Mem. Fac. Sci. Shimane Univ.,
26, pp. 189–200 Dec. 25, 1992

On the Goodness of a Criterion for the Existence of MLE's Based on Interval-censored Data from Some Three-parameter Distribution with a Shifted Origin

Tadashi NAKAMURA and Yasuhisa HIRAI

Department of Information Science, Faculty of Science, Shimane University
Matsue, 690 *JAPAN*

and

Faculty of Education, Okayama University
Okayama, 700 *JAPAN*

(Received September 10, 1992)

By a method called the probability contents boundary analysis, Nakamura [1991] derived a criterion for the existence of a maximum likelihood estimate when observations are interval-censored. We show that this criterion is good if the Hessian matrix of the log-likelihood is negative definite on a convex subset of the parameter space. A detaild discussion on the log-likelihood is also given.

1. Introduction

Let $F(x)$ be a twice continuously differentiable distribution function (d.f.) on the real line $\mathfrak{R}$ with positive and continuously differentiable density function $f(x)$ and let $\mathfrak{R}_+ = (0, \infty)$, $\bar{\mathfrak{R}} = [-\infty, \infty]$ and $\Theta_0 = \mathfrak{R}_+ \times \mathfrak{R} \times [-\infty, \infty)$. Define a transformation $t(x, \theta)(x \in \bar{\mathfrak{R}}; \theta = (\alpha, \beta, \lambda) \in \Theta_0)$ by

$$t(x, \theta) = \begin{cases} -\infty, & x = -\infty \text{ and } \lambda = -\infty, \\ \alpha x - \beta, & x \in \mathfrak{R} \text{ and } \lambda = -\infty, \\ \infty, & x = \infty \text{ and } \lambda = -\infty; \end{cases}$$

$$t(x, \theta) = \begin{cases} -\infty, & x \in [-\infty, \lambda] \text{ and } \lambda \in \mathfrak{R}, \\ \alpha \log(x - \lambda) - \beta, & x \in (\lambda, \infty) \text{ and } \lambda = \mathfrak{R}, \\ \infty, & x = \infty \text{ and } \lambda \in \mathfrak{R}; \end{cases}$$

Here we adopt the rule: $F(-\infty) = 0$ and $F(\infty) = 1$. Consider a family $\mathscr{F}(\Theta) = \{F(t(x, \theta)); \theta \in \Theta\}$, where Θ is a nonempty subset of Θ_0. Suppose that n independent observations $X_1, \cdots, X_n$ have the distribution in $\mathscr{F}(\Theta)$ and that each X_i is known only to lie in a proper subinterval C_i of $\mathfrak{R}$ with nonempty interior. The collection $\mathscr{C} = \{C_1, \cdots, C_n\}$ is called an interval-censored (i.c.) data. Arrange all finite terminal points of $C_i's$ in order of magnitude and denote them by $x_1, \cdots, x_m$. Throughout this paper it is assumed that $m \geq 3$. For the sake of convenience, put $x_0 = -\infty$ and $x_{m+1} = -\infty$. Denote by n_{ij}, $0 \leq i \leq j \leq m+1$, the number of C_k, $1 \leq k \leq n$, such that the small (resp. large) one of extreme points of C_k is x_i(resp. x_j). Then the log-likelihood $l(\theta)$ of the i.c. data $\mathscr{C}$ is expressed as

$$\ell(\theta) = const. + \sum_{j=1}^{m+1} \sum_{i=0}^{j-1} n_{ij} \log(F(t(x_j, \theta)) - F(t(x_i, \theta))).$$

In computing this function, the following rules are used: $\log 0 = -\infty$, $0 \cdot \log 0 = 0$, $(-\infty) + (-\infty) = -\infty$ and $t \cdot (-\infty) = -\infty$. A maximum likeklihood estimate (MLE) $\hat{\theta}$ for the family $\mathscr{F}(\Theta)$ is

$$\hat{\theta} = \operatorname{Arg} \max_{\theta \in \Theta} \ell(\theta).$$

Nakamura [1991] derived a criterion for the existence of an MLE for $\mathscr{F}(\mathfrak{R}_+ \times \mathfrak{R} \times \mathfrak{R})$; that is, to see whether an MLE for $F(\mathfrak{R}_+ \times \mathfrak{R} \times \mathfrak{R})$ exists or not, it suffices to verify the following conditions:

(1.1) $$\sum_{j=1}^{h} n_{\cdot j} + \sum_{j=h+2}^{m} n_{i\cdot} \neq 0, \qquad h = 0, \cdots, m-2.$$

(1.2) $$\sum_{j=1}^{h} n_{\cdot j} + \sum_{h+2 \leq i < j \leq m}^{m} n_{ij} \neq 0, \qquad h = 0, \cdots, m-3.$$

(1.3) $$\sum_{j=1}^{m-1} n_{\cdot j} = 0 \text{ or}$$

(*) $$\sum_{j=2}^{m} \sum_{i=1}^{j-1} n_{ij} \frac{x_j^2 f(\hat{\alpha} x_j - \hat{\beta}) - x_i^2 f(\hat{\alpha} x_i - \hat{\beta})}{F(\hat{\alpha} x_j - \hat{\beta}) - F(\hat{\alpha} x_i - \hat{\beta})}$$
$$< \sum_{i=1}^{m} n_{i\,m+1} \frac{x_i^2 f(\hat{\alpha} x_i - \hat{\beta})}{1 - F(\hat{\alpha} x_i - \hat{\beta})} - \sum_{j=1}^{m} n_{0j} \frac{x_j^2 f(\hat{\alpha} x_j - \hat{\beta})}{F(\hat{\alpha} x_i - \hat{\beta})}$$

