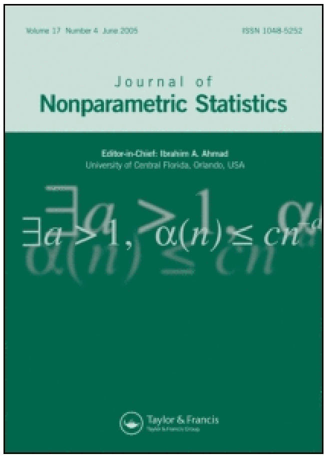

Journal of Nonparametric Statistics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713645758

L-estimators and m-estimators for doubly censored data

Jian-Jian Ren [a]; Mai Zhou [b]
[a] Department of Mathematics& Statistics, University of Nebraska, Lincoln, NE
[b] Department of Statistics, University of Kentucky, Hong Kong Baptist University, Lexington, KY

Online Publication Date: 01 January 1997
To cite this Article: Ren, Jian-Jian and Zhou, Mai (1997) 'L-estimators and m-estimators for doubly censored data', Journal of Nonparametric Statistics, 8:1, 1 - 20
To link to this article: DOI: 10.1080/10485259708832712
URL: http://dx.doi.org/10.1080/10485259708832712

Nonparametric Statistics, Vol. 8, pp. 1–20
Reprints available directly from the publisher
Photocopying permitted by license only

L-ESTIMATORS AND M-ESTIMATORS FOR DOUBLY CENSORED DATA

JIAN-JIAN REN[a,*] and MAI ZHOU[b,†]

[a]*Department of Mathematics & Statistics, University of Nebraska, Lincoln, NE 68588-0323;* [b]*Department of Statistics, University of Kentucky, Lexington, KY 40506-0027 and Hong Kong Baptist University*

(Received 12 October 1996; In final form 10 December 1996)

Motivated by estimation and testing with doubly censored data, we study (robust) L-estimators and M-estimators based on such data. These estimators are given through functionals of the self-consistent estimator $S_X^{(n)}$ for the survival function with doubly censored data. We show that a Hadamard differentiable functional of $S_X^{(n)}$ is asymptotically normal. Thereby, the asymptotic normality of the proposed L-estimators and M-estimators for doubly censored data are derived via their Hadamard differentiability properties. The estimates for the asymptotic variances of the proposed estimators are also given and shown to be strong consistent. The proposed estimators are applied to a doubly censored data set encountered in breast cancer research.

Keywords: Asymptotic normality; asymptotic variance; Fredholm integral equation; Hadamard differentiability; self-consistent estimator; strong consistency

AMS 1991 subject classification: 62E20; 62G05

1 INTRODUCTION, NOTATION AND PRELIMINARIES

Partly powered by the demand in medical research for such tools, analysis of right censored data has been one of the focal points of statistics in the last two decades. Some of the milestones are the Kaplan-Meier estimator, Log-rank test and Cox models.

*Research supported in part by a Faculty Fellowship of University of Nebraska and by NSF grants DMS-9510376 and DMS-9626532.
†Research supported in part by NSF grants DMS-9204380 and SBR-9321103.

Besides the right censored data, there are often other more complicated types of censoring that occur in medical research, such as interval censored data and doubly censored data. Recently the analysis of these data has started to catch the attention of statisticians, as these data are encountered in important clinical trials. For instance, doubly censored data are presented in the study of age-dependent growth rate of primary breast cancer (Peer *et al.*, 1993). The statistical research of these complicated types of data generally lags behind that for right censored data. The analog of the Kaplan-Meier estimator in the doubly censored case has been studied by Chang and Yang (1987), Chang (1990) and Gu and Zhang (1993). But the analog of Log-rank tests has not been developed.

This paper concerns the estimation and testing of the classical location model with doubly censored data, which tries to take advantage of the nonparametric distribution estimators already studied by the papers mentioned above. In this paper we suggest that estimation and testing procedures be based on the analog of the Kaplan-Meier estimators for doubly censored data.

The term *doubly censored data* has been used in the literature to refer to several quite different scenarios. The situation we shall study in this paper is along the lines of Turnbull (1974), Chang and Yang (1987), Chang (1990), Gu and Zhang (1993).

Let $X_i, i = 1, 2, \ldots, n$, be nonnegative independent random variables denoting the lifetime under investigation with distribution function F. In this study, one observes not X_i but a doubly censored sample:

$$W_i = \begin{cases} X_i, & \text{if } Z_i < X_i \leq Y_i \quad \delta_i = 1 \\ Y_i, & \text{if } X_i > Y_i \quad\quad\;\; \delta_i = 2 \\ Z_i, & \text{if } X_i \leq Z_i \quad\quad\;\; \delta_i = 3 \end{cases} \tag{1.1}$$

where independent from X_i, the i.i.d. random vectors (Y_i, Z_i) satisfy $P\{Y_i > Z_i\} = 1$, and Y_i and Z_i are called *right* and *left censoring variables*, respectively. This means that X_i is observable only when it lies in the observation window $[Z_i, Y_i]$, otherwise we only observe Z_i (Y_i) and know X_i is below (above) it. Another way to see this is that we always observe the middle value of the three and know the configuration. The problem considered here is statistical inference concerning the distribution of X_i, F, using the sample (W_i, δ_i), $i = 1, \ldots, n$, in (1.1).

We shall use the random variable without the subscript to denote a typical such random variable. Let

$$S_X(t) = P\{X > t\}, \ \ S_Y(t) = P\{Y > t\}, \ \ S_Z(t) = P\{Z > t\},$$

and let

$$Q_j(t) = P\{W \leq t, \delta = j\}, \quad j = 1, 2, 3$$

$$Q_j^{(n)}(t) = \frac{1}{n}\sum_{i=1}^{n} I\{W_i \leq t, \delta_i = j\}, \quad j = 1, 2, 3.$$

The self-consistent estimators $S_X^{(n)}, S_Y^{(n)}, S_Z^{(n)}$ *of* S_X, S_Y, S_Z (Chang and Yang, 1987) are given (implicitly) by the solutions of the following equations:

$$\begin{aligned} Q_1^{(n)}(t) &= -\int_0^t [S_Y^{(n)} - S_Z^{(n)}]\, dS_X^{(n)} \\ Q_2^{(n)}(t) &= -\int_0^t S_X^{(n)} dS_Y^{(n)} \\ Q_3^{(n)}(t) &= -\int_0^t [1 - S_X^{(n)}]\, dS_Z^{(n)} \end{aligned} \tag{1.2}$$

It has been shown that $S_X^{(n)}, S_Y^{(n)}$ and $S_Z^{(n)}$ are strong uniform consistent, asymptotically normal and efficient (Chang and Yang, 1987; Chang, 1990; Gu and Zhang, 1993). The solutions $S_X^{(n)}$, $S_Y^{(n)}$ and $S_Z^{(n)}$ of the system (1.2) can be calculated numerically (see Mykland and Ren, 1996).

