Journal of Multivariate Analysis, *1999, vol. 69, 155-166.*

GENERALIZED MLE OF A JOINT DISTRIBUTION FUNCTION WITH MULTIVARIATE INTERVAL-CENSORED DATA

George Y. C. Wong[1] and Qiqing Yu[2]

Strang Cancer Prevention Center, 428 E 72nd Street, NY 10021, USA

and

Department of Mathematical Sciences, SUNY at Binghamton, NY 13902, USA,

First version 6/24/97; Current version 1/28/98

Short Title: Multivariate survival analysis

AMS 1991 subject classification: Primary 62 G05; Secondary 62 G20.

Key words and phrases: multivariate interval-censored data, asymptotic normality, asymptotic variance, consistent estimate, generalized MLE, multivariate survival analysis.

Abstract We consider the problem of estimation of a joint distribution function of a multivariate random vector with interval-censored data. The generalized maximum likelihood estimator of the distribution function is studied and its consistency and asymptotic normality are established under the case 2 multivariate interval censorship model and discrete assumptions on the censoring random vectors.

[1] Partially supported by DAMD17-94-J-4332.
[2] Partially supported by NSF Grant DMS-9402561 and DAMD17-94-J-4332.

1. Introduction

We consider the estimation of a joint distribution function F_0 of a multivariate random vector $\mathbf{X} = (X_1, ..., X_d)$ which is subject to interval censoring. In interval censoring, the value of each coordinate variable X_i may not be directly observable; instead, a pair of extended real numbers L_i and R_i such that $L_i \leq X_i \leq R_i$ are always observed. The observations L_i and R_i satisfy one of the following four conditions: $L_i = R_i$ (exact), $0 = L_i < R_i$ (left censored), $L_i < R_i = \infty$ (right censored), and $0 < L_i < R_i < \infty$ (strictly interval censored). A $d-$dimensional interval-censored observation corresponding to $\mathbf{X}$ is represented by the $2d$-dimensional vector $(L_1, R_1, ..., L_d, R_d)$.

Multivariate interval-censored data arise in a variety of life testing situations and biomedical studies. We describe a clinical study in the following example that gives rise to bivariate ($d = 2$) interval-censored data.

> **Example 1.1** (The Italian-American Cataract Study Group (1994)). A total of 1399 persons, between 45 and 79 years of age, who had been identified in a clinic-based case control study were enrolled in a follow-up study between 1985 and 1988. The follow-up study was designed to estimate the rate of incidence and progression of cortical, nuclear, and posterior subcapsular cataracts and to evaluate the usefulness of the Lens Opacities Classification System II in a longitudinal study. Beginning in 1989, follow-up lens photographs were taken and graded at a six-month interval. Patients might skip some visits. Data were obtained from Zeiss slit-lamp and Neitz retroillumination lens photographs at each patient's visit. The exact time that the event of interest occurred was only known to lie within the period between two consecutive visits, or was right censored if by the end of the study the event still had not taken place. Consequently, bivariate interval-censored data were encountered.

At present, nonparametric estimation of a joint distribution function with multivariate interval-censored data has not been considered. A current practice is to take the midpoint of the interval (L, R) as an exact observation unless it is right censored. Then Dabrowska's (1988) Kaplan-Meier estimator on the plane or van der Laan's (1996) repaired generalized maximum likelihood estimator can be applied to such data. Another practice is to treat the right endpoints of the interval-censored data as exact observations unless they are right censored (see, Samuelsen and Kongerud (1994)). However, these two practices will introduce bias in the analysis (Samuelsen and Kongerud (1994)).

Multivariate right-censored data are special cases of multivariate interval-censored data. References for nonparametric estimation of distribution functions with multivariate right-censored data can be found in Campbell (1981), Hanley and Parnes (1983), Tsai *et al.* (1986), Dabrowska (1988), Gill (1992), Prentice and Cai (1992), Lin and Ying (1993), and van der Laan (1996), etc.

Nonparametric estimation of a distribution function with univariate interval-censored data has been studied by Peto (1973), Turnbull (1976), Tsai and Crowley (1985), Chang and Yang (1987), Groeneboom and Wellner (1992), Gu and Zhang (1993), and Yu *et al* (1996 and 1998), among others.

In Section 2, we discuss generalized maximum likelihood estimation of F_0 based on multivariate interval-censored data and formulate the case 2 multivariate interval censorship model. We establish consistency of the generalized maximum likelihood estimator (GMLE) of F_0 in Section 3 and asymptotic normality of the GMLE in Section 4.

