Asymptotic Normality of L-Statistics with Randomly Censored Data

by

Harry Joe¹ and Frank Proschan²

FSU Statistics Report M613
AFOSR Technical Report No. 82-139

April, 1982

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

¹Research supported by an NSERC Canada Postgraduate Scholarship.
²Research performed under the support of the Air Force Office of Scientific Research, AFSC, USAF, under Grant AFOSR-82-K-0007.

Key Words: linear combination of order statistics, randomly censored data, Kaplan-Meier product-limit estimator, asymptotic normality.

Asymptotic Normality of L-Statistics With Randomly Censored Data

by

Harry Joe and Frank Proschan

ABSTRACT

An asymptotic normality theorem is proved under the random censoring model for statistics of the form

$$T_n = \sum_{j=1}^{k} c_j F_n^{-1}(\alpha_j) + \int_{0}^{\infty} B(F_n(x))dx,$$

where $F_n$ is the Kaplan and Meier (1958) product-limit estimator. When there is no censoring, the theorem reduces to an asymptotic normality theorem for L-statistics (linear combinations of order statistics).
1. Introduction

The class of L-statistics, that is, linear combinations of order statistics, has received considerable attention in the statistical literature. Papers dealing with asymptotic normality of L-statistics include those by Chernoff, Gastwirth and Johns (1967), Shorack (1969, 1972), Stigler (1969, 1974) and Boos (1979). In this paper, we prove an asymptotic normality theorem for a class of statistics which are extensions of L-statistics to the censored data situation. Sander (1975b) has studied asymptotic normality of linear combinations of functions of order statistics with censored data, but our asymptotic normality result holds under weaker conditions than those required by Sander.

Let $X_1 \leq X_2 \leq \ldots \leq X_n$ be the order statistics corresponding to a random sample of size $n$ from a distribution function $F$. A convenient subclass of L-statistics broad enough for many applications consists of statistics of the form

$$T_n = \sum_{j=1}^{k} c_j X_{[na_j]} + \sum_{i=1}^{n} a_i X_i,$$

where $k$ is a nonnegative integer (if $k = 0$, the sum is 0), $0 < a_j < 1$, $j = 1, \ldots, k$, $c_1, \ldots, c_k$ are reals, $[\cdot]$ is the greatest integer function, and for some integrable function $J(u)$ defined on $[0, 1]$, $a_{in} = \int_{i/n}^{(i+1)/n} J(u)du$, $i = 1, \ldots, n$. For extension to the censored data situation, the third form of $a_{in}$ above is
most convenient to work with. If the observations are distinct, $F_n$ is the empirical distribution function of $X_1, \ldots, X_n$ and $B(t) = \int_0^t J(u)du$, $0 \leq t \leq 1$, then

$$
\frac{n}{i/n} \sum_{i=1}^n \left( \int J(u)du \right) X_i = \sum_{i=1}^n \left[ B\left(\frac{i-1}{n}\right) - B\left(\frac{i}{n}\right) \right] X_i
$$

$$
= \sum_{i=1}^n B\left(\frac{i-1}{n}\right) [X_i - X_{i-1}] = \int_0^\infty B(F_n(x))dx,
$$

where $X_0 = 0$.

Suppose now that $F$ is a life distribution ($F(0^-) = 0$) and that $F_n$ is the Kaplan and Meier (1958) product-limit estimator of $F$. In this paper, we prove an asymptotic normality theorem for statistics of the form

$$
T_n = \sum_{j=1}^k c_j F_n^{-1}(\alpha_j) + \int_0^\infty E(F_n(x))dx
$$

under the random censoring model, where $F_n^{-1}$ is the left continuous inverse of $F_n$. If $\xi_1 < \ldots < \xi_m$ are the ordered distinct uncensored observations, then