In Section 2, we outline the application of the von Mises method to doubly censored data, and show that with a mild condition, a Hadamard differentiable functional of $S_X^{(n)}$ is asymptotically normal. In Section 3, we give the estimates for the asymptotic variances, which include the L- and M-estimators proposed later on as special cases, and establish the strong consistency of these variance estimators. In Section 4 and Section 5, we propose doubly censoring L-estimators and doubly censoring M-estimators and derive their asymptotic normality via Hadamard differentiability properties, respectively. Some simulation results on the α-trimmed mean estimator and

Huber's M-estimator of location for doubly censored data are also presented. The testing procedures for the location parameter obviously follow easily from our results, which will not be discussed further in this paper. In Section 6, we apply the proposed α-trimmed mean estimator and Huber's M-estimator to a doubly censored data set which was encountered in the study of breast cancer (Peer *et al.*, 1993).

2 HADAMARD DIFFERENTIABLE FUNCTIONALS OF $S_X^{(n)}$

In nonparametric models, a parameter $\theta (= T(F))$ is often regarded as a functional $T(\cdot)$ on a space $\mathscr{F}$ of distribution functions F. Thus, when data are not censored, the same functional of the empirical d.f. F_n for the random sample $X_1, \ldots, X_n$, i.e., $T(F_n)$, is regarded as a natural estimator of θ. In particular, (robust) L-estimators and M-estimators can be expressed in this form (Serfling, 1980; Huber, 1981). Using a form of the Taylor expansion involving the derivative of the functional, von Mises (1974) expressed $T(F_n)$ as

$$T(F_n) = T(F) + T'_F(F_n - F) + \mathrm{Rem}\,(F_n - F; T), \qquad (2.1)$$

where T'_F is the derivative of the functional $T(\cdot)$ at F and $\mathrm{Rem}(F_n - F; T)$ is the remainder term in this first-order expansion. Note that for a continuous and strictly increasing F, a statistical functional $T(\cdot)$ induces a functional τ on the space $D[0,1]$ (of right continuous functions having left-hand limits) as below (Fernholz, 1983):

$$\tau(G) = T(G \circ F), \quad G \in D[0,1]. \qquad (2.2)$$

Thus,

$$T(F) = \tau(U), \quad T(F_n) = \tau(U_n) \qquad (2.3)$$

where U_n is the empirical d.f. of $F(X_i)$, $1 \leq i \leq n$, U is the uniform d.f. on [0,1], and (2.1) can be written equivalently as

$$\tau(U_n) = \tau(U) + \tau'_U(U_n - U) + \mathrm{Rem}\,(U_n - U; \tau), \qquad (2.4)$$

where τ'_U is a linear functional of $D[0,1]$. The asymptotic normality of $T(F_n)$ usually follows from $\tau'_U(U_n - U) = \frac{1}{n}\Sigma_{i=1}^n IC(X_i; F, T)$, where $IC(x; F, T)$ is the *influence curve* of $T(\cdot)$ at F (Fernholz, 1983), and $\sqrt{n}\,\mathrm{Rem}\,(U_n - U; \tau) \xrightarrow{P} 0$, as $n \to \infty$, which is implied by the Hadamard differentiability condition along with some other regularity conditions (Reeds, 1976). More work on Hadamard differentiability can be found in Gill (1989), Ren and Sen (1991, 1995).

When data are doubly censored, a natural extension of the estimator $T(F_n)$ should be given by $T(\widehat{F}_n)$, where $\widehat{F}_n = 1 - S_X^{(n)}$. In Section 4 and Section 5, our generalized L-estimators and M-estimators for doubly censored data are precisely given in the form of $T(\widehat{F}_n)$. We observe that based on (W_i, δ_i), (2.4) becomes

$$\tau(\widehat{U}_n) = \tau(U) + \tau'_U(\widehat{U}_n - U) + \mathrm{Rem}(\widehat{U}_n - U; \tau) \tag{2.5}$$

where $\widehat{U}_n = \widehat{F}_n \circ F^{-1}$ with $T(\widehat{F}_n) = \tau(\widehat{U}_n)$. Note that Chang (1990) has established the following proposition under the condition:

Assumption A

(*A1*) *The random variable X_i and the vector (Y_i, Z_i) are independent for each i and the vectors $(X_i, Y_i, Z_i), i = 1, \ldots, n$, are independently and identically distributed;*

(*A2*) $P\{Z \leq Y\} = 1$;

(*A3*) $S_Y(t) - S_Z(t) > 0$ *on* $(0, \infty)$;

(*A4*) S_X, S_Y, S_Z *are continuous for* $t \geq 0$, *and* $0 < S_X(t) < 1$ *for* $t > 0$;

(*A5*) $S_X(0) = S_Y(0) = 1, S_X(\infty) = S_Y(\infty) = S_Z(\infty) = 0$;

(*A6*) *There exist* δ_0 *and* $\Delta, 0 < \delta_0 < \Delta < \infty$, *such that* $S_Z(t) =$ *constant* < 1 *on* $[0, \delta_0]$ *and* $S_Z(\Delta) = 0$, *i.e.,* $P\{Z = 0\} > 0$, $P\{Z \in (0, \delta_0)\} = 0$ *and* $P\{Z \leq \Delta\} = 1$.