2. Method of estimation

Let $\mathbf{X} = (X_1, ..., X_d)$ be a d-dimensional random survival vector with a joint distribution function $F_0(\mathbf{x})$, where $\mathbf{x} = (x_1, ..., x_d)$. The observable random vector is $(L_1, R_1, ..., L_d, R_d)$, where $L_i \leq R_i$ for all i. Suppose that

$$(L_{11}, R_{11}, ..., L_{1d}, R_{1d}), ..., (L_{n1}, R_{n1}, ..., L_{nd}, R_{nd})$$

are i.i.d. copies of $(L_1, R_1, ..., L_d, R_d)$. We want to estimate the joint distribution function $F_0(\mathbf{x})$ (or the survival function $S_0(\mathbf{x}) = P\{X_1 > x_1, ..., X_d > x_d\}$). Each univariate interval-censored data (L_{ij}, R_{ij}) can be viewed as an interval I_{ij}, where

$$I_{ij} = \begin{cases} [L_{ij}, R_{ij}] & \text{if } L_{ij} = R_{ij}, \\ (L_{ij}, R_{ij}] & \text{if } L_{ij} < R_{ij}; \end{cases}$$

therefore, each multivariate interval-censored observation can be viewed as a rectangular set $\mathcal{I}_i = I_{i1} \times \cdots \times I_{id}$, $i = 1, ..., n$.

Define a *maximal intersection* (MI), A, with respect to the $\mathcal{I}_i$'s to be a nonempty finite intersection of the $\mathcal{I}_i$'s such that for each i, $A \cap \mathcal{I}_i = \emptyset$ or A. For example, let $\mathcal{I}_1 = (0,2] \times (1,3]$, $\mathcal{I}_2 = (0,4] \times (1,5]$, $\mathcal{I}_3 = (3,5] \times (4,8]$ and $\mathcal{I}_4 = (3,5] \times (4,8]$. Then the possible MI's are $(0,2] \times (1,3]$ and $(3,4] \times (4,5]$. Let $\{A_1, ..., A_m\}$ be the collection of all possible distinct MIs.

Using an argument similar to Hanley and Parnes (1983), it can be shown that the GMLE of $F_0(\mathbf{x})$ which maximizes the generalized likelihood function, Λ_n, must assign all the probability masses $s_1, ..., s_m$ to the sets $A_1, ..., A_m$. Thus the generalized likelihood function is as follows:

$$\Lambda_n = \prod_{i=1}^{n} \mu_F(\mathcal{I}_i) = \prod_{i=1}^{n} [\sum_{j=1}^{m} \mathbf{1}(A_j \subset \mathcal{I}_i) s_j], \tag{2.1}$$

where μ_F is the measure induced by a distribution function F, $\mathbf{1}(\cdot)$ is a indicator function, $\mathbf{s}$ $(= (s_1, ..., s_{m-1})^t)$ $\in D_s$, $s_m = 1 - s_1 - \cdots - s_{m-1}$, $\mathbf{s}^t$ is the transpose of the vector $\mathbf{s}$, and $D_s = \{\mathbf{s};\ s_i \geq 0, s_1 + \cdots + s_{m-1} \leq 1\}$. Denote the GMLE of $\mathbf{s}$ by $\hat{\mathbf{s}}$ and that of F_0 by $\hat{F}_n$.

The s_j's can be obtained by the self-consistent algorithm described by Turnbull (1976) for univariate interval-censored data as follows: Let $s_j^{(0)} = 1/m$ for $j = 1, ..., m$. Denote $\delta_{ij} = \mathbf{1}(A_j \subset \mathcal{I}_i)$. At the h-step, $s_j^{(h)} = \sum_{i=1}^{n} \frac{1}{n} \frac{\delta_{ij} s_j^{(h-1)}}{\sum_{k=1}^{m} \delta_{ik} s_k^{(h-1)}}$, $j = 1, ..., m$, $h \geq 1$. Repeat until the s_j's converge. The justification of the convergence of this method for multivariate interval-censored data is similar to that given in Turnbull (1976) for univariate data.

Given a GMLE $\hat{\mathbf{s}}$, the GMLE of $F_0(\mathbf{x})$ is not uniquely defined on an MI unless the MI is a singleton. A GMLE of $F_0(\mathbf{x})$ can be obtained as follows:

$$\hat{F}_n(\mathbf{x}) = \sum_{A_j \subset [0,x_1] \times \cdots \times [0,x_d]} \hat{s}_j. \tag{2.2}$$

Remark 1 *The GMLE of* $\mathbf{s}$ *may not be unique, as the following example demonstrates.*

Suppose that a sample of size 4 consists of two-dimensional interval-censored observations $(1,6,1,3)$, $(1,6,4,6)$, $(1,3,1,6)$ *and* $(4,6,1,6)$. *Then the MIs are* $A_1 = (1,3] \times (1,3]$, $A_2 = (1,3] \times (4,6]$, $A_3 = (4,6] \times (1,3]$ *and* $A_4 = (4,6] \times (4,6]$, *and* $(\hat{s}_1, \hat{s}_2, \hat{s}_3, \hat{s}_4) = r(1/2, 0, 0, 1/2) + (1-r)(0, 1/2, 1/2, 0)$ *is a GMLE of* $\mathbf{s}$, *for all* $r \in [0,1]$. *Thus there are infinitely many expressions for GMLE. However,* $\mu_{\hat{F}_n}(\mathcal{I}_i) = 1/4$, $i = 1, ..., 4$, *for all* $r \in [0,1]$.