$$
T_n = \sum_{j=1}^k c_j F_n^{-1}(\alpha_j) + \sum_{i=1}^m \int J(u)\xi_i.
$$

That is, $T_n$ is a (random) linear function of the uncensored observations. Several test statistics in life testing are $L$-statistics (e.g. Hollander and Proschan (1975), Joe (1982)). Application of our theorem can be used to obtain asymptotic normality of these test statistics when they are extended to the randomly censored data situation. Sander (1975b) obtains an asymptotic normality theorem for statistics
of the form
\[ T_n^* = \int_0^1 x J(F_n(x)) dF_n(x), \]
where \( M \) is such that \((1 - F(M))(1 - H(M)) > 0\) and \( H \) is the censoring distribution, and applies the theorem to an estimator of the trimmed mean. Our theorem can also be used in the randomly censored data situation to obtain asymptotic normality of the estimator of the trimmed mean and other \( L \)-statistics used as robust estimates.

In the random censoring model, instead of observing a complete sample \( X_1, \ldots, X_n \) from the life distribution \( F \), we are able to observe the pairs \((Z_i, \delta_i)\), \( i = 1, \ldots, n \), where \( Z_i = \min(X_i, Y_i) \), \( \delta_i = 1 \) if \( Z_i = X_i \) (the \( i \)-th observation is uncensored) and \( \delta_i = 0 \) otherwise (the \( i \)-th observation is censored), \( i = 1, \ldots, n \). We assume that the censoring random variables \( Y_1, \ldots, Y_n \) are independently and identically distributed according to the distribution \( H \) and that the \( X \)'s and \( Y \)'s are mutually independent. Therefore, \( Z_1, \ldots, Z_n \) are independently and identically distributed according to the distribution \( K \), where \( 1 - K = \bar{K} = \bar{F} \bar{H} = (1 - F)(1 - H) \).

The product-limit estimator \( \hat{F}_n \) of \( F \), introduced by Kaplan and Meier (1958), is:

\[ 1 - \hat{F}_n(x) = \bar{F}_n(x) = \prod_{\{i: Z_i(x) \leq x\}} \delta_i \left[ 1 - (n-i)/(n-i+1) \right], \quad 0 \leq x < Z_{(n)} , \]
where $Z_{(1)}, \ldots, Z_{(n)}$ are the ordered observations and $\delta_{(i)}$ is the $\delta$ corresponding to $Z_{(i)}$, $i = 1, \ldots, n$. When censored observations are tied with uncensored observations, the convention is to treat uncensored members of the tie as preceding the censored members of the tie. Also, we follow the convention of some authors in treating $Z_{(n)}$ as an uncensored observation whether or not it is uncensored and defining $\hat{F}_n(x) = 0$ for $x \geq Z_{(n)}$. Kaplan and Meier show that $\hat{F}_n$ is the "nonparametric maximum likelihood" estimator of $F$ when no parametric assumptions about $F$ are made. $\hat{F}_n$ reduces to the empirical distribution function when all observations are uncensored. Asymptotic properties (strong consistency, weak convergence) of $\hat{F}_n$ are studied by Efron (1967), Breslow and Crowley (1974), Meier (1975), Peterson (1977) and Gill (1981). Throughout the remainder of this paper, $\hat{F}_n$ will denote the product-limit estimator.

In Section 2, we state the asymptotic normality theorem and make a few remarks. A sufficient condition for $J$ is that either $J$ is continuous or $J$ is continuous except on a finite subset of $(0, 1)$. This type of $J$ function occurs in the work of Joe (1982). The proof of the theorem is given in Section 3.