Proposition 2.1 (Chang, 1990) *Under Assumption A, we have that for $t \in [0, M]$*

$$\sqrt{n}\,[\widehat{F}_n(t) - F(t)] = n^{-1/2} \sum_{i=1}^n \xi_i(t) + o_p^{(n)}(1), \tag{2.6}$$

where $M > 0$ is any large enough real number, $o_p^{(n)}(1)$ almost surely converges to 0 uniformly on $[0, M]$ as $n \to \infty$,

$$\xi_i(t) = -\sum_{j=1}^3 \int_0^M IC_j(t, s)\, d(I\{W_i \leq s, \delta_i = j\} - Q_j(s)), \tag{2.7}$$

and $IC_j(t,s)$ is the solution of the integral equation

$$IC_j(t, s) = F_j(t,s) + \int_0^M g(t,u,du)\, IC_j(u,s), \quad j = 1,2,3 \tag{2.8}$$

for

$$F_1(t,s) = -\frac{I\{0 \le s \le t\}}{S_Y(s) - S_Z(s)}, \quad F_2(t,s) = \frac{I\{0 \le s \le t\}}{S_X(s)} \int_s^t \frac{dS_X}{S_Y - S_Z},$$

$$F_3(t,s) = \frac{1}{1 - S_X(s)} \int_0^{s \wedge t} \frac{dS_X}{S_Y - S_Z}, \tag{2.9}$$

$$g(t,s,ds) = F_2(t,s)\, dS_Y(s) - F_3(t,s)\, dS_Z(s).$$

Suppose that we further assume:

ASSUMPTION B *F is strictly increasing on $(0,\infty)$.*
Then, we have that for any $0 < 1 - \delta \le F(M) < 1$ and for $t \in [0, 1-\delta]$,

$$\sqrt{n}\,[\widehat{U}_n(t) - t] = n^{-1/2} \sum_{i=1}^{n} \eta_i(t) + o_p^{(n)}(1), \tag{2.10}$$

where $\eta_i(t) = \xi_i(F^{-1}(t))$.

Under weaker conditions that Assumption A, Gu and Zhang (1993) showed that $\sqrt{n}[\widehat{U}_n - U]$ weakly converges to a centered Gaussian process on [0,1]. Hence, if τ is Hadamard differentiable at U, from Theorem II.8.1. of Andersen, Borgan, Gill and Keiding (1993), we know

$$\sqrt{n}[\tau(\widehat{U}_n) - \tau(U)] \overset{P}{\approx} \tau'_U(\sqrt{n}\,[\widehat{U}_n - U]), \quad \text{as } n \to \infty. \tag{2.11a}$$

Noting that $\eta_i(\cdot)$ are i.i.d. random elements, we know that if (2.10) is true on [0,1] instead of $[0,1-\delta]$, from the linearity of τ'_U we will have

$$\tau'_U(\sqrt{n}\,[\widehat{U}_n - U]) = n^{-1/2} \sum_{i=1}^{n} \tau'_U(\eta_i) + o_p^{(n)}(1) \xrightarrow{D} N(0, \sigma_\tau^2), \tag{2.11}$$

as $n \to \infty$, where $\sigma_\tau^2 = \mathrm{Var}\{\tau'_U(\eta_i)\}$. Nonetheless, for certain classes of estimators, such as L-estimators and M-estimators discussed later on, we often have that for some $0 < \delta < 1$,

$$\tau'_U(G) = \tau'_U(H), \quad \text{if } G(t) \equiv H(t), \quad t \in [0, 1 - \delta]. \tag{2.12}$$

Clearly, we still have (2.11) for a functional τ satisfying (2.12). We state this result in the following theorem.

THEOREM 2.1 *Let $\widehat{\theta}_n = \tau(\widehat{U}_n)$ be an estimator of $\theta = \tau(U)$. If the functional $\tau: D[0,1] \to R$ is Hadamard differentiable at U with $E\{\tau'_U(\eta_i)\} = 0$, $\mathrm{Var}\{\tau'_U(\eta_i)\} = \sigma_\tau^2$, and if (2.12) holds, then under Assumptions A and B, we have*

$$\sqrt{n}[\widehat{\theta}_n - \theta] = \sqrt{n}\,[\tau(\widehat{U}_n) - \tau(U)] \xrightarrow{D} N(0, \sigma_\tau^2), \quad \text{as } n \to \infty$$

where $0 < \sigma_\tau^2 < \infty$.

Remark 1 We should note that Theorem 2.1 is established for any Hadamard differentiable τ defined on $D[0,1]$ satisfying (2.12). Hence, the theorem has broader applications than what is considered in this paper. We should also note that the von Mises method provides a direct approach for obtaining the asymptotic normality of $T(\widehat{F}_n) = \tau(\widehat{U}_n)$ for doubly censored data. Since $S_X^{(n)}$ given by (1.2) can only be obtained numerically, it may be difficult to study the asymptotic properties of $T(\widehat{F}_n)$ through other approaches.

Remark 2 If there is no left censoring, i.e., if $S_Z \equiv 0$, (A6) is satisfied and $S_X^{(n)}$ is the product-limit estimator of S_X in the right censored case (Chang and Yang, 1987). Hence, all of our results in this paper hold in the right censored case. When $S_Z \equiv 0$, our estimators proposed in Section 4 and Section 5 are *right censoring L-estimators* and *right censoring M-estimators*, respectively.

Remark 3 For comments on Assumption A, one may see Chang (1990). Assumption B is to ensure $T(\widehat{F}_n) = \tau(\widehat{U}_n)$ and $T(F) = \tau(U)$ for the functional τ given by (2.2).

Remark 4 Although the weak convergence of $S_X^{(n)}$ has been established by Gu and Zhang (1993) under weaker conditions than those by Chang (1990), the method of calculating the asymptotic variance of $\sqrt{n}\,[S_X^{(n)} - S_X]$ developed by Chang (1990), stated in our Proposition 2.1, enables us to obtain Theorem 2.1 and provides a way to estimate the asymptotic variance σ_τ^2 of the estimator $\widehat{\theta}_n = \tau(\widehat{U}_n)$. Using Chang's results (1990) stated in Proposition 2.1, Ren (1995) has obtained strong uniform consistent estimators for $IC_j(\cdot,\cdot)$ given by (2.8). Hence, a strong consistent estimator of σ_τ^2 can be easily constructed through τ'_U. This is studied in the next section.

