mathbf{z})$ is continuous on $\bar{Z}$ and $L(\mathbf{F}(\theta))$ is equal to the contrast function of Problem (*P*) up to a constant. Hence an MCE exists if and only if $L(\mathbf{z})$ attains its minimum on $\mathbf{F}(\Theta)$.

PROOF OF THEOREM 3.1. Put $M_c = \inf\{L(\mathbf{z});\, \mathbf{z} \in \overline{\mathbf{F}(\Theta)}\}$ We show

$$M = M_c \leq M_b. \tag{1}$$

Because of the relation $\min(M, M_b) \geq M_c$, it suffices to prove that $M \leq M_c$ in case $M_c < \infty$. Let $t > M_c$ and choose $\hat{\mathbf{z}} \in \overline{\mathbf{F}(\Theta)}$ with $L(\hat{\mathbf{z}}) = M_c$. Since $L(\mathbf{z})$ is continuous on $\overline{\mathbf{F}(\Theta)}$, there exists a neighborhood $\mathscr{U}$ of $\hat{\mathbf{z}}$ such that $L(\mathbf{z}) < t$ for all $\mathbf{z} \in \mathscr{U}$. Since $\mathbf{F}(\Theta) \cap \mathscr{U} \neq \emptyset$, $M < t$. By letting $t \to M_c$, we have $M \leq M_c$. The relation (1) proves the "only if" part of the theorem. Let $\mathbf{z}'$ be that of (3.2) (see Section 3). If $L(\mathbf{z})$ attains its minimum at $\mathbf{z} = \mathbf{z}'$, then an MCE exists. Assume that $M < L(\mathbf{z}')$. From (1), $L(\hat{\mathbf{z}}) = M$ and $M < L(\mathbf{z}') \leq M_b$. This implies that $\hat{\mathbf{z}} \in \mathbf{F}(\Theta)$ and $L(\mathbf{z})$ attains its minimum at $\mathbf{z} = \hat{\mathbf{z}}$. Hence an MCE exists.

References

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

[2] Bourbaki, N. (1965). Topologie Générale, Chap. 1 et 2. Hermann, Paris.

[3] Carter, W. H., Bowen, J. V., and Myers, R. H. (1971). Maximum likelihood estimation from grouped Poisson data. J. Amer. Statist. Ass. **66**, 351–353.

[4] Grossman, W. (1982). On the asymptotic properties of minimum contrast estimates. Probability and Statistical Inference, 115–127.

[5] Kulldorff, G. (1962). Contribution to the theory of estimation from grouped and partially grouped samples. John Wiley and Sons, New York.

[6] Nakamura, T. and Kariya, T. (1975). On the weighted least squares estimation and the existence theorem of its optimal solution. Kawasaki Medical J. (Liberal Arts and Science Course), **1**, 1–11.

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

[8] Nakamura, T. (1985a). The probability contents boundary analysis. In Statistical Theory and Data Analysis (ed. K. Matushita), 485–497. North-Holland.

[9] Nakamura, T. (1985b). The probability contents inner boundary of an interval-censored data sample for families of distributions. Technical Report Series of Okayama Statisticians Group, No. **1**, Kawasaki Medical School, Japan.

[10] Nakamura, T. (1985c). Probability contents inner boundary of an interval-censored data. Keio Science and Technology Reports, **38**, 1–13.

[11] Nakamura, T. (1990). Existence of maximum likelihood estimates for interval-censored data from some three-parameter models with shifted origin. J. R. Statist. Soc. B, **52**, in press.

[12] Rao, C. R. (1955). Theory of the method of estimation by minimum chi-square. Bull. Inst. Inter. Statist. **35**, 25–32.

[13] Rao, C. R. (1957). Maximum likelihood estimation for the multinomial distribution. Sankhya, **18**, 139–148.