Here $n_{\cdot j} = \sum_{i=0}^{j-1} n_{ij}$, $1 \leq j \leq m+1$ and $n_{i\cdot} = \sum_{j=i+1}^{m+1} n_{ij}$, $0 \leq i \leq m$. When n is sufficiently large, the inequality (*) plays an important role to check whether the criterion (conditions (1.1)-(1.3)) is satisfied or not. We say that the criterion is good if there exist a real number $\lambda_0 \in (-\infty, x_1)$ and path $\tau(\lambda) = (\alpha(\lambda), \beta(\lambda), \lambda)$, $\lambda \in (-\infty, \lambda_0)$ such that:

(i) $\{\tau(\lambda); \lambda \in (-\infty, \lambda_0)\} \subset \mathfrak{R}_+ \times \mathfrak{R} \times \mathfrak{R}$.

(ii) $\ell(\tau(\lambda)) = \max\{\ell(\theta); \theta \in \mathfrak{R}_+ \times \mathfrak{R} \times \{\lambda\}\}$, $\lambda \in (-\infty, \lambda_0)$.

(iii) $\alpha(\lambda)$ and $\beta(\lambda)$ are continuously differentiable on $(-\infty, \lambda_0)$.

(iv) The inequality (*) is satified if and only if $\frac{d\ell(\tau(\lambda))}{d\lambda} < 0$ whenever λ is sufficiently near $-\infty$.

The aim of this paper is to show that the criterion is good under the following

(H) condition (H): For every $(m+2)$-tuples $(y_0, y_1, \cdots y_m, y_{m+1})$ with $y_0 = -\infty < y_1 < \cdots < y_m < y_{m+1} = \infty$, the Hessian matrix of the function

$$\sum_{j=1}^{m+1} \sum_{i=0}^{j-1} n_{ij} \log(F(\alpha y_j - \beta) - F(\alpha y_i - \beta))$$

is negative define on $\mathfrak{R}_+ \times \mathfrak{R}$.

The proof of the goodness of the crierion is given in Section 2. A sufficient condition for which condition (H) is satisfied is discussed in Section 3.

2. Goodness of the criterion

Nakamura (1991) derived the inequality (*) by using an artificial path. For this reason the criterion seems to be very restricted one. As is shown later, the criterion turns out to be good. Before proving the goodness of the criterion, we prepare some definitions and notation. Put $\mathbf{F}(\theta) = (F(t(x_1, \theta)), \cdots, F(t(x_m, \theta)))$ and $Z = \{(z_1, \cdots z_m); 0 \leq z_1 \leq \cdots \leq z_m \leq 1\}$. Define $L(\mathbf{z})(\mathbf{z} = (z_1, \cdots, z_m) \in Z)$ by

$$L(\mathbf{z}) = \sum_{j=0}^{m+1} \sum_{i=0}^{j-1} n_{ij} \log(F(t(x_j, \theta)) - F(t(x_i, \theta))),$$

where $z_0 = 0$ and $z_1 = 1$. For notational simplicity, put $\Theta_1 = \mathfrak{R}_+ \times \mathfrak{R} \times \mathfrak{R}$, $\Theta_2 = \mathfrak{R}_+ \times \mathfrak{R} \times \{-\infty\}$ and $\hat{\mathbf{z}}_k = \mathbf{F}(\hat{\theta}_k)$, where $\hat{\theta}_k$ is an MLE for $\mathscr{F}(\Theta_k)$, $k = 1, 2$. We say that the i.c. data $\mathscr{C}$ is lr-censored data if all observations are left-censored or right-censored (see Peto (1973)). Hence the i.c. data $\mathscr{C}$ is an lr-censored data if and only if $\sum_{1 \leq i \leq j \leq m} n_{ij} = 0$.

The following proposition is due to Nakamura (1984).

PROPOSITION 2.1. *Let $\lambda \in [-\infty, x_1)$ and the i.c.data $\mathscr{C}$ be not an lr-censored data. Then an MLE for $\mathscr{F}(\mathfrak{R}_+ \times \mathfrak{R} \times \{\lambda\})$ exists if and only if*

$$\sum_{j=1}^{h} n_{\cdot j} + \sum_{i=h+2}^{m} n_{i\cdot} \neq 0, \qquad h = 0, 1, \cdots, m-1. \tag{2.1}$$

LEMMA 2.1. *Let condition* (H) *be satified, the i.c. data $\mathscr{C}$ be not an lr-censored data and $\lambda_0 \in (-\infty, x_1)$. If condition* (2.1) *is satisfied, then there exists the unique path $\tau(\lambda) = (a(\lambda), b(\lambda), \lambda)$, $\lambda \in (-\infty, \lambda_0)$ such that:*

(i) $\{\tau(\lambda); \lambda \in (-\infty, \lambda_0)\} \subset \Theta_1$.

(ii) $L(\mathbf{F}(\tau(\lambda))) = \max\{L(\mathbf{z}); \mathbf{z} \in \mathbf{F}(\mathfrak{R}_+ \times \mathfrak{R} \times \{\lambda\})\}$.
(iii) $\alpha(\lambda)$ *and* $\beta(\lambda)$ *are continuously differentiable on* $(-\infty, \lambda_0)$.