3 VARIANCE ESTIMATORS

For the functional τ satisfying (2.12), we often have that for some $M > 0$,

$$\tau'_U(G) = \left(\int_0^M a(x)b(F(x))G(F(x))dx\right) \Big/ \left(\int_0^M c(x)dF(x)\right), \quad G \in D[0,1] \tag{3.1}$$

where a, b and c are bounded functions on $[0, M]$. From (2.7) and (2.10), we have

$$\begin{aligned}\sigma_\tau^2 &= \mathrm{Var}\{\tau'_U(\eta_i)\} \\ &= \left(\sum_{j=1}^{3}\int_0^M\left\{\int_0^M a(x)\,b\,(F(x))\,IC_j(x,y)\,dx\right\}^2 dQ_j(y)\right. \\ &\quad \left.-\left\{\sum_{j=1}^{3}\int_0^M\int_0^M a(x)b(F(x))IC_j(x,y)\,dx\,dQ_j(y)\right\}^2\right)\Big/ \\ &\quad \left\{\int_0^M c(x)dF(x)\right\}^2.\end{aligned} \tag{3.2}$$

Naturally, the variance estimator may be constructed as the following:

(V-1) In (2.8) – (2.9), replace S_X, S_Y, S_Z by their estimators $S_X^{(n)}$, $S_Y^{(n)}, S_Z^{(n)}$ (given by (1.2)), respectively, to compute the estimators $IC_j^{(n)}$ for IC_j;

(V-2) In (3.2), replace F, IC_j, Q_j by $\widehat{F}_n, IC_j^{(n)}, Q_j^{(n)}$, respectively, to compute the ***variance estimator*** $\widehat{\sigma}_n^2$, which we call the *Influence Curve Variance Estimator* (ICVE).

From Chang (1990), we know that IC_j are bounded. Ren (1995, Lemma 4.1 – (v)) has shown

$$IC_j^{(n)} \stackrel{a.s}{\rightarrow} IC_j, \quad \text{as } n \rightarrow \infty \tag{3.3}$$

uniformly on $[0, M]^2$. From an argument similar to that in the proof of Lemma 2.1 by Ren (1995), we know that $\int_0^M a(x)b(F(x))IC_j(x, y)dx$ is continuous in y. Noting that $Q_j^{(n)}$ are empirical processes, we have that if b is piecewise continuous, then

$$\sum_{j=1}^{3} \int_0^M \left\{ \int_0^M a(x)b(\widehat{F}_n(x))IC_j^{(n)}(x, y)dx \right\}^2 dQ_j^{(n)}(y)$$
$$\stackrel{a.s.}{\rightarrow} \sum_{j=1}^{3} \int_0^M \left\{ \int_0^M a(x)b(F(x))IC_j(x, y)dx \right\}^2 dQ_j(y), \quad \text{as } n \rightarrow \infty \tag{3.4}$$

Moreover, if $c \in D[0,M]$, from Lemma 4.1.1 of Fernholz (1983) we have

$$\int_0^M c(x)d\widehat{F}_n(x) \stackrel{a.s.}{\rightarrow} \int_0^M c(x)dF(x), \quad \text{as } n \rightarrow \infty. \tag{3.5}$$

Therefore, we have the strong consistency of the variance estimator $\widehat{\sigma}_n^2$ and we state this result in the following theorem.

THEOREM 3.1 *In additional to Assumption A and B, we assume that in (3.1) the function b is piecewise continuous and $c \in D[0, M]$. Then, the variance estimator $\widehat{\sigma}_n^2$ obtained by (V-1) and (V-2) converges to σ_T^2 with probability 1.*

Remark 5 The equation (2.8) is a general form of the Fredholm integral equation of the second kind. Hence, in step (V-1) one needs to solve the Fredholm integral equation to compute $IC_j^{(n)}$. In numerical analysis, Fredholm integral equations have been well studied. One may see Delves and Walsh (1974) or Delves and Mohamed (1985) for references. Some discussions on the computation of $IC_j^{(n)}$ can be found in Ren (1995).

Remark 6 Our simulation studies in Section 4 and Section 5 show that our method proposed here computes $\widehat{\sigma}_n^2$ very fast and gives satisfactory estimation.

4 L-ESTIMATORS FOR DOUBLY CENSORED DATA

A linear combination of a function of order statistics is called an *L-estimator* (Huber, 1981; Fernholz, 1983). For any function $G : R \to R$, let G^{-1} denote a function given by

$$G^{-1}(t) = \inf\{x; G(x) \geq t\}. \tag{4.1}$$

A von Mises functional representation of L-estimators is usually given by (Fernholz, 1983):

$$T(F_n) = \int_0^1 h\,(F_n^{-1}(t))\, m(t)dt = \sum_{i=1}^{n} \omega_{ni}\, h(X_{(i)}), \tag{4.2}$$

where F_n is the empirical d.f. of a random sample $X_1, \ldots, X_n$ (from a d.f. F), $X_{(i)}$ is the i^{th} order statistic of $X_1, \ldots, X_n$, $h(\cdot)$ is a real valued function and the weights ω_{ni} are generated by the weight function $m(\cdot)$.