In general, $\hat{\mathbf{s}}$ *may not be consistent under discrete assumptions. However, the consistency of* $\hat{F}_n$ *on a certain set will not be affected (for more details, see Section 3).*

The derivation of the GMLE only requires that the observations $\mathcal{I}_1$, ..., $\mathcal{I}_n$ are i.i.d. To derive the asymptotic properties of the GMLE, we need further assumptions on F_0 and the distribution function of $(L_1, R_1, ..., L_d, R_d)$.

A set of univariate interval-censored data are referred to as *case 2* data if they consist of strictly interval-censored, right-censored or left-censored observations, but do not contain exact observations. For such type of data, Groeneboom and Wellner (1992) formulate the case 2 univariate interval censorship model. We consider a natural multivariate extension of the case 2 univariate interval censorship model in the following.

Suppose $(U_1, V_1, ..., U_d, V_d)$ is a random censoring vector and is independent of $\mathbf{X}$. The observable random vector $(L_1, R_1, ..., L_d, R_d)$ is generated by the following formula.

$$(L_i, R_i) = \begin{cases} (0, U_i) & \text{if } X_i \leq U_i, \\ (U_i, V_i) & \text{if } U_i < X_i \leq V_i, \quad i = 1, ..., d. \\ (V_i, +\infty) & \text{if } X_i > V_i, \end{cases}$$

We call this model a *case 2 multivariate interval censorship model* (C2M model). In the next two sections, we shall discuss the asymptotic properties of the GMLE under the C2M model. For ease of presentation and without loss of generality (WLOG), we assume $d = 2$ hereafter.

3. Consistency of GMLE

In this section, we make the following assumptions under the C2M model:

$$\text{The censoring vector } (\mathbf{U}, \mathbf{V}) \text{ is discrete.} \tag{3.1}$$

Let $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$, $\mathbf{U} = (U_1, U_2)$ and $\mathbf{V} = (V_1, V_2)$. Define

$$\mathcal{B} = \{(\mathbf{a}, \mathbf{b}) : g(\mathbf{a}, \mathbf{b}) > 0\}, \text{ where } g(\mathbf{a}, \mathbf{b}) = P(\mathbf{U} = \mathbf{a}, \mathbf{V} = \mathbf{b}).$$

Note that each point in $\mathcal{B}$ induces a grid of nine cells in $\mathcal{R}^2$. Let

$$\mathcal{A}_* = \{(x_1, x_2) : \; x_i \in \{a_i, b_i, \pm\infty\}, \; i = 1, 2, \; (\mathbf{a}, \mathbf{b}) \in \mathcal{B}\}.$$

be the set of all such grid points. We shall establish the strong consistency of the GMLE at each point in $\mathcal{A}_*$. From this we can then infer the uniform strong consistency of the GMLE if F_0 is continuous and $\mathcal{A}_*$ is dense in $[0, \infty)^2$.

Let $(\mathbf{X}_i, \mathbf{U}_i, \mathbf{V}_i)$, $i = 1$, ..., n be i.i.d. copies of $(\mathbf{X}, \mathbf{U}, \mathbf{V})$. For $(\mathbf{a}, \mathbf{b}) \in \mathcal{B}$, let

$$\begin{array}{lll}
I_{11}(\mathbf{a}, \mathbf{b}) = (-\infty, a_1] \times (-\infty, a_2], & \cdots, & \cdots, \\
I_{21}(\mathbf{a}, \mathbf{b}) = (a_1, b_1] \times (-\infty, a_2], & \cdots, & \cdots, \\
I_{31}(\mathbf{a}, \mathbf{b}) = (b_1, +\infty) \times (-\infty, a_2], & \cdots, & I_{33}(\mathbf{a}, \mathbf{b}) = (b_1, +\infty) \times (b_2, +\infty).
\end{array}$$

Let $\mathcal{A}$ be the set of all vertexes of B_1, ..., B_h, where $B_1, \cdots, B_h$ are all possible MIs with respect to $I_{ij}(\mathbf{a}, \mathbf{b})$, i, $j = 1, 2, 3$, and $(\mathbf{a}, \mathbf{b}) \in \mathcal{B}$. Note that $\mathcal{A}_*$ is the set of vertexes of the rectangles $I_{ij}(\mathbf{a}, \mathbf{b})$s. Thus $\mathcal{A}_* \neq \mathcal{A}$ in general. Let

$$N_{nik}(\mathbf{a}, \mathbf{b}) = \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}(\mathbf{X}_j \in I_{ik}(\mathbf{a}, \mathbf{b}), \mathbf{U}_j = \mathbf{a}, \mathbf{V}_j = \mathbf{b}), \; i, k = 1, 2, 3.$$