PROOF. Choose an arbitrary $\lambda \in (-\infty, \lambda_0)$. By Prposition 2.1, an MLE for $\mathscr{F}(\mathfrak{R}_+ \times \mathfrak{R} \times \{\lambda\})$ exists. This implies that there exists a point of $\mathfrak{R}_+ \times \mathfrak{R} \times \{\lambda\}$ which maximizes the log-likelihood $\ell(\theta)$ over $\mathfrak{R}_+ \times \mathfrak{R} \times \{\lambda\}$. On the other hand, $\ell(\theta)$ is strictly concave on $\mathfrak{R}_+ \times \mathfrak{R} \times \{\lambda\}$, since condition (H) is satisfied (cf. [1; Theorem II]). Hence such a maximizing points is unique, and is denoted by $\tau(\lambda) = (\alpha(\lambda), \beta(\lambda), \lambda)$. From its definition, (i) and (ii) follow. We show (iii). Define $d(\theta) = (\partial\ell(\theta)/\partial\alpha, \partial\ell(\theta)/\partial\beta)$, $\theta = (\alpha, \beta, \lambda) \in \Theta_1$, and consider the equation $d(\theta) = 0$. By Theorem 1 of [12], this equation has the solution. Hence $\tau(\lambda)$ is the unique stationary point of $\ell(\theta)$. Put $H(\theta) = (a_{ij}(\theta))$, where $a_{11}(\theta) = \partial^2\ell(\theta)/\partial\alpha^2$, $a_{12}(\theta) = a_{21}(\theta) = \partial^2\ell(\theta)/\partial\alpha\partial\beta$ and $a_{22}(\theta) = \partial^2\ell(\theta)/\partial\beta^2$. Condition (H) yields that $\det(H(\theta)) \neq 0$ at $\theta = \tau(\lambda)$. By the implicit function theorem, there exist unique continuously differentiable functions $a(t)$ and $b(t)$ in a neighborhood U of λ such that $d(a(t), b(t), t) = 0$ for all $t \in U$ and $(a(\lambda), b(\lambda)) = (\alpha(\lambda), \beta(\lambda))$. The uniqueness of the stationary point yields that $a(t) = \alpha(t)$ and $b(t) = \beta(t)$ on U which proves (iii).

REMARK 2.1. Let us make the following condition (H*):

(H*) For every $(m+1)$-tuple $(y_0, y_1, \ldots, y_m, y_{m+1})$ with $y_0 = -\infty < y_1 < \cdots < y_m < y_{m+1} = \infty$, the Hessian matrix of

$$\sum_{j=1}^{m+1} \sum_{i=0}^{j-1} n_{ij} \log(F(\alpha y_j - \beta) - F(\alpha y_i - \beta))$$

is negative definite at every stationary point $(\alpha, \beta) \in \mathfrak{R}_+ \times \mathfrak{R}$ of this function. With the aid of Theorem 2.1 of Mäkeläinen et. al. (1981), or Theorem 1 of Barndoff-Nielsen and Blasild (1980), or Proposition 1 of Gabrielsen (1982), we can prove Lemma 2.1 under a slightly weaker condition (H*).

LEMMA 2.2. *Assume that condition* (H) *is satisfied and the i.c. data* $\mathscr{C}$ *is not an lr-censored data. Let* $\tau(\lambda) = (\alpha(\lambda), \beta(\lambda), \lambda)$ *be as in the previous lemma. Assume that conditions* (1.1), (1.2) *and* (2.1) *are satisfied. Then*

$$\lim_{\lambda \to -\infty} t(x_i, \tau(\lambda)) = t(x_i, (\hat{\alpha}, \hat{\beta}, -\infty)), \qquad 1 \leq i \leq m,$$

where $\hat{\alpha}$ and $\hat{\beta}$ *are uniquely determined by* $\mathbf{F}((\hat{\alpha}, \hat{\beta}, -\infty)) = \hat{\mathbf{z}}_2$.

PROOF. We show that

$$(2.2) \qquad -\infty < L(\hat{z}_2) \leq \liminf_{\lambda \to -\infty} L(\mathbf{F}(\tau(\lambda))).$$

Define a path $\rho(s) = (\alpha(s), \beta(s), \lambda(s))$ from $(0, \min_{1 \leq i \leq m} 1/|x_i|)$ into $\mathfrak{R}_+ \times \mathfrak{R} \times \mathfrak{R}$ by

$$\alpha(s) = \hat{\alpha}/s,\ \beta(s) = \hat{\beta} - \alpha(s)\log s,\ \lambda(s) = -1/s.$$

It is easy to see that

$$t(x_i, \rho(-\lambda^{-1})) = \hat{\alpha}x_i - \hat{\beta} + O(-\lambda^{-1}), \qquad 1 \le i \le m.$$

Hence $\lim_{\lambda\to-\infty} t(x_i, \rho(-\lambda^{-1})) = \hat{\alpha}x_i - \hat{\beta}$, $1 \le i \le m$. This derives that $\lim_{\lambda\to-\infty} L(\mathbf{F}(\rho(-\lambda^{-1}))) = L(\hat{z}_2)$. While, $L(\mathbf{F}(\rho(-\lambda^{-1}))) \le L(\mathbf{F}(\tau(\lambda)))$ for all $\lambda \in (-\infty, x_1)$. We show that

(2.3) there exists no sequence $\{\lambda_n\}$ in $(-\infty, x_1)$ with $\lim_n \lambda_n = -\infty$ such that $\lim_n \mathbf{F}(\tau(\lambda_n)) \in \mathbf{F}(\Theta_1)$.