A natural extension of L-estimators to doubly censored data is given by

$$T(\widehat{F}_n) = \int_0^1 h(\widehat{F}_n^{-1}(t))\, m\,(t)\, dt. \tag{4.3}$$

Since $S_X^{(n)}$ given by (1.2) can only be obtained numerically, we do not have a clear numerical structure of $T(\widehat{F}_n)$ except for the large sample case. From Theorem 1 of Ren (1994), we know that for sufficiently large n and $t \in [0, M]$,

$$\begin{aligned}\widehat{F}_n(t) = {} & \frac{I\{A_n \leq t\}}{n[S_Y^{(n)}(A_n) - S_Z^{(n)}(A_n)]} + \frac{1}{n} \sum_{A_n < W_i < C_n} \frac{I\{\delta_i = 1, W_i \leq t\}}{[S_Y^{(n)}(W_i) - S_Z^{(n)}(W_i)]} \\ & + \frac{I\{B_n \leq \min(t, M)\}}{n[S_Y^{(n)}(B_n) - S_Z^{(n)}(B_n)]} + \frac{I\{B_n > M, \delta_{C_n} = 1, C_n \leq t\}}{n[S_Y^{(n)}(C_n) - S_Z^{(n)}(C_n)]},\end{aligned} \tag{4.4}$$

with probability 1, where $M > A_n, A_n = \min\{W_i; \delta_i = 1 \text{ or } 3\}$, $B_n = \max\{W_i; \delta_i = 1 \text{ or } 2\}, C_n = \max\{W_i; W_i \leq \min(B_n, M)\}$ and δ_{C_n} is the index of C_n. Hence, if $m(t) = 0$ on the interval $(F^{-1}(M), 1]$, we know that for sufficiently large n, the estimator given by (4.3) can be written as

$$T(\widehat{F}_n) = \sum_{i=1}^{n} \lambda_{ni} h(W_{(i)}), \tag{4.5}$$

with probability 1, where $\lambda_{ni} = \int_{t_{i-1}}^{t_i} m(t)dt$ with $t_i = \widehat{F}_n(W_{(i)}), t_0 = 0$, $t_n = 1$ for $W_{(i)}$'s being the i^{th} order statistic of $W_1, \ldots, W_n$. We can see that our *doubly censoring L-estimator* given by (4.3) is still some kind of a linear combination of a function of order statistics with weights also depending on the data.

By (2.2), the functional τ induced by functional $T(\cdot)$ of (4.3) is defined on $D[0,1]$ and is given by

$$\tau(G) = \int_0^1 h(F^{-1}(G^{-1}(t)))\, m(t)\, dt, \tag{4.6}$$

for $G \in D[0,1]$.

In the following theorem, we show that the doubly censoring L-estimators given by (4.3) are asymptotically normal.

THEOREM 4.1 *Let h be continuous and piecewise differentiable with bounded derivative, and let $m \in L^2[0, 1]$ have support in $[\alpha, 1 - \alpha]$ for some $\alpha > 0$. Suppose that Assumptions A and B hold and that F is absolutely continuous. Then,*

$$\sqrt{n}[T(\widehat{F}_n) - T(F)] \xrightarrow{D} N(0, \sigma^2), \quad \text{as } n \to \infty \tag{4.7}$$

where for some $M > F^{-1}(1 - \alpha/2)$,

$$\sigma^2 = \sum_{j=1}^{3} \int_0^M \left\{ \int_0^M h'(x) IC_j(x, y) m(F(x)) dx \right\}^2 dQ_j(y)$$
$$- \left\{ \sum_{j=1}^{3} \int_0^M \int_0^M h'(x) IC_j(x, y) m(F(x))\, dQ_j(y)\, dx \right\}^2. \tag{4.8}$$

Proof By Proposition 7.2.1 of Fernholz (1983), we know that τ is Hadamard differentiable at U. By the chain rule (Fernholz, 1983, Proposition 3.1.2), we can compute the derivative as below:

$$\tau'_U(G) = -\int_0^1 \frac{h'(F^{-1}(t))}{F'(F^{-1}(t))} G(t)\, m(t)\, dt, \quad \text{for } G \in D[0,1]. \tag{4.9}$$

Clearly, from the conditions on $m(\cdot)$ we know that (2.12) holds for τ'_U. Since Proposition 2.1 implies that for $M \geq F^{-1}(1-\delta) > F^{-1}(1-\alpha)$,

$$\begin{aligned}
\tau'_U(\eta_i) &= -\int_0^1 \frac{h'(F^{-1}(t))}{F'(F^{-1}(t))} \eta_i(t) m(t) dt \\
&= -\int_0^1 \frac{h'(F^{-1}(t))}{F'(F^{-1}(t))} \xi_i(F^{-1}(t)) m(t) dt \\
&= \int_0^1 \frac{h'(F^{-1}(t))}{F'(F^{-1}(t))} \left\{ \sum_{j=1}^{3} \int_0^M IC_j(F^{-1}(t), s) d(I\{W_i \leq s, \delta_i = j\} \right. \\
&\qquad\qquad \left. - Q_j(s)) \vphantom{\sum_{j=1}^3} \right\} m(t) dt \\
&= \int_0^1 \frac{h'(F^{-1}(t))}{F'(F^{-1}(t))} \left\{ \sum_{j=1}^{3} (IC_j(F^{-1}(t), W_i) I\{W_i \leq M, \delta_i = j\} \right. \\
&\qquad \left. - \int_0^M IC_j(F^{-1}(t), s) dQ_j(s)) \right\} m(t) dt \\
&= \int_0^M h'(x) \left\{ \sum_{j=1}^{3} (IC_j(x, W_i) I\{W_i \leq M, \delta_i = j\} \right. \\
&\qquad \left. - \int_0^M IC_j(x, s) dQ_j(s)) \right\} m(F(x)) dx,
\end{aligned}$$

it is easy to see that $E\{\tau'_U(\eta_i)\} = 0$ and that we can compute

$$\sigma_\tau^2 = \mathrm{Var}\{\tau'_U(\eta_i)\} = \sum_{j=1}^{3}\int_0^M\left\{\int_0^M h'(x)IC_j(x,y)m(F(x))dx\right\}^2 dQ_j(y)$$
$$-\left\{\sum_{j=1}^{3}\int_0^M\int_0^M h'(x)IC_j(x,y)m(F(x))dQ_j(y)dx\right\}^2.$$

Therefore, (4.7) follows from Theorem 2.1 and $T(\widehat{F}_n) = \tau(\widehat{U}_n)$. □

From (4.9) and (3.2), one may see that through (V-1) and (V-2) described in Section 3, the variance estimator $\widehat{\sigma}_n^2$ for σ_τ^2 can be computed with $a(x) = -h'(x), b(x) = m(x), c(x) \equiv 1$.