2. Asymptotic Normality Theorem.

Theorem 1. Suppose $F$ and $H$ are continuous distributions. Let

$$T_n = \frac{1}{k} \sum_{j=1}^{k} c_j \frac{F_n^{-1}(a_j) + \int B(F_n(x))dx}{a_j - a_{j-1}}$$

and

$$T = \frac{1}{k} \sum_{j=1}^{k} c_j \frac{F^{-1}(a_j) + \int B(F(x))dx}{a_j - a_{j-1}}$$

where $k$ is a nonnegative integer, $0 < a_1 < \ldots < a_k < 1$, $c_1, \ldots, c_k$
are reals, $B(t) = \int_0^1 J(u) \, du$, $0 \leq t \leq 1$, and $J$ is a bounded real-valued function on $[0, 1]$. Assume that (for $k \geq 1$) $F'$ exists at $F^{-1}(a_j)$ and $F'(F^{-1}(a_j))$ is positive, $j = 1, \ldots, k$. If

(i) $J$ is continuous, or $J$ is continuous except on a finite subset $A$ of $(0, 1)$ and $F^{-1}$ is continuous on $A$,

(ii) $\int_0^1 [\tilde{H}(x)]^{-1} dF(x) < \infty$, and

(iii) $\int_0^x \int_0^y [F^2 H]^{-1} dF \, dy < \infty$,

then $n^{1/2} T = n^{1/2} (T - T_n) \overset{D}{\rightarrow} N(0, \sigma^2)$ as $n \to \infty$, where

$$
\sigma^2 = \int_0^x \int_0^y J(F(x)) J(F(y)) \tilde{F}(x) \tilde{F}(y) [F^2 \tilde{H}]^{-1} dF \, dy
+ \sum_{i=1}^k \sum_{j=1}^k a_i a_j \int_0^x \int_0^y [F^2 \tilde{H}]^{-1} dF
+ 2 \sum_{j=1}^k a_j \int_0^x J(F(x)) \tilde{F}(x) \int_0^y [F^2 \tilde{H}]^{-1} dF \, dy
$$

(1)

and $a_j = c_j a_j / F'(F^{-1}(a_j))$, $j = 1, \ldots, k$. (Here $\wedge$ is the symbol for minimum.)

Some remarks on Theorem 1 are:

1. Condition (ii) implies that $H^{-1}(1) \geq F^{-1}(1)$; it is satisfied if $[\tilde{H}(x)]^{-1} = O([\tilde{F}(x)]^{-1})$ as $x \to F^{-1}(1)$ for some $\theta \in [0, 1]$.

2. If $[\tilde{H}(x)]^{-1} = O([\tilde{F}(x)]^{-1+\theta})$ as $x \to F^{-1}(1)$ for some $\theta \in [0, 1)$,

then a sufficient condition for condition (iii) is that

$$
\int_0^x [F^{1-\theta}(x)]^{1/2} dF(x) < \infty.
$$
(3) Conditions (i) and (iii) imply that the integral
\[
\int_0^\infty B(F(x))dx \text{ is finite.}
\]

(4) Since \( F \) is continuous, \( \sigma^2 \) can be written as

\[
\sigma^2 = \int_0^1 J(u)J(v)(1-u)(1-v) \int_0^{\alpha^v} (1-t)^{-2}[\hat{H}(F^{-1}(t))]^{-1}dt \int_0^{\alpha^u} dt \int_0^{\alpha^w} dF^{-1}(v)
\]

\[
+ \sum_{i=1}^k \sum_{j=1}^k a_i a_j \int_0^{\alpha_j} (1-t)^{-2}[\hat{H}(F^{-1}(t))]^{-1}dt
\]

\[
+ 2 \sum_{j=1}^k \frac{1}{a_j} \int_0^{\alpha_j} J(u)(1-u) \int_0^{\alpha_j} (1-t)^{-2}[\hat{H}(F^{-1}(t))]^{-1}dF^{-1}(u)
\]

\[
= \int_0^1 J(u)(1-u)dF^{-1}(u) \sum_{j=1}^k a_j \int_0^{\alpha_j} (1-t)^{-2}[\hat{H}(F^{-1}(t))]^{-1}dt
\]

(5) Thus, if \( H_1 \) is stochastically smaller than \( H_2 \) (that is, there tend to be more censored observations with censoring distribution \( H_1 \) than with \( H_2 \)), then \( \sigma^2(F, H_1) \geq \sigma^2(F, H_2) \). \( \sigma^2(F, H) \) is minimized when \( H = 0 \) (that is, no censoring).