Then the generalized likelihood (2.1) is equal to

$$\Lambda_n(F) = \prod_{(\mathbf{a}, \mathbf{b}) \in \mathcal{B}} \prod_{i=1}^{3} \prod_{j=1}^{3} [\mu_F(I_{ij}(\mathbf{a}, \mathbf{b}))]^{nN_{nij}(\mathbf{a}, \mathbf{b})},$$

where

$$\mu_F((c, d] \times (e, f]) = F(d, f) + F(c, e) - F(c, f) - F(d, e). \tag{3.2}$$

Moreover, the normalized generalized log-likelihood function is

$$\mathcal{L}_n(F) = \sum_{(\mathbf{a}, \mathbf{b}) \in \mathcal{B}} \sum_{i=1}^{3} \sum_{j=1}^{3} N_{nij}(\mathbf{a}, \mathbf{b}) \ln[\mu_F(I_{ij}(\mathbf{a}, \mathbf{b}))].$$

Here and below we interpret $0 \log 0 = 0$ and $\log 0 = -\infty$. For this likelihood function, we let F range over the set $\mathcal{F}^*$ of all functions F on $[-\infty, +\infty]^2$ such that

$$F(+\infty, +\infty) = 1, \tag{3.3}$$

$$F(-\infty, x) = F(x, -\infty) = 0 \text{ for each } x, \tag{3.4}$$

and

$$\mu_F(I) \geq 0 \text{ for all rectangle sets } I \text{ in } (-\infty, +\infty]^2. \tag{3.5}$$

In view of (3.2), $\Lambda_n(F)$ and $\mathcal{L}_n(F)$ depend on F only through the values of F at the points $\mathbf{x} \in \mathcal{A}_*$. Because the GMLE of F_0 is not unique, we adopt expression (2.2) for the GMLE in our proofs below.

Theorem 1. *Under Assumption (3.1), the GMLE $\hat{F}_n$ satisfies $\hat{F}_n(\mathbf{a}) \to F_0(\mathbf{a})$ almost surely for all $\mathbf{a} \in \mathcal{A}_*$.*

proof. Verify that

$$\text{Ł}(F) := E(\mathcal{L}_n(F)) = \sum_{(\mathbf{a},\mathbf{b})\in\mathcal{B}} g(\mathbf{a},\mathbf{b})h_{\mathbf{a},\mathbf{b}}(F) \tag{3.6}$$

with

$$h_{\mathbf{a},\mathbf{b}}(F) = \sum_{i=1}^{3}\sum_{j=1}^{3} \mu_{F_0}(I_{ij}(\mathbf{a},\mathbf{b}))\ln[\mu_F(I_{ij}(\mathbf{a},\mathbf{b}))].$$

Verify that the expression $h_{\mathbf{a},\mathbf{b}}(F)$ is maximized by a function $F \in \mathcal{F}^*$ if and only if

$$\mu_F(I_{ij}(\mathbf{a},\mathbf{b})) = \mu_{F_0}(I_{ij}(\mathbf{a},\mathbf{b})),\ i,\ j = 1,\ 2,\ 3. \tag{3.7}$$

Equations (3.2) and (3.4) imply that (3.7) is equivalent to $F(\mathbf{x}) = F_0(\mathbf{x})$ for each vertex $\mathbf{x}$ of rectangles $I_{ij}(\mathbf{a},\mathbf{b})$, i, $j = 1, 2, 3$. Thus F_0 maximizes $\text{Ł}(F)$ and any other function in $\mathcal{F}^*$ that maximizes $\text{Ł}(F)$ will coincide with F_0 on $\mathcal{A}_*$.

Note that $\mathcal{L}_n(F_0) = \frac{1}{n}\sum_{j=1}^{n}\psi(\mathbf{X}_j,\mathbf{U}_j,\mathbf{V}_j)$, where ψ is the map defined by

$$\psi(\mathbf{x},\mathbf{a},\mathbf{b}) = \sum_{i=1}^{3}\sum_{j=1}^{3} \mathbf{1}(\mathbf{x}\in I_{ij}(\mathbf{a},\mathbf{b}))\ln(\mu_F(I_{ij}(\mathbf{a},\mathbf{b}))).$$

Thus it follows from the SLLN and (3.2) that $\mathcal{L}_n(F_0) \to \text{Ł}(F_0)$ almost surely. By the definition of the GMLE, $\mathcal{L}_n(\hat{F}_n) \geq \mathcal{L}_n(F_0)$. Consequently,

$$\underline{\lim}_{n\to\infty}\ \mathcal{L}_n(\hat{F}_n) \geq \underline{\lim}_{n\to\infty}\ \mathcal{L}_n(F_0) = \text{Ł}(F_0) \text{ almost surely.}$$