Suppose the contray, and put $\lim_n \mathbf{F}(\tau(\lambda_n)) = (z_1, \ldots, z_m)$. Consider the case $0 < z_i \le z_j < 1$ for some pair (i, j) with $1 \le i \le j \le m$. Then $\lim_n t(x_i, \tau(\lambda_n)) = F^{-1}(z_i)$ and $\lim_n t(x_j, \tau(\lambda_n)) = F^{-1}(z_j)$. By L'Hospital rule,

$$\lim_n t(x, \tau(\lambda_n)) = F^{-1}(z_i) + \frac{x - x_i}{x_j - x_i}(F^{-1}(z_j) - F^{-1}(z_i))$$

for all $x \in \mathfrak{R}$. This is a contradiction, since $\mathbf{F}(\Theta_1) \cap \mathbf{F}(\Theta_2) = \emptyset$. The rest possible cases are $z_1 = \cdots = z_m = 0$ and $z_{m-1} = 0 < z_m < 1$. By (2.1), $\lim_n L(\mathbf{F}(\tau(\lambda_n))) = -\infty < L(\hat{z}_2)$ for both cases, which contradicts (2.2). Hence (2.3) holds. Let $\{\lambda_{1n}\}$ and $\{\lambda_{2n}\}$ be sequences in $(-\infty, x_1)$ with $\lim_n \lambda_{1n} = \lim_n \lambda_{2n} = -\infty$ such that $\lim_n \mathbf{F}(\tau(\lambda_{1n})) = \mathbf{z}'$ and $\lim_n \mathbf{F}(\tau(\lambda_{2n})) = \mathbf{z}''$. From (2.3) it follows that $\mathbf{z}', \mathbf{z}'' \in \overline{\mathbf{F}(\Theta_1)} - \mathbf{F}(\Theta_1)$. This, together with Proposition 2 and Theorem 1 of [9], (1.1) and (1.2), yields that $\mathbf{z}', \mathbf{z}'' \in \mathbf{F}(\Theta_2)$. Since $\hat{\mathbf{z}}_2$ is the unique maximizing point $L(\mathbf{z})$ over $\mathbf{F}(\Theta_2)$, we see that $\mathbf{z}' = \mathbf{z}'' = \hat{\mathbf{z}}_2$. Thus $\lim_{\lambda\to-\infty} \mathbf{F}(\tau(\lambda)) = \hat{\mathbf{z}}_2$, which is equivalent to $\lim_{\lambda\to-\infty} t(x_i, \tau(\lambda)) = \hat{\alpha}x_i - \hat{\beta} = t(x_i, (\hat{\alpha}, \hat{\beta}, -\infty))$, $1 \le i \le m$.

Hereafter we assume that condition (H) and conditions (1.1), (1.2) and (2.1) are satisfied. Then by Lemmas 2.1 and 2.2, there exists a unique path $\tau(\lambda) = (\alpha(\lambda), \beta(\lambda), \lambda)$, $\lambda \in (-\infty, x_1)$, such that (i)-(iii) in Lemma 2.1 are satisfied and $\lim_{\lambda\to-\infty} t(x_i, \tau(\lambda)) = \hat{\alpha}x_i - \hat{\beta}$, $1 \le i \le m$. Here, $\hat{\alpha}$ and $\hat{\beta}$ satisfy $\mathbf{F}((\hat{\alpha}, \hat{\beta}, -\infty)) = \hat{\mathbf{z}}_2$. For notational simplicity, put $F_i(\lambda) = F(t(x_i, \tau(\lambda)))$, $f_i(\lambda) = f(t(x_i, \tau(\lambda)))$, $1 \le i \le m$, and

$$a_{ij}^{(k)}(\lambda) = \begin{cases} \dfrac{f_j(\lambda)x_j^k}{F_j(\lambda)}, & i = 0;\ 1 \le j \le m;\ 0 \le k \le 2, \\[2ex] \dfrac{f_j(\lambda)x_j^k - f_i(\lambda)x_i^k}{F_j(\lambda) - F_i(\lambda)}, & 1 \le i \le j \le m;\ 0 \le k \le 2, \\[2ex] \dfrac{-f_i(\lambda)x_i^k}{1 - F_i(\lambda)}, & 1 \le i \le m;\ j = m+1;\ 0 \le k \le 2; \end{cases}$$

$$b_{ij}(\lambda)=\begin{cases}\dfrac{f_j(\lambda)\log(x_j-\lambda)}{F_j(\lambda)}, & i=0;\ 1\le j\le m,\\[2ex] \dfrac{f_j(\lambda)\log(x_j-\lambda)-f_i(\lambda)\log(x_i-\lambda)}{F_j(\lambda)-F_i(\lambda)} & 1\le i\le j\le m,\\[2ex] \dfrac{-f_i(\lambda)\log(x_i-\lambda)}{1-F_i(\lambda)}, & 1\le i\le m;\ j=m+1;\end{cases}$$

$$c_{ij}^{(k)}(\lambda)=\begin{cases}\dfrac{\dfrac{f_j(\lambda)}{1+x_js}}{F_j(\lambda)}s^k, & i=0;\ 1\le j\le m;\ 0\le k\le 1,\\[3ex] \dfrac{\dfrac{f_j(\lambda)}{1+x_js}-\dfrac{f_i(\lambda)}{1+x_is}}{F_j(\lambda)-F_i(\lambda)}s^k, & 1\le i\le j\le m;\ 0\le k\le 1,\\[3ex] \dfrac{\dfrac{f_i(\lambda)}{1+x_is}}{1-F_i(\lambda)}s^k, & 1\le i\le m;\ j=m+1;\ 0\le k\le 1,\end{cases}$$

where $s=-1/\lambda$.