Remark 7 One may note that although Assumptions A and B are the sufficient conditions required for (4.8), the asymptotic normality of the L-estimators given by (4.3) holds under weaker conditions by Gu and Zhang (1993). To see this, one just need to notice that (4.9) does not rely on Assumption A and that the asymptotic normality of $T(\widehat{F}_n)$ in (4.3) follows from (4.9), (2.11a), page 154 – 157 of Iranpour and Chacon (1988), and the weak convergence of $\sqrt{n}\,[\widehat{F}_n - F]$ under weaker conditions by Gu and Zhang (1993). Similarly, because of (5.9), the asymptotic normality of the M-estimators studied in Theorem 5.1 can also be established under weaker conditions by Gu and Zhang (1993).

Example 4.1 The α-trimmed mean for $0 < \alpha < 0.5$ is an L-estimator with $h(x) = x$ and $m(x) = 1/(1-2\alpha)$ for $x \in (\alpha, 1-\alpha)$ and $m(x) = 0$ elsewhere. Our (4.3) gives a generalized α-trimmed mean estimator for doubly censored data. This estimator is asymptotically normal by our Theorem 4.1.

A Small Simulation Study We did a small simulation study on 5%-trimmed mean with sample size 100. In our study, 4 doubly censored samples of size 100 are generated, and the trimmed mean estimators, the *influence curves s.d. estimators* $\sqrt{\mathrm{ICVE}}$ and the bootstrap mean and s.d. are computed. The results are displayed in Table 4.1, where we note that $\sqrt{\mathrm{ICVE}}$ and the bootstrap s.d. are quite close.

Remark 8 Our experiences show that the computation for proposed $\widehat{\sigma}_n^2$ is faster than that using the nonparametric bootstrap. Specifically, for the results displayed in Table 5.1 for M-estimators in Section 5, we

TABLE 4.1 5%-trimmed mean with $n = 100$

$X \sim N(10,4), Z \sim N(8,4)$ $Y = Z + 3 + V,\ V \sim \text{Exp}(1)$	*Sample 1*	*Sample 2*	*Sample 3*	*Sample 4*
Trimmed Mean Estimator	9.741753	9.834141	10.145756	9.829009
$\sqrt{\text{ICVE}}$	0.230878	0.258776	0.272084	0.215944
Bootstrap mean based on 5000 bootstrap samples	9.727450	9.842031	10.14416	9.833629
Bootstrap s.d. based on 5000 bootstrap samples	0.225070	0.247348	0.265843	0.200058

$N(\mu, \sigma^2)$ = normal d.f. with mean μ and s.d. σ; $\text{Exp}(\mu)$ = exponential d.f. with mean μ.

used an HP-730 workstation with 80 MB of RAM and a combination of C and S-plus program, and found that $\sqrt{\text{ICVE}}$ took about 4 seconds while bootstrap took about 30 minutes. Generally, if the sample size n is larger or the sample is more heavily censored or we require high accuracy in EM algorithm for computing $\widehat{F}_n$, the nonparametric bootstrap method can be much more time consuming, but our proposed ICVE method is not affected by all these factors unless n is very large, say $n > 1000$. Another advantage of our approach is that for the same data set, the variance estimators for different estimators can be easily obtained using the influence function estimator already obtained in (V-1) if an analytic expression for $\widehat{\sigma}_n^2$ such as (3.2) is available, while the bootstrap approach needs to start from scratch. Nonetheless, we note that the programming of our proposed method is generally more complicated than that using bootstrap.

5 M-ESTIMATORS FOR DOUBLY CENSORED DATA

The M-estimator is defined implicitly as a solution of the equation (Huber, 1981):

$$\sum_{i=1}^{n} \psi(X_i, \theta) = 0,$$

where $\psi(x, \theta)$ is the *score function* and $X_1, \ldots, X_n$ is a random sample from F. For the M-estimator of location, ψ is given by

$$\psi(x, \theta) = \psi(x - \theta).$$

A von Mises functional representation of M-estimator of location is given by a root $\widehat{\theta}_n = T(F_n)$ of the equation (Fernholz, 1983)

$$\int_0^1 \psi(F_n^{-1}(t) - \theta)dt = \frac{1}{n}\sum_{i=1}^{n} \psi(X_i - \theta) = 0, \tag{5.1}$$

where F_n is the empirical d.f. for $X_1, \ldots, X_n$. The functional $T(\cdot)$ corresponding to (5.1) is defined to be a root $\theta_0 = T(F)$ of the equation

$$\int_0^1 \psi(F^{-1}(t) - \theta)dt = 0, \tag{5.2}$$

or equivalently, of the equation

$$\int_{-\infty}^{\infty} \psi(x - \theta)dF(x) = 0. \tag{5.3}$$

A natural extension of M-estimators to doubly censored data is given by a root $T(\widehat{F}_n)$ of the equation

$$\int_0^1 \psi(\widehat{F}_n^{-1}(t) - \theta)\, dt = 0. \tag{5.4}$$

If for some $M > 0, \psi(x) = 0$ on (M,∞), then from (4.4) we know that for sufficiently large n our *doubly censoring M-estimator* given by (5.4) can be written as the root $T(\widehat{F}_n)$ of the equation

$$\sum_{i=1}^{n} c_{ni}\{[S_Y^{(n)}(W_i) - S_Z^{(n)}(W_i)]\}^{-1}\psi(W_i - \theta) = 0 \tag{5.5}$$

with probability 1, where $c_{ni} = I\{W_i = A_n \text{ or } C_n\} + I\{\delta_i = 1,\ A_n < W_i < C_n\}$. By (2.2), the functional τ induced by functional $T(\cdot)$ of (5.2) is defined on $D[0,1]$ and is given by a root $\tau(G)$ of the equation

$$\int_0^1 \psi(F^{-1}(G^{-1}(t)) - \theta)dt = 0, \tag{5.6}$$

for $G \in D[0, 1]$.

From Theorem 6.2.1. of Fernholz (1983), we know that $\tau(G)$ exists uniquely in a neighborhood of U under certain regularity conditions. Hence, by the strong consistency of $S_X^{(n)}$ (Chang and Yang, 1987), we know that $\tau(\widehat{U}_n) = T(\widehat{F}_n)$ is uniquely defined when n is sufficiently large.