(6) When \( H \equiv 0 \), Theorem 1 reduces to an asymptotic normality theorem for the usual L-statistic.

3. Proof of Asymptotic Normality

Under the conditions of Theorem 1, we will show that
\[ n^{-\frac{1}{2}}(T_n - \mu) = n^{-\frac{1}{2}} \left[ \sum_{j=1}^{k} c_j (F_n^{-1}(\alpha_j) - F^{-1}(\alpha_j)) \right] \]

\[ + n^{-\frac{1}{2}} \int_{0}^{\infty} \left[ B(F_n(x)) - B(F(x)) \right] dx \]

\[ = - \sum_{j=1}^{k} \left[ c_j / F''(F_n^{-1}(\alpha_j)) \right] n^{-\frac{1}{2}} \left[ F_n(F_n^{-1}(\alpha_j)) - \alpha_j \right] \]

\[ - \int_{0}^{\infty} J(F(x)) n^{-\frac{1}{2}} (F_n(x) - F(x)) dx + R_n \]

\[ = - \sum_{j=1}^{k} \left[ c_j / F''(F_n^{-1}(\alpha_j)) \right] \zeta_n(F_n^{-1}(\alpha_j)) \]

\[ - \int_{0}^{\infty} J(F(x)) \zeta_n(x) dx + R_n, \] (2)
(1968), p. 30) and Slutsky's theorem,

\[ n \left( T_n - \mu \right) \overset{D}{\to} - \sum_{j=1}^{k} \left[ c_{j}/F^{-1}(F^{-1}(\alpha_j)) (F^{-1}(\alpha_j)) - \int_{0}^{\infty} J(F(x)) \zeta(x) dx \right] \]

as \( n \to \infty \). By the theory of stochastic integration (cf. Parzen (1962), Section 3.3), the limiting random variable is normal with mean 0 and variance given by (1).

We will show that (2) holds through a sequence of lemmas.

If \( J \) is continuous, we make the decomposition:

\[ n^{-1} \int_{0}^{\infty} [B(F_n(x)) - B(F(x))] dx \]

\[ = n^{-1} \int_{0}^{\infty} B^{\ast} \left[ \theta_n(x) F_n(x) + (1 - \theta_n(x)) F(x) \right] (F_n(x) - F(x)) dx \]

\[ = - \int_{0}^{\infty} J(F(x)) \zeta_n(x) dx - \int_{0}^{\infty} \left[ J[\theta_n(x) F_n(x) + (1 - \theta_n(x)) F(x)] - J(F(x)) \right] \zeta_n(x) dx, \]

where \( 0 \leq \theta_n(x) \leq 1 \) for \( x \geq 0 \) (\( \theta_n \) exists by the mean value theorem).

In Lemma 2, we show that

\[ R_n = - \int_{0}^{\infty} \left[ J[\theta_n(x) F_n(x) + (1 - \theta_n(x)) F(x)] - J(F(x)) \right] \zeta_n(x) dx \]

converges in probability to 0 as \( n \to \infty \). If \( J \) is continuous except on a finite set \( A \), then we assume without loss of generality that
A = \{\alpha\}$, where $0 < \alpha < 1$, and make the decomposition:

$$
\begin{align*}
\int_0^n \left[ B(F_n(x)) - B(F(x)) \right] dx &= \int_0^n J(F(x)) \zeta_n(x) dx \\
&\quad - \int_{I_n^C} \left\{ J[\theta_n(x)F_n(x) + (1-\theta_n(x))F(x)] - J(F(x)) \right\} \zeta_n(x) dx \\
&\quad + \int_{I_n} J(F(x)) \zeta_n(x) dx + n^\frac{1}{2} \int_{I_n} \left[ B(F_n(x)) - B(F(x)) \right] dx,
\end{align*}
$$

where $I_n = \left[ F^{-1}(\alpha - \epsilon_n), F^{-1}(\alpha + \epsilon_n) \right]$, $I_n^C = [0, \infty) \setminus I_n$, $
\epsilon_n = \sup \{|F_n(x) - F(x)| : 0 \leq x < \infty\}$ and $0 \leq \theta_n(x) \leq 1$, $x \in I_n^C$.