We prove

THEOREM 2.1. *The first few terms of Taylor's expansion of* $L(\mathbf{F}((\tau(\lambda)))/d\lambda$ *is given by*

$$\frac{dL(\mathbf{F}(\tau(\lambda))}{d\lambda}=-\frac{\alpha}{2}\sum_{j=1}^{m+1}\sum_{i=0}^{j-1}n_{ij}a_{ij}^{(2)}(\lambda)\lambda^{-2}+O(\lambda^{-3}). \tag{2.4}$$

PROOF. Put $L_\alpha(\theta)=\partial L(\mathbf{F}(\theta))/\partial\alpha$ and $L_\beta(\theta)=\partial L(\mathbf{F}(\theta))/\partial\beta$. Then $L_\alpha(\tau(\lambda))=L_\beta(\tau(\lambda))=0$ for all $\lambda\in(-\infty,\lambda_0)$, $L_\alpha(\tau(\lambda))=\sum_{j=1}^{m+1}\sum_{i=0}^{j-1}n_{ij}b_{ij}(\lambda)$ and $L_\beta(\tau(\lambda))=-\sum_{j=1}^{m+1}\sum_{i=0}^{j-1}n_{ij}a_{ij}^{(0)}(\lambda)$. It is easy to see that

$$\begin{aligned}\frac{dL(\mathbf{F}(\tau(\lambda)))}{d\lambda}&=L_\alpha(\tau(\lambda))\alpha'(\lambda)+L_\beta(\tau(\lambda))\beta'(\lambda)-\alpha(\lambda)\sum_{j=1}^{m+1}\sum_{i=0}^{j-1}n_{ij}c_{ij}^{(1)}(\lambda)\\ &=-\alpha(\lambda)\sum_{j=1}^{m+1}\sum_{i=0}^{j-1}n_{ij}c_{ij}^{(1)}(\lambda)\\ &=-\alpha(\lambda)s\sum_{j=1}^{m+1}\sum_{i=0}^{j-1}n_{ij}c_{ij}^{(0)}(\lambda).\end{aligned}\tag{2.5}$$

We show

(2.6) $$\alpha(\lambda)s = \hat{\alpha} + O(s).$$

Note that $t(x_j, \tau(\lambda)) - t(x_i, \tau(\lambda)) = \alpha(\lambda)\log\frac{1 + x_j s}{1 + x_i s} = (x_j - x_i)\alpha(\lambda)s(1 + O(s))$. On the other hand, $\lim_{\lambda\to-\infty}(t(x_j, \tau(\lambda)) - t(x_i, \tau(\lambda))) = (x_j - x_i)\hat{\alpha}$ by Lemma 2.2. Combining these facts we obtain (2.6). It can be easily seen that

$$\sum_{j=1}^{m+1}\sum_{i=0}^{j-1} n_{ij}c_{ij}^{(0)}(\lambda) = \sum_{j=1}^{m+1}\sum_{i=0}^{j-1} n_{ij}(\alpha_{ij}^{(0)}(\lambda) - a_{ij}^{(1)}(\lambda) + a_{ij}^{(2)}(\lambda) + O(s^3))$$

(2.7) $$= -L_\beta(\tau(\lambda)) - s\sum_{j=1}^{m+1}\sum_{i=0}^{j-1} n_{ij}a_{ij}^{(1)}(\lambda) + s^2\sum_{j=1}^{m+1}\sum_{i=0}^{m+1} n_{ij}a_{ij}^{(2)}(\lambda) + O(s^3)$$

$$= -s\sum_{j=1}^{m+1}\sum_{i=0}^{j-1} n_{ij}a_{ij}^{(1)}(\lambda) + s^2\sum_{j=1}^{m+1}\sum_{i=0}^{j-1} n_{ij}a_{ij}^{(2)}(\lambda) + O(s^3).$$

From the expression

$$x_i s = \log s + \log(x_i - \lambda) + \frac{x_i^2 s^2}{2} + O(s^3), \qquad 1 \le i \le m,$$

it follows that

(2.8) $$s\sum_{j=1}^{m+1}\sum_{i=0}^{j-1} n_{ij}a_{ij}^{(1)}(\lambda) = L_\alpha(\tau(\lambda)) - L_\beta(\tau\lambda))\log s + \frac{s^2}{2}\sum_{j=1}^{m+1}\sum_{i=0}^{j-1} n_{ij}a_{ij}^{(2)}(\lambda) + O(s^3)$$

$$= \frac{s^2}{2}\sum_{j=1}^{m+1}\sum_{i=0}^{j-1} n_{ij}a_{ij}^{(2)}(\lambda) + O(s^3).$$

Conbining (2.5)-(2.8), we obtain (2.4). This completes the proof.

It can be easily seen that inequality (∗) is equivalent to

$$\lim_{\lambda\to-\infty}\sum_{j=1}^{m+1}\sum_{i=0}^{j-1} n_{ij}a_{ij}^{(2)}(\lambda) < 0.$$

The expression (2.4) implies that the sign of the left-hand side of the above inequality determines the behavior of $L(\mathrm{F}(\tau(\lambda)))$ when λ is near $-\infty$. Because of the optimality of $\tau(\lambda)$ (see Lemma 2.1), the criterion is good, though it is derived by using a special path (see the proof of Theorem 3 in Nakamura (1991)).

3. Sufficient condition for condition (H)

In this section we shall give, by using the log-concavity of the density function $f(x)$, a sufficient condition for which condition (H) is satisfied. Put $g(x) = f'(x)/f(x)$.

The density function $f(x)$ is said to be *strictly log-concave* if $\log f(x)$ is strictly concave. The concavity of the log-likelihood was discussed briefly in Burridge (1981).

The following Lemmas 3.1-3.3 are fundamental to our purpose.