Let Ω' denote the event on which $\underline{\lim}_{n\to\infty}\ \mathcal{L}_n(\hat{F}_n) \geq \text{Ł}(F_0)$. Fix an $\omega \in \Omega'$, let $F^* \in \mathcal{F}^*$ be a limit point of $\hat{F}_{k_n}(\cdot,\omega)$ in the sense that $\hat{F}_{k_n}(\mathbf{a},\omega) \to F^*(\mathbf{a})$ for all $\mathbf{a}\in\mathcal{A}_*$ and for some sequence $\{k_n\}$ of positive integers tending to infinity. We now show that

$$\text{Ł}(F^*) \geq \text{Ł}(F_0).$$

Let $t_{k_n}(\mathbf{a},\mathbf{b})$ denote the value of the random variable $\sum_{i=1}^{3}\sum_{j=1}^{3} N_{k_n ij}(\mathbf{a},\mathbf{b})\ln[\mu_{\hat{F}_{k_n}}(I_{ij})]$ at the point ω. By the definition of Ω',

$$\underline{\lim}_{n\to\infty} \sum_{(\mathbf{a},\mathbf{b})\in\mathcal{B}} t_{k_n}(\mathbf{a},\mathbf{b}) \geq \text{Ł}(F_0).$$

Next, verify that

$$t_{k_n}(\mathbf{a},\mathbf{b}) \to g(\mathbf{a},\mathbf{b})h_{\mathbf{a},\mathbf{b}}(F^*)$$

for each $(\mathbf{a},\mathbf{b})\in\mathcal{B}$. Note also that $t_{k_n}(\mathbf{a},\mathbf{b}) \leq 0$ for all $(\mathbf{a},\mathbf{b})\in\mathcal{B}$. From Fatou's Lemma,

$$\begin{aligned}
\overline{\lim}_{n\to\infty} \sum_{(\mathbf{a},\mathbf{b})\in\mathcal{B}} t_{k_n}(\mathbf{a},\mathbf{b}) &= -\ \underline{\lim}_{n\to\infty} \sum_{(\mathbf{a},\mathbf{b})\in\mathcal{B}} -t_{k_n}(\mathbf{a},\mathbf{b}) \\
&\leq -\sum_{(\mathbf{a},\mathbf{b})\in\mathcal{B}} \underline{\lim}_{n\to\infty}(-t_{k_n}(\mathbf{a},\mathbf{b})) \\
&= \sum_{(\mathbf{a},\mathbf{b})\in\mathcal{B}} g(\mathbf{a},\mathbf{b})h_{\mathbf{a},\mathbf{b}}(F^*) \\
&= \text{Ł}(F^*).
\end{aligned}$$

Combining the above yields $\text{Ł}(F_0) \leq \text{Ł}(F^*)$. As F_0 maximizes Ł, we conclude that $\text{Ł}(F^*) = \text{Ł}(F_0)$ and therefore $F^*(\mathbf{a}) = F_0(\mathbf{a})$ for all $\mathbf{a}\in\mathcal{A}_*$. Since ω is arbitrary and Ω' has probability one, the consistency result is thus established □

If $\mathcal{A}_*$ is a finite set, then it follows from the theorem that the GMLE is uniformly strongly consistent on $\mathcal{A}_*$. For arbitrary $\mathcal{A}_*$, the uniform strong consistency of the GMLE requires additional assumptions.

Theorem 2. *Suppose that (3.1) holds, F_0 is continuous and $\mathcal{A}_*$ is dense in $[0,+\infty)^2$. Then $\sup_{\mathbf{x}\in\mathcal{R}^2}|\hat F_n(\mathbf{x})-F_0(\mathbf{x})|\to 0$ almost surely.*

Proof. Let $F_1, F_2, \ldots$ be functions in $\mathcal{F}^*$ such that $F_n(\mathbf{a})\to F_0(\mathbf{a})$ for all $\mathbf{a}\in\mathcal{A}_*$. Let M be a positive integer. Since F_0 is continuous, there is a grid which partitions the space $(-\infty,+\infty]^2$ into M disjoint rectangles $I=(c,d]\times(e,f]$ with grid points (upper-right vertexes of Is) $\mathbf{x}_1,\ldots,\mathbf{x}_M$ in $(-\infty,+\infty]^2$ and $\mu_{F_0}(I)\le 1/M$ for each grid cell I. The continuity of F_0 and the fact that $\mathcal{A}_*$ is dense in $[0,+\infty)^2$ imply that there are points $\mathbf{a}_1,\ldots,\mathbf{a}_M$ in $\mathcal{A}_*$ such that $|F_0(\mathbf{a}_i)-F_0(\mathbf{x}_i)|\le 1/M^2$. Using this and the facts $F_0, F_n\in\mathcal{F}^*$ and that $F_0(c,e)\le F_0(\mathbf{x})\le F_0(d,f)$ and $F_n(c,e)\le F_n(\mathbf{x})\le F_n(d,f)$ for each $\mathbf{x}\in I$, we derive that

$$|F_n(\mathbf{x})-F_0(\mathbf{x})|\le \max_{1\le i\le M}|F_n(\mathbf{a}_i)-F_0(\mathbf{a}_i)|+\frac{3}{M},\ \ \mathbf{x}\in\mathcal{R}^2.$$

This shows that F_n converges to F_0 uniformly.