Let $R_{2n} = -\int_{I_n^C} \left\{ J[\theta_n(x)F_n(x) + (1-\theta_n(x))F(x)] - J(F(x)) \right\} \zeta_n(x) dx$,

$R_{3n} = \int_{I_n} J(F(x)) \zeta_n(x) dx$ and $R_{4n} = n^\frac{1}{2} \int_{I_n} \left[ B(F_n(x)) - B(F(x)) \right] dx$.

In Lemmas 3 to 5, we show that $R_{2n}$, $R_{3n}$ and $R_{4n}$ all converge in probability to 0 as $n \to \infty$.

Finally, in Lemma 6, we show that if $0 < p < 1$, $F'$ exists at $F^{-1}(p)$ and $F'(F^{-1}(p)) > 0$, then

$$
\begin{align*}
n^\frac{1}{2}(F^{-1}(p); F^{-1}(p)) &= n^\frac{1}{2}(T_n(F^{-1}(p)) - (1-p)) / F'(F^{-1}(p)) + R_{5n},
\end{align*}
$$

where $R_{5n} \overset{P}{\to} 0$ as $n \to \infty$ (cf. results of Sander (1975a)). Lemma 6 is an extension of a theorem of Ghosh (1971) to the random censoring situation. The proof is similar to that of Ghosh.

**Lemma 2.** If $J$ is continuous, then $R_{jn} \overset{P}{\to} 0$ as $n \to \infty$. 
Proof. \(|R_{1n}| \leq \sup_{x} |J[\theta_n(x)F_n(x) + (1-\theta_n(x))F(x)] - J(F(x))| \int_{0}^{\infty} |\zeta_n(x)|\,dx.\)

Peterson (1977) shows that \(F_n(x) \rightarrow F(x)\) a.s. for all \(0 \leq x < \tau = \min(F^{-1}(1), H^{-1}(1))\). By condition (ii), \(\tau = F^{-1}(1)\).

Since \(F\) is continuous, it follows by a standard argument (Chung (1974), pp. 132-134) that \(F_n(x) \rightarrow F(x)\) uniformly in \(x\) a.s.. Since \(J\) is continuous, \(\sup_{x} |J[\theta_n(x)F_n(x) + (1-\theta_n(x))F(x)] - J(F(x))| \rightarrow 0\) a.s.. Next note that \(\int_{0}^{\infty} |\zeta(x)|\,dx\) is a proper random variable since

\[
E \left\{ \int_{0}^{\infty} |\zeta(x)|\,dx \right\} = \int_{0}^{\infty} E|\zeta(x)|\,dx \leq \int_{0}^{\infty} [E(\zeta^2(x))]^{\frac{1}{2}}\,dx = \int_{0}^{\infty} [\int_{0}^{\infty} \zeta(x)^2\,dx]^{\frac{1}{2}}\,dF^{\frac{1}{2}} \,dx < \infty
\]

by condition (iii). By the continuous mapping theorem,

\[
\int_{0}^{\infty} |\zeta_n(x)|\,dx \overset{D}{\rightarrow} \int_{0}^{\infty} |\zeta(x)|\,dx
\]

and hence \(\int_{0}^{\infty} |\zeta_n(x)|\,dx\) is bounded in probability. Thus, \(R_{1n} \overset{P}{\rightarrow} 0\) as \(n \rightarrow \infty\).

**Lemma 3.** If \(J\) is bounded and \(J\) is continuous except at \(a\), then \(R_{2n} \overset{P}{\rightarrow} 0\) as \(n \rightarrow \infty\).