LEMMA 3.1. *Let $f(x)$ be strictly log-concave on $\mathfrak{R}$. Then*

$$f(x) \leq f(y)\exp(g(y)(x-y)) \tag{3.1}$$

for every $(x, y) \in \mathfrak{R}^2$. Moreover, the equality in the above inequality holds if and only if $x = y$.

PROOF. Let $x \neq y$ and put $A(x, y) = \dfrac{\log(f(x)/f(y))}{x-y}$. It is well-known that $A(z, y)$ is strictly decreasing in z on (y, ∞)(see [13; p.2]). Hence $A(x, y) < \lim_{z \to y} A(z, y) = g(y)$. From this,

$$f(x) < f(y)\exp(g(y)(x-y))$$

for all $(x, y) \in \mathfrak{R}^2$ with $y < x$. It is also well known that $A(z, y)$ is strictly decreasing in z on $(-\infty, y)$(see [13; p.2]). Therefore $A(x, y) > \lim_{z \to y} A(z, y) = g(y)$ and hence

$$f(x) < f(y)\exp(g(y)(x-y))$$

for all $(x, y) \in \mathfrak{R}^2$ with $x < y$. This completes the proof.

The following result is due to Berstein and Toupin (1962).

PROPOSITION 3.1. *Let S be a convex domain in $\mathfrak{R}^p$ and $w: \mathfrak{R} \to \mathfrak{R}^q$ be a strictly convex function and twice continuously differentiable on S. Then the Hessian matrix of w is positive definite except on a nowhere dense subset of S.*

REMARK 3.1. Let V be a topological space and W be a subset of V. The set W is called a *border set* if $V - W$ is dense in V. The set W is said to be *nowhere dense* $\bar{W}$ is a border set.

LEMMA 3.2. *Let $f(x)$ be strictly log-concave on $\mathfrak{R}$. Then*:

(i) $g(x)(F(y) - F(x)) + f(x) > 0$ *for all (x, y) with* $-\infty < x < y \leq \infty$.

(ii) $g(y)(F(y) - F(x)) - f(y) < 0$ *for all (x, y) with* $-\infty \leq x \leq y < \infty$.

(iii) $g(x)g(y)(F(y) - F(x)) < g(x)f(y) - g(y)f(x)$ *for all (x, y) with* $-\infty < x < y < \infty$.

PROOF. Proof of (i): If $f'(x) \geq 0$, then the inequality (i) is Assume that $f'(x) < 0$. Lemma 3.1 gives

$$F(y)-F(x)=\int_x^y f(t)dt<f(x)\int_x^y \exp(g(x)(t-x))dt=\frac{f(x)(\exp(g(x)(y-x))-1)}{g(x)}.$$

From this, $g(x)(F(y)-F(x))+f(x)>f(x)\exp(g(x)(y-x))\geq 0$. Hence (i) is established.

Proof of (ii): If $f'(y)\leq 0$, then (ii) is obvious. Assume that $f'(y)>0$. Lemma 3.1 gives

$$F(y)-F(x)=\int_x^y f(t)dt<f(y)\int_x^y \exp(g(y)(t-y))dt=\frac{f(x)(1-\exp(g(y)(y-x)))}{g(y)}.$$

From this, $g(y)(F(y)-F(x))-f(x)<-\exp(g(y)(x-y))\leq 0$. Hence (ii) is proved.

Proof of (iii): Choose an arbitrary $x\in\mathfrak{R}$ and fix it. Put $A(y)=g(x)f(y)-g(y)f(x)-g(x)g(y)(F(y)-F(x))$, $y\in[x,\infty)$. It is obvious that $A(x)=0$. We show that $A(y)>0$ on (x,∞). Since $f(x)$ is strictly log-concave, $g(y)$ is strictly decreasing (see [13; p.5]), and hence $g'(y)\leq 0$. This and (i) in Lemma 3.2 derive that $A'(y)=-g'(y)(f(x)+g(x)(F(y)-F(x)))\geq 0$. Hence $A(y)$ is nondecreasing on (x,∞). Assume that there exists $y'\in(x,\infty)$ such that $A(y')=0$. Since $A(y)$ is nondecreasing (x,∞) and $A(x)=0$, $A(y)=0$ on $[x,y']$ and hence $A'(y)=0$ on $[x,y']$. From this and (i) in Lemma 3.2, $g'(y)=0$ on $[x,y']$. Noting that $g'(y)=d^2\log f(y)/dy^2$, we see that this contradicts the assertion of Proposition 3.1. Now the positiveness of $A(y)$ on (x,∞) is proved. This completes the proof.

REMARK 3.2. Feller (1957, Lemma 2 in Chap. 12) proved the inequality (iii) in the case where $F(x)$ is the standard normal distribution function (see also Haberman (1974, p.308-309)).

LEMMA 3.3. *Let $f(x)$ be strictly log-concave on $\mathfrak{R}$. Then:*

(i) *The second derivative of $\log F(x)$ is negative on $\mathfrak{R}$.*

(ii) *The second derivative of $\log(1-F(x))$ is negative on $\mathfrak{R}$.*

(iii) *The Hessian matrix of $\log(F(y)-F(x))$ is negative define on the set $\{(x,y)\in\mathfrak{R}^2: x<y\}$.*

PROOF. Proof of (i): From $d^2\log F(x)/dx^2=f(x)F(x)^{-2}(g(x)F(x)-f(x))$ and (ii) in Lemma 3.2, our assertion follows.

Proof of (ii): From $d^2\log(1-F(x))/dx^2=-f(x)(1-F(x))^{-2}(g(x)(1-F(x))+f(x))$ and (i) in Lemma 3.2, our assertion follows.