In the following theorem, we show that our doubly censoring M-estimators of location given by (5.4) is asymptotically normal.

THEOREM 5.1 *Let ψ be nondecreasing, continuous, and piecewise differentiable, with bounded derivative ψ' such that $0 < m < \psi'(x)$ for a constant m and x in some neighborhood of 0, and $\psi'(x) = 0$ outside of some bounded interval. Suppose that Assumptions A and B hold and that F has a piecewise continuous density $F'(x) > 0$ on $(0, \infty)$. Then*

$$\sqrt{n}[T(\widehat{F}_n) - T(F)] \xrightarrow{D} N(0, \sigma^2), \quad \textit{as } n \to \infty \tag{5.7}$$

where for $\theta_0 = T(F)$ and some large enough $M > 0$,

$$\sigma^2 = \left(\sum_{j=1}^{3} \int_0^M \left\{ \int_0^M \psi'(x - \theta_0) IC_j(x, y) dx \right\}^2 dQ_j(y) - \left\{ \sum_{j=1}^{3} \int_0^M \int_0^M \psi'(x - \theta_0) IC_j(x, y) dQ_j(y) dx \right\}^2 \right) \Big/ \left\{ \int_0^\infty \psi'(x - \theta_0) dF(x) \right\}^2. \tag{5.8}$$

Proof Let $-\infty < a < b < \infty$ be such that $\psi'(x) \equiv 0$ outside of $[a, b]$, and denote

$$\Psi(G, \theta) = \int_0^1 \psi(F^{-1}(G^{-1}(t)) - \theta) dt,$$

for $G \in D[0, 1], \theta \in R$. By a similar proof of Proposition 7.1.1 given by Fernholz (1983), we know that Ψ is Hadamard differentiable at (U, θ_0) for $\theta_0 = \tau(U)$. By the chain rule (Fernholz, 1983, Proposition 3.1.2), we can compute the partial derivatives of Ψ as below:

$$D_1\Psi_{(U,\theta_0)}(G) = -\int_0^1 \frac{\psi'(F^{-1}(t)-\theta_0)}{F'(F^{-1}(t))}G(t)\,dt, \quad \text{for } G \in D\,[0,1]$$

$$D_2\Psi_{(U,\theta_0)}(\theta) = -\theta\int_0^1 \psi'(F^{-1}(t)-\theta_0)\,dt, \quad \text{for } \theta \in R.$$

From Theorem 5.1.2 and Theorem 6.2.1 of Fernholz (1983), we know that τ is Hadamard differentiable at U with derivative

$$\begin{aligned}\tau'_U(G) &= -(D_2\Psi_{(U,\theta_0)})^{-1} \circ D_1\Psi_{(U,\theta_0)}(G)\\ &= -\left\{\int_0^1 \frac{\psi'(F^{-1}(t)-\theta_0)}{F'(F^{-1}(t))}G(t)dt\right\}\Big/\\ &\quad \left\{\int_0^1 \psi'(F^{-1}(t)-\theta_0)dt\right\} \end{aligned} \tag{5.9}$$

for $G \in D\,[0,1]$. Clearly, from the conditions on the score function ψ we know that (2.12) holds for τ'_U.

Let us denote $D = \int_0^1 \psi'(F^{-1}(t)-\theta_0)dt$. Since by Proposition 2.1, we have that for $M \geq F^{-1}(1-\delta) > b+\theta_0$,

$$\begin{aligned}\tau'_U(\eta_i) &= -\frac{1}{D}\int_0^1 \frac{\psi'(F^{-1}(t)-\theta_0)}{F'(F^{-1}(t))}\eta_i(t)dt\\ &= -\frac{1}{D}\int_0^1 \frac{\psi'(F^{-1}(t)-\theta_0)}{F'(F^{-1}(t))}\xi_i(F^{-1}(t))dt\\ &= \frac{1}{D}\int_0^1 \frac{\psi'(F^{-1}(t)-\theta_0)}{F'(F^{-1}(t))}\\ &\quad \times \left\{\sum_{j=1}^3 \int_0^M IC_j(F^{-1}(t),s)\,d(I\{W_i \leq s, \delta_i = j\} - Q_j(s))\right\}dt\\ &= \frac{1}{D}\int_0^1 \frac{\psi'(F^{-1}(t)-\theta_0)}{F'(F^{-1}(t))}\left\{\sum_{j=1}^3\left\{IC_j\,(F^{-1}(t),W_i)I\{W_i \leq M, \delta_i = j\}\right.\right.\\ &\quad \left.\left.-\int_0^M IC_j(F^{-1}(t),s)dQ_j(s)\right\}\right\}dt\end{aligned}$$

$$= \frac{1}{D}\int_0^M \psi'(x-\theta_0)\left\{\sum_{j=1}^{3}\left\{IC_j(x,W_i)I\{W_i \le M, \delta_i = j\}\right.\right.$$
$$\left.\left. - \int_0^M IC_j(x,s)\,dQ_j(s)\right\}\right\}dx$$

it is easy to see that $E\{\tau'_U(\eta_i)\} = 0$ and that we can compute

$$\sigma_\tau^2 = \mathrm{Var}\{\tau'_U(\eta_i)\}$$
$$= \frac{1}{D^2}\left(\sum_{j=1}^{3}\int_0^M\left\{\int_0^M \psi'(x-\theta_0)IC_j(x,y)dx\right\}^2 dQ_j(y)\right.$$
$$\left. - \left\{\sum_{j=1}^{3}\int_0^M\int_0^M \psi'(x-\theta_0)IC_j(x,y)\,dQ_j(y)dx\right\}^2\right).$$

Therefore, (5.7) follows from Theorem 2.1 and $T(\widehat{F}_n) = \tau(\widehat{U}_n)$. □

From (5.9) and (3.2), one may see that through (V-1) and (V-2) described in Section 3, the variance estimator $\widehat{\sigma}_n^2$ for σ_τ^2 can be computed with $a(x) = -\psi'(x-\widehat{\theta}_n), b(x) \equiv 1, c(x) = \psi'(x-\widehat{\theta}_n)$ for $\widehat{\theta}_n = \tau(\widehat{U}_n)$, and that from Section 3 and Theorem 5.1, $\widehat{\sigma}_n^2$ is at least consistent in probability.