Proof. \(|R_{2n}| \leq \sup_{x \in I_n^c} |J[\theta_n(x)F_n(x) + (1-\theta_n(x))F(x)] - J(F(x))| \int_{0}^{\infty} |\zeta_n(x)|\,dx.\)

Since \(F_n(x) \rightarrow F(x)\) uniformly in \(x\) a.s., \(J\) is bounded and continuous on \(I_n^c\) for all \(n\), and by the definition of \(I_n^c\), \((F_n(x) - a)(F(x) - a) > 0\) for \(x \in I_n^c\), then \(\sup_{x \in I_n^c} |J[\theta_n(x)F_n(x) + (1-\theta_n(x))F(x)] - J(F(x))| \rightarrow 0\) a.s.. By condition (iii), \(\int_{0}^{\infty} |\zeta_n(x)|\,dx\) is bounded in probability (see proof of Lemma 2). Thus, \(R_{2n} \overset{P}{\rightarrow} 0\) as \(n \rightarrow \infty\).
Lemma 4. If J is bounded and $F^{-1}$ is continuous at $a$, then $R_{3n} \xrightarrow{P} 0$ as $n \to \infty$.

Proof. $|R_{3n}| \leq \sup |J| \sup |\zeta_n(x)| \cdot [F^{-1}(a+\epsilon_n) - F^{-1}(a-\epsilon_n)]$. Since $\sup |\zeta_n(x)| \overset{D}{\leq} \sup |\zeta(x)|$, $\{\sup |\zeta_n(x)|\}$ is bounded in probability. Also, since $F^{-1}$ is continuous at $a$ and $\epsilon_n \to 0$ a.s., $F^{-1}(a+\epsilon_n) - F^{-1}(a-\epsilon_n) \to 0$ a.s.. Thus, $R_{3n} \xrightarrow{P} 0$ as $n \to \infty$.

Lemma 5. If J is bounded and $F^{-1}$ is continuous at $a$, then $R_{4n} \xrightarrow{P} 0$ as $n \to \infty$.

Proof. $|R_{4n}| \leq \frac{1}{n} \sup |B(F_n(x)) - B(F(x))| \cdot [F^{-1}(a+\epsilon_n) - F^{-1}(a-\epsilon_n)]$.

Since $|B(F_n(x)) - B(F(x))| = \int \limits_{F_n(x)}^{F(x)} |J(u) du| \leq \sup |J| \cdot |F_n(x) - F(x)|$ and

By the boundedness of $\zeta_n(x)$ and $\sup |\zeta_n(x)|$, $\{\sup |\zeta_n(x)|\}$ is bounded in probability, $\{\sup |B(F_n(x)) - B(F(x))|\}$ is bounded in probability. Thus, $R_{4n} \xrightarrow{P} 0$ as $n \to \infty$.

Lemma 6. Let $0 < p < 1$ be such that $H(F^{-1}(p)) < 1$. Suppose that $F^*$ exists at $F^{-1}(p)$ and $F^*(F^{-1}(p)) > 0$. Then

$$n^p(F_n^{-1}(p) - F^{-1}(p)) = n^p(F_n^{-1}(p) - (1-p))/F^*(F^{-1}(p)) + R_{5n},$$

where $R_{5n} \xrightarrow{P} 0$ as $n \to \infty$. (Here $F^{-1}$ and $F^*$ are the left continuous inverses.)

For the proof of Lemma 6, we need the following lemma, which is stated and proved in Ghosh (1971).

Lemma 7. Let $\{A_n\}$ and $\{C_n\}$ be two sequences of random variables satisfying the following conditions:
(a) The sequence \( \{C_n\} \) is bounded in probability, that is, for every \( \delta > 0 \), there exists \( M \) such that \( P(|C_n| > M) < \delta \) for all \( n \).