Proof of (iii): It can be easily calculated that

$$A=\partial^2\log(F(y)-F(x))/\partial x^2=-f(x)(F(y)-F(x))^{-2}(g(x)(F(y)-F(x))+f(x)),$$

$$B=\partial^2\log(F(y)-F(x))/\partial x\partial y=-f(x)f(y)(F(y)-F(x))^{-2},$$

$$C=\partial^2\log(F(y)-F(x))/\partial y^2=f(x)(F(y)-F(x))^{-2}(g(y)(F(y)-F(x))-f(y)).$$

By (i) and (ii) in Lemma 3.2, A < 0 and $C < 0$. The relation $AC - B^2 = f(x)f(y)(F(y) - F(x))^{-3}(g(x)f(y) - g(y)f(x) - g(x)g(y)(F(y) - F(x)))$ and (iii) in Lemma 3.2 yield that $AC - B^2 > 0$. Hence that assertion (iii) is established.

Let $y_0 = -\infty < y_1 < \cdots < y_m < y_{m+1} = \infty$. For every pair (i, j) of integers with $0 \leq i \leq j \leq m+1$ and for every $\theta = (\alpha, \beta) \in \mathfrak{R}_+ \times \mathfrak{R}$, define $F_{ij}(\theta)$ by $F_{ij}(\theta) = F(\alpha y_j - \beta) - F(\alpha y_i - \beta)$.

We prove

THEOREM 3.1. *Let $f(x)$ be strictly log-concave. Then*:

(i) *The Hessian matrix of* $\log F_{0j}(\theta)$, $1 \leq j \leq m$, *is not negative definite but negative semi-define.*

(ii) *The Hessian matrix of* $\log F_{im+1}(\theta)$, $1 \leq i \leq m$, *is not negative definite but negative semi-define.*

(iii) *The Hessian matrix of* $\log F_{ij}(\theta)$, $1 \leq i \leq j \leq m$, *is negative definite.*

PROOF. Proof of (i): By (i) in Lemma 3.3, $\log F(x)$ is strictly concave. Hence $\log F_{0j}(\theta) = \log F(\alpha y_j - \beta)$ is concave on $\mathfrak{R}_+ \times \mathfrak{R}$, and thus the Hessian matrix of $\log F_{0j}(\theta)$ is negative semi-definite. Let s be an arbitrary real number and put $D = \{(\alpha, \beta) \in \mathfrak{R}_+ \times \mathfrak{R};\ \alpha y_j - \beta = s\}$. Choose distinct points $\theta_1 = (\alpha_1, \beta_1)$, $\theta_2 = (\alpha_2, \beta_2) \in D$ and put $z_i = \alpha_i y_j - \beta_i$, $i = 1, 2$. It is easy to see that for every $\lambda \in (0, 1)$, $\log F_{0j}(\lambda\theta_1 + (1-\lambda)\theta_2) = \log F(\lambda z_1 + (1-\lambda)z_2) = \lambda \log F(z_1) + (1-\lambda)\log F(z_2) = \lambda \log F_{0j}(\theta_1) + (1-\lambda)\log F_{0j}(\theta_2)$. This implies that $\log F_{0j}(\theta)$ is not strictly concave. Hence (i) is established.

Proof of (ii): By the same argument as above, we can prove (ii).

Proof of (iii): Let $1 \leq i \leq j \leq m$. Put $A(x, y) = \log(F(y) - F(x))$ and $B(\theta) = \log F_{ij}(\theta)$. Denote by $(a_{pq}(x, y))$(resp. $b_{pq}(\theta)$)) the Hessian matrix of $A(x, y)$(resp. $B(\theta)$). It suffices to show that

$$b_{22}(\theta) < 0 \quad \text{and} \quad b_{11}(\theta)b_{22}(\theta) > b_{12}(\theta)^2 \tag{3.2}$$

for all $\theta \in \mathfrak{R}_+ \times \mathfrak{R}$. Note that $B(\theta) = A(\alpha y_j - \beta, \alpha y_j - \beta)$. It is easily calculated that

$$\partial B/\partial\alpha = y_i \partial A/\partial x + y_j \partial A/\partial y,$$

$$\partial B/\partial\beta = -\partial A/\partial x - \partial A/\partial y,$$

$$\partial(\partial A/\partial x)/\partial\alpha = a_{11}(\alpha y_i - \beta, \alpha y_j - \beta)y_i + a_{12}(\alpha y_j - \beta, \alpha y_j - \beta)y_j,$$

$$\partial(\partial A/\partial y)/\partial\alpha = a_{12}(\alpha y_i - \beta, \alpha y_j - \beta)y_i + a_{22}(\alpha y_i - \beta, \alpha y_j - \beta)y_j,$$

$$\partial(\partial A/\partial x)/\partial\beta = -a_{11}(\alpha y_i - \beta, \alpha y_j - \beta) - a_{12}(\alpha y_i - \beta, \alpha y_j - \beta),$$

$$\partial(\partial A/\partial y)/\partial\beta = -a_{12}(\alpha y_i - \beta, \alpha y_j - \beta) - a_{22}(\alpha y_i - \beta, \alpha y_j - \beta).$$