By the above, the events $\bigcap_{\mathbf{a}\in\mathcal{A}_*}\{\hat F_n(\mathbf{a})\to F_0(\mathbf{a})\}$ and $\{\sup_{\mathbf{x}\in\mathcal{R}^2}|\hat F_n(\mathbf{x})-F_0(\mathbf{x})|\to 0\}$ are identical and thus have probability 1 by Theorem 1. $\square$

Remark 2 In the case of the bivariate right censorship model, under the assumptions in Theorem 2, it is well known that the GMLE is not a consistent estimate of a continuous F_0 (see Tsai *et al.* (1986)).

4. Asymptotic normality of the GMLE

Under the univariate case 2 interval censorship model, Groeneboom and Wellner (1992) conjecture that if the censoring distribution is continuous, then the GMLE of a continuous F_0 is not asymptotically normally distributed and the convergence rate is not in $\sqrt{n}$. Yu *et al.* (1998) prove that if the censoring vector takes on finitely many values, then under an additional assumption the GMLE is asymptotically normally distributed and the convergence rate is in $\sqrt{n}$. In the multivariate case, the situation is more complicated. In this section we shall obtain the asymptotic normality of the GMLE under the C2M model and the assumptions that

$$\mathcal{A}_* \text{ contains finitely many elements,} \tag{4.1}$$

$$\mu_{F_0}((a_1,b_1]\times(a_2,b_2])>0 \ \text{ if } \mathbf{a},\ \mathbf{b}\in\mathcal{A}_* \text{ and } a_i<b_i,\ i=1,2. \tag{4.2}$$

and

$$\mathcal{A}_*=\mathcal{A} \text{ (see Section 3).} \tag{4.3}$$

Note that under the current assumptions the standard method for finite parametric models can be used.

Remark 3 The GMLE of $\mathbf{s}$ may not be unique (see Remark 1) and Theorem 1 does not ensure the consistency of the GMLE $\hat{\mathbf{s}}$ as $\mathcal{A}$ and $\mathcal{A}_*$ are not the same in general. Note that the consistency of the GMLE $\hat F_n$ on $\mathcal{A}_*$ is mainly due to Eq. (3.7), since $\mathcal{A}_*$ is the set of all vertexes of the rectangles $I_{ij}(\mathbf{a},\mathbf{b})$'s.

By Theorem 1 and (4.3), the GMLE $\hat F_n$ is consistent on the set $\mathcal{A}$. Since $\hat s_j=\mu_{\hat F_n}(A_j)$, where the vertexes of the MI A_j belong to $\mathcal{A}$, $\hat{\mathbf{s}}$ is consistent by (3.2).

Let $s_j^o=\mu_{F_0}(A_j)$. Then (4.2) yields $s_j^o>0$ for all j. Verify that (3.6) yields

$$\begin{aligned}
\text{Ł}(F) &= \sum_{(\mathbf{a},\mathbf{b})\in\mathcal{B}} g(\mathbf{a},\mathbf{b})\sum_{i=1}^{3}\sum_{l=1}^{3}\sum_k s_k^o\mathbf{1}(A_k\subset I_{il}(\mathbf{a},\mathbf{b}))\ln\sum_j s_j\mathbf{1}(A_j\subset I_{il}(\mathbf{a},\mathbf{b}))\\
&= \sum_{(\mathbf{a},\mathbf{b})\in\mathcal{B}}\sum_{i=1}^{3}\sum_{l=1}^{3}\big[g(\mathbf{a},\mathbf{b})\sum_k s_k^o\mathbf{1}(A_k\subset I_{il}(\mathbf{a},\mathbf{b}))\big]\ln\sum_j s_j\mathbf{1}(A_j\subset I_{il}(\mathbf{a},\mathbf{b})).
\end{aligned}\tag{4.4}$$

Let

$$\{I_1,\ldots,I_\beta\}=\{I_{ij}(\mathbf{a},\mathbf{b}):\ i,j=1,2,3,\ (\mathbf{a},\mathbf{b})\in\mathcal{B}\},$$

and

$$p_h=g(\mathbf{a},\mathbf{b})\sum_k s_k^o\mathbf{1}(A_k\subset I_{il}(\mathbf{a},\mathbf{b})).$$

We can rewrite (4.4) as

$$Ł(F) = \sum_{h=1}^{\beta} p_h \ln \sum_{j=1}^{m} s_j \mathbf{1}(A_j \subset I_h) = \sum_{h=1}^{\beta} p_h \ln \sum_{j=1}^{m} s_j \delta_{hj}.$$