Example 5.1 The M-estimator of location proposed by Huber has the score function:

$$\psi(x) = \begin{cases} -c & \text{if } \ x < -c \\ x & \text{if} -c \le x \le c \\ c & \text{if } \ x > c \end{cases}$$

where c is a positive constant. Our (5.4) gives a generalized Huber's M-estimator of location for doubly censored data. This estimator is asymptotically normal by our Theorem 5.1.

A Small Simulation Study We repeated the simulation study of Table 4.1 for Huber's M-estimator of location with c=1.345 and report the results in Table 5.1. Again, we note that $\sqrt{\text{ICVE}}$ and the bootstrap s.d. are quite close.

TABLE 5.1 Huber's M-estimator with $c = 1.345$ and $n = 100$

$X \sim N(10,4), Z \sim N(8,4)$ $Y = Z + 3 + V, V \sim \text{Exp}(1)$	*Sample 1*	*Sample 2*	*Sample 3*	*Sample 4*
Huber's M-estimator	9.801246	9.956626	10.075554	9.823428
$\sqrt{\text{ICVE}}$	0.223949	0.227287	0.298073	0.218175
Bootstrap mean based on 5000 bootstrap samples	9.787651	9.966692	10.08516	9.820347
Bootstrap s.d. based on 5000 bootstrap samples	0.228318	0.264017	0.2591486	0.191176

$N(\mu, \sigma^2)$ = normal d.f. with mean μ and s.d. σ; $\text{Exp}(\mu)$ = exponential d.f. with mean μ.

6 AN EXAMPLE

To illustrate our proposed method in this paper, we apply the α-trimmed mean estimator and Huber's M-estimator discussed in Section 4 and Section 5, respectively, to a real data set as follows.

Example In a recent study of the age-dependent growth rate of primary breast cancer (Peer *et al.*, 1993), a doubly censored sample is encountered. The age X (in months), at which a tumor volume is developed, is observed among 236 women with age ranged from 41–84 years. From 1981 to 1990, serial screening mammograms with a mean screening interval of 2 years were obtained. Among the tumor volumes detected by the screening mammograms, 45 women had tumor volumes observed at the first screening mammograms – yielding left censored observations, 79 did not have tumor volume observed at the last screening mammograms – yielding right censored observations, and 112 were observed growth during the period of the serial screening mammograms – yielding uncensored observations. For this data set, 10%-trimmed mean estimator described in Example 4.1 is calculated as 773.705, and Huber's M-estimator with $c = 100$ described in Example 5.1 is calculated as 771.753. Evidently, they appear to be very close. In our studies, we find that for a different choice of c, Huber's M-estimator varies, but not much.

Acknowledgement

The authors would like to thank the referees and the associate editor for their comments and suggestions on the manuscript, and would like to thank Dr. P.G.M. Peer at Department of Medical Informatics and Epidemiology, University of Nijmegen, The Netherlands, for providing the original breast cancer data set for our research.

References

Andersen, P. K., Borgan, Ø, Gill, R. D. and Keiding, N. (1993) *Statistical Models Based on Counting Processes*. Springer-Verlag.

Chang, M. N. and Yang, G. L. (1987) Strong consistency of a nonparametric estimator of the survival function with doubly censored data. *Ann. Stat.*, **15**, 1536–1547.

Chang, M. N. (1990) Weak convergence of a self-consistent estimator of the survival function with doubly censored data. *Ann. Stat.*, **18**, 391–404.

Delves, L. M. and Walsh, J. (1974) *Numerical Solution of Integral Equations*, Clarendon Press, Oxford.

Delves, L. M. and Mohamed, J. L. (1985) *Computational Methods for Integral Equations*, Cambridge University Press, Massachusetts.

Fernholz, L. T. (1983) *Von Mises Calculus for Statistical Functional*. Lecture Notes in Statistics 19. Springer, New York. [1.3f, 2.1b, 2.3a].

Gill, R. D. (1989) Non- and semi-parametric maximum likelihood estimators and the von Mises method I. *Scand. J. Statis.*, **16**, 97–128.

Gu, M. G. and Zhang, C. H. (1993) Asymptotic properties of self-consistent estimators based on doubly censored data. *Ann. Stat.*, **21**(2), 611–624.

Huber, P. J. (1981) *Robust Statistics*. John Wiley & Sons, New York.

Iranpour,, R. and Chacon,, P. (1988) *Basic Stochastic Processes – The Mark Kac Lectures*. MacMillan Publishing Company, New York.

Mykland, P. A. and Ren, J. (1996) Algorithms for computing self-consistent and maximum likelihood estimators with doubly censored data. *Ann. Statist.*, **24**, 1740–1764.

Peer, P. G., Van Dijck, J. A., Hendriks, J. H., Holland, R. and Verbeek, A. L. (1993) Age-dependent growth rate of primary breast cancer. *Cancer*, **71**(11), 3547–3551.

Reeds, J. A. (1976) On the definition of Von Mises functional. *Ph. D. dissertation*, Harvard University, Cambridge, MA.

Ren, J. (1995) Generalized Cramr-von Mises tests of goodness of fit for doubly censored data. *Ann. Inst. Statist. Math.*, **47**(3), 525–549.

Ren, J. (1994) On Self-consistent estimators and kernel density estimators with doubly censored data. *J. Statist. Plan. Infer.* (To appear).

Ren, J. and Sen, P. K. (1991) On Hadamard differentiability of extended statistical functional. *J. of Multi. Analys.*, **39**(1), 30–43.

Ren, J. and Sen, P. K. (1995) Hadamard differentiability on $D[0,1]^p$. *J. Multi. Analys.*, **55**(1), 14–28.

Serfling, R. J. (1980) *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.

Turnbull,, B.W. (1974) Nonparametric estimation of a survivorship function with doubly censored data. *JASA*, **69**, 169–173.

Von Mises, R. (1947) On the asymptotic distribution of differentiable statistical functions. *AMS*, **18**, 309–348.