From this

$$b_{11}(\theta) = \hat{a}_{11}(\theta)y_i^2 + 2\hat{a}_{12}(\theta)y_i y_j + \hat{a}_{22}(\theta)y_j^2,$$

$$b_{21}(\theta) = -\hat{a}_{11}(\theta)y_i - \hat{a}_{12}(\theta)(y_i + y_j) - \hat{a}_{22}(\theta)y_j,$$

$$b_{22}(\theta) = \hat{a}_{11}(\theta) + 2\hat{a}_{12}(\theta) + \hat{a}_{22}(\theta),$$

where $\hat{a}_{pq}(\theta) = a_{pq}(\alpha y_i - \beta, \alpha y_j - \beta)$. By Lemma 3.3, $(\hat{a}_{pq}(\theta))$ is negative define. Hence $\hat{a}_{11}(\theta)\hat{a}_{22}(\theta) > \hat{a}_{12}(\theta)^2$, $\hat{a}_{11}(\theta) < 0$ and $\hat{a}_{22}(\theta) < 0$. Since $-\hat{a}_{11}(\theta) - \hat{a}_{22}(\theta) \geq 2(\hat{a}_{11}(\theta))\hat{a}_{12}(\theta))^{1/2} > 2|\hat{a}_{12}(\theta)|$, we see that $b_{22}(\theta) < 0$. The inequality $b_{11}(\theta)b_{22}(\theta) > b_{12}(\theta)^2$ follows from the negative definiteness of $(\hat{\alpha}_{pq}(\theta))$ and the relation $b_{11}(\theta)b_{22}(\theta) - b_{12}(\theta)^2 = (\hat{a}_{11}(\theta)\hat{a}_{22}(\theta) - \hat{a}_{12}(\theta)^2)(y_i - y_j)^2$. This completes the proof.

REMARK 3.3. Theorem 3.2 shows that condition (H) is violated when the i.c. data $\mathscr{C}$ is an lr-censored data, i.e., $\sum_{1 \leq i \leq j \leq m} n_{ij} = 0$.

We now prove the main result in this section.

THEOREM 3.3. *Let $f(x)$ be strictly log-concave on $\mathfrak{R}$. Then condition* (H) *is satisfied if the i.c. data $\mathscr{C}$ is not an lr-censored data.*

PROOF. Let $y_0 = -\infty < y_1 < \cdots < y_m < y_{m+1} = \infty$ and denote by $H_{ij}(\theta)$ the Hessian matrix of $\log F_{ij}(\theta)$, $0 \leq i < j \leq m+1$; $(i, j) \neq (0, m+1)$. By Theorem 3.2, $H_{0j}(\theta)$, $1 \leq j \leq m$ and $H_{im+1}(\theta)$, $1 \leq i \leq m$, are nonnegative definite. Since the i.c. data $\mathscr{C}$ is not an lr-censored data, and since the Hessian matrix of $\sum_{j=2}^{m} \sum_{i=1}^{j-1} n_{ij} \log F_{ij}(\theta)$ is $\sum_{j=2}^{m} \sum_{i=1}^{j-1} n_{ij} H_{ij}(\theta)$, we conclude that condition (H) is satisfied.

Now we shall give examples of d.f.'s having the strictly log-density functions (cf. Burridge (1981), Pratt (1981)).

EXAMPLE 3.1. The standard normal density:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \qquad -\infty < x < \infty.$$

EXAMPLE 3.2. The logistic density:

$$f(x) = \frac{\exp(x)}{(1 + \exp(x))^2}, \qquad -\infty < x < \infty.$$

EXAMPLE 3.3. The external value type of density:

$$f(x) = \Gamma(\gamma)\exp(\gamma x)\exp(-\exp(x)), \qquad -\infty < x < \infty.$$

in which $\gamma > 0$ is known. Note that this may be regarded as the density function of the logarithm of a gamma random variable. When $\gamma = 1$, $f(x)$ is the extreme value of density.

References

[1] Bernstein, B. and Toupin, R. A. (1962). Some properties of the Hessian matrix of a strictly convex functions. J. Reine Angew. Math., **210**, 65–72.
[2] Barndorff-Neilsen, O. and Blasild, P. (1980). Global maxima, and likelihood in linear models. Research Reports, No. **57**, Dept. Theoretical Statist., Univ. Aarhus, Denmark.
[3] Burridge, J. (1981). A note on maximum likelihood estimation for regression models using grouped data. J. R. Statist. Soc. B, **43**, 41–45.
[4] Feller, W. (1957). An introduction to probability theory and its application. New York: Wiley.
[5] Gabrielsen, G. (1982). On the unimodality of the likelihood for the Cauchy distribution: some comments. Biometrika, **69**, 677–688.
[6] Haberman, S. J. (1974). The analysis of frequency data. Univ. of Chicago Press.
[7] Mäkeläinen, K., Schmidt, K. and Styanm G. H. P. (1981). On the existence and uniqueness of the maximum likelihood estiamte of a vector-valued parameter in fixed-size samples. Ann. Statist., **9**, 758–767.
[8] Nakamura, T. (1984). Existence theorems of a maximum likelihood estimate from a generalized censored data sample. Ann. Inst. Statist. Math., **36**, 375–393.
[9] Nakamura, T. (1991). Existence of maximum likelihood estimates for interval-censored data from some three-parameter models with shifted origin. J. R. Statist. Soc. B. **52**, 211–220.
[10] Pratt, J. W. (1981). Concavity of the log likelihood. J. Amer. Statist. Assoc., **76**, 103–106.
[11] Peto, R. (1973). Experimental survival curves for interval-censored data. Appl. Statist., **22**, 86–91.
[12] Rai, K. and Ryzin, J. W. (1982). A note on multivariate version of Roll's theorem and uniqueness of maximum likelihood roots. Commun. Statist.-Theor. Meth., **11**, 1505–1510.
[13] Robert, A. W. and Varberg, D. E. (1973). Convex functions. New York: Academic Press.