(b) For all real \( y \) and all \( \epsilon > 0 \),
\[
\lim_{n \to \infty} P(A_n \leq y, C_n \geq y + \epsilon) = 0 \quad \text{and} \quad \lim_{n \to \infty} P(A_n > y, C_n \leq y - \epsilon) = 0.
\]
Then \( A_n \xrightarrow{P} C_n \) as \( n \to \infty \).

**Proof of Lemma 6.** Let \( A_n = n^b(F^{-1}_n(p) - F^{-1}(p)) \) and
\[
C_n = n^b[\bar{F}_n(F^{-1}(p)) - \bar{F}(F^{-1}(p))] / F'(F^{-1}(p)).
\]
Since \( C_n \) converges in distribution, condition (a) of Lemma 7 is satisfied. Let \( y \) be an arbitrary real number. We will show that \( \lim_{n \to \infty} P(A_n \leq y, C_n \geq y + \epsilon) = 0 \) and \( \lim_{n \to \infty} P(A_n > y, C_n \leq y - \epsilon) = 0 \) for all \( \epsilon > 0 \). Hence, by Lemma 7, \( A_n \xrightarrow{P} C_n \) as \( n \to \infty \), which is the conclusion of Lemma 6.

Note that \( A_n \leq y \) if and only if \( p \leq F_n(F^{-1}(p) + n^{-b}y) \) or
\[
D_{y,n} = n^b[\bar{F}_n(F^{-1}(p) + n^{-b}y) - \bar{F}(F^{-1}(p) + n^{-b}y)] / F'(F^{-1}(p)) \leq y_n, \quad \text{where}
\]
\[
y_n = n^b(F^{-1}(p) + n^{-b}y - p) / F'(F^{-1}(p)).
\]
Using Young's form of Taylor's theorem (cf. Serfling (1980), p. 45),
\[
y_n = n^b(F^{-1}(p)n^{-b}y + o(n^{-b})) / F'(F^{-1}(p)) = y + o(1). \quad \text{Thus,} \quad y_n \to y \quad \text{as} \quad n \to \infty.
\]

Let \( x^* \in (F^{-1}(p), \min\{F^{-1}(1), H^{-1}(1)\}) \). Breslow and Crowley (1974) prove the weak convergence of \( \{\xi_n(x), 0 \leq x \leq x^*\} \) to \( \{\xi(x), 0 \leq x \leq x^*\} \) in \( D[0, x^*] \) with the Skorohod metric \( d \), where \( D[0, x^*] \) is the space of real-valued functions on \([0, x^*]\) which are right continuous and have left-hand limits. We can introduce Skorohod (1956) equivalent random elements \( \xi_n^*(n = 1, 2, \ldots), \xi^* \)
such that \( \xi_n^* \) is equal in distribution to \( \xi_n \) in \( D[0, x^*] \), \( n = 1, 2, \ldots \). \( \xi^* \) is equal in distribution to \( \xi \) in \( D[0, x^*] \) and \( d(\xi_n^*, \xi) \to 0 \) a.s. as \( n \to \infty \). Since \( F \) is continuous, \( \{\xi^*(x), \, 0 \leq x \leq x^*\} \) can be assumed to be continuous a.s., so that \( \rho(\xi_n^*, \xi^*) \to 0 \) a.s. as \( n \to \infty \), where \( \rho \) is the supremum metric on \( D[0, x^*] \) (cf. Billingsley (1968), pp. 150-153).

Consider the difference \( D_{y,n} - C_n \)

\[
D_{y,n} - C_n = -\left[F^{-1}(p)\right]^{-1}\left\{\xi_n \left(F^{-1}(p) + n^{-\frac{1}{2}}y\right) - \xi_n \left(F^{-1}(p)\right)\right\}.
\]

Let \( \xi_n = F^{-1}(p) + n^{-\frac{1}{2}}y \) and \( \xi = F^{-1}(p) \), and let \( \varepsilon > 0 \). Then

\