From (4.2), $p_h > 0$, $h = 1, ..., \beta$. Set $J = -E(\frac{\partial^2 \mathcal{L}(F_0)}{\partial \mathbf{s} \partial \mathbf{s}^t})$, where $\frac{\partial \mathcal{L}}{\partial \mathbf{s}}$ is an $(m-1) \times 1$ vector and $\frac{\partial^2 \mathcal{L}}{\partial \mathbf{s} \partial \mathbf{s}^t}$ is an $(m-1) \times (m-1)$ matrix. Verify that

$$J = nE(\frac{\partial \mathcal{L}(F_0)}{\partial \mathbf{s}} \frac{\partial \mathcal{L}(F_0)}{\partial \mathbf{s}^t}) = -\frac{\partial^2 Ł}{\partial \mathbf{s} \partial \mathbf{s}^t} = (\sum_{h=1}^{\beta} p_h \frac{(\delta_{hi} - \delta_{hm})(\delta_{hj} - \delta_{hm})}{(\sum_{k=1}^{m} \delta_{hk} s_k^o)^2})_{(m-1)\times(m-1)} = UU^t,$$

where

$$U = \begin{pmatrix} \frac{(\delta_{11} - \delta_{1m})\sqrt{p_1}}{\sum_{k=1}^{m} \delta_{1k} s_k^o} & \cdots & \frac{(\delta_{\beta 1} - \delta_{\beta m})\sqrt{p_\beta}}{\sum_{k=1}^{m} \delta_{\beta k} s_k^o} \\ \vdots & \cdots & \vdots \\ \frac{(\delta_{1(m-1)} - \delta_{1m})\sqrt{p_1}}{\sum_{k=1}^{m} \delta_{1k} s_k^o} & \cdots & \frac{(\delta_{\beta(m-1)} - \delta_{\beta m})\sqrt{p_\beta}}{\sum_{k=1}^{m} \delta_{\beta k} s_k^o} \end{pmatrix}.$$

We now show that J is nonsingular. Let $\mathbf{x}_j$ be the upper-right vertex of A_j, $j = 1, ..., m-1$. By reordering the I_j's, WLOG, we can assume that the upper-right vertex of I_i is equal to $\mathbf{x}_i$, $i = 1, ..., m-1$. Thus $I_i \cap A_j = \emptyset$ for $j > i$, $i = 1, ..., m-1$. Then the matrix U has the upper triangle matrix form

$$U = \begin{pmatrix} \frac{\sqrt{p_1}}{s_1^o} & \cdot & \cdots & \cdot & \cdots & \frac{(\delta_{\beta 1} - \delta_{\beta m})\sqrt{p_\beta}}{\sum_{k=1}^{m} \delta_{\beta k} s_k^o} \\ 0 & \frac{\sqrt{p_2}}{s_2^o + \delta_{21} s_1^o} & \cdots & \cdot & \cdots & \frac{(\delta_{\beta 2} - \delta_{\beta m})\sqrt{p_\beta}}{\sum_{k=1}^{m} \delta_{\beta k} s_k^o} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{\sqrt{p_{m-1}}}{s_{m-1}^o + \sum_{k=1}^{m-2} \delta_{(m-1)k} s_k^o} & \cdots & \frac{(\delta_{\beta(m-1)} - \delta_{\beta m})\sqrt{p_\beta}}{\sum_{k=1}^{m} \delta_{\beta k} s_k^o} \end{pmatrix}.$$

Recall $s_i^o > 0$ and $p_i > 0$ for $i = 1, ..., m-1$. It follows that the matrix U is of full rank and $J = UU^t$ is nonsingular.

It is easy to verify that

$$\frac{\partial^2 \mathcal{L}(\hat{F}_n)}{\partial \mathbf{s} \partial \mathbf{s}^t} \to E(\frac{\partial^2 \mathcal{L}(F_0)}{\partial \mathbf{s} \partial \mathbf{s}^t}) = -J.$$

It thus follows that

$$\frac{\partial \mathcal{L}(\hat{F}_n)}{\partial \mathbf{s}} = \frac{\partial \mathcal{L}(F_0)}{\partial \mathbf{s}} - J\Delta_n + o_p(\|\Delta_n\|),$$

where Δ_n is the $(m-1)$-dimensional column vector with entries $\hat{s}_i - s_i^o = \mu_{\hat{F}_n}(A_i) - \mu_{F_0}(A_i)$, $i = 1, \ldots, m-1$. Let $\Omega_n = \{\inf_{i \le m} \hat{s}_i = 0\}$. Verify that

$$0 = \frac{\partial \mathcal{L}(\hat{F}_n)}{\partial \mathbf{s}} \text{ except on the event } \Omega_n,$$

and by Theorem 1 and Assumptions (4.1) and (4.2),

$$\mathrm{P}(\Omega_n) \to 0 \text{ as } n \to \infty.$$

It follows from the CLT that $\sqrt{n}\frac{\partial \mathcal{L}(F_0)}{\partial \mathbf{s}}$ is asymptotically normal with mean 0 and dispersion matrix J. This shows that $\Delta_n = J^{-1}\frac{\partial \mathcal{L}(F_0)}{\partial \mathbf{s}} + o_p(n^{-1/2})$. Thus we have the following result.

Theorem 3. *Under Assumptions (4.1), (4.2) and (4.3),* $\sqrt{n} \begin{pmatrix} \hat{s}_1 - s_1^o \\ \vdots \\ \hat{s}_{m-1} - s_{m-1}^o \end{pmatrix}$ *is asymptotically normal with mean 0 and dispersion matrix* J^{-1}. *A strongly consistent estimator of* J *is given by* $\hat{J} = -\frac{\partial^2 \mathcal{L}(\hat{F}_n)}{\partial \mathbf{s} \partial \mathbf{s}^t}$.

Furthermore, $\sqrt{n}[\hat{F}_n(\mathbf{x}) - F_0(\mathbf{x})]$ *is asymptotically normally distributed for all* $\mathbf{x} \in \mathcal{A}_*$*. A consistent estimate of the asymptotic variance of* $\hat{F}_n(\mathbf{x})$ *is* $\frac{1}{n}\mathbf{c}^t \hat{J}^{-1}\mathbf{c}$*, where* $\mathbf{c}$ *is a* $(m-1) \times 1$ *vector with the* i*-th entry* $c_i = \mathbf{1}(A_i \subset [0, x_1] \times [0, x_2])$ *unless* $F_0(\mathbf{x}) = 1$*.*

Under the assumptions in Theorem 3, the GMLE is also asymptotically efficient. The proof of this assertion is straightforward and is omitted.

ACKNOWLEDGMENT

We thank the referee and an editor for their invaluable suggestions and opinions.

5. References.

* Campbell, G. (1981). Nonparametric bivariate estimation with randomly censored data. *Biometrika.* 68, 417-422.
* Chang, M.N. and Yang, G. (1987). Strong consistency of a self-consistent estimator of the survival function with doubly censored data. *Ann. Statist.*, 15, 1536-1547.
* Clayton, D. G. (1978). A model of association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika.* 65, 141-151.
* Clayton, D. G. and Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model (with discussion). *J. Roy. Statist. Soc. Ser.* A. 148, 82-117.
* Dabrowska, D. M. (1988). Kaplan-Meier estimate on the plane. *Ann. Statist.* 16, 1475-1489.
* Gill, R. (1992). Multivariate survival analysis. *Theory Prob. Applic.* 37, 18-31.
* Groeneboom, P. and Wellner, J. A. (1992). Information bounds and nonparametric maximum likelihood estimation. *Birkhäuser Verlag, Basel.*
* Gu, M.G. and Zhang, C.H. (1993). Asymptotic properties of self-consistent estimator based on doubly censored data. *Ann. of Statist.* 21, 611-624.
* Hanley, J. A. and Parnes, M. N. (1983). Nonparametric estimation of a multivariate distribution in the presense of censoring. *Biometrics.* 39. 129-139.
* The Italian-American Cataract Study Group (1994). Incidence and progression of cortical, nuclear, and posterior subcapsular cataracts. *Amer. J. Ophthal.* , 118, 623-631.
* Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Stat. Assoc.* 53, 457-481.
* Lin, D.Y. and Ying, Z.L. (1993). A simple nonparametric estimator of the bivariate survival function under univariate censoring. *Biometrika.* 80, 573-581.
* Peto, R. (1973). Experimental survival curves for interval-censored data. *Applied Statist.* 22, 86-91.
* Prentice, R. L. and Cai, J. (1992). Covariance and survival function estimation using censored multivariate failure time data. *Biometrika*, 79, 3, 495-512.
* Samuelsen, S. O. and Kongerud, J. (1994). Interval censoring in longitudinal data of respiratory symptoms in aluminium potroom workers: A comparison of methods. *Statist. in Medicine.* 13. 1771-1780.
* Tsai, W. Y., Leurgans, S. and Crowley,J. (1986). Nonparametric estimation of a bivariate survival function in the presence of censoring. *Ann. Statist.* 14, 1351-1365.
* Tsai, W.Y. and Crowley, J. (1985). A large sample study of the generalized maximum likelihood estimators from incomplete data via self-consistency. *Annals of Statistics.* 13, 1317-1334.
* Turnbull, B. W. (1976). The empirical distribution function with arbitrary grouped, censored and truncated data. *J. Roy. Statist. Soc. Ser.* B. 38. 290-295.
* Van der Laan, M. J. (1996) Efficient estimation in the bivariate censoring model and repairing NPMLE. *Ann. Statist.* 24, 596-627.
* Yu, Q.Q., Schick, A., Li, L.L. and Wong, G.Y.C. (1996). Asymptotic properties of the GMLE in the case 1 interval-censorship model with discrete inspection times. *Canadian Journal of Statistics.* (To appear).
* Yu, Q.Q., Schick, A., Li, L.L. and Wong, G.Y.C. (1998). Asymptotic properties of the GMLE of a survival function with case 2 interval-censored data. *Statist. & Prob. Let.* 37, 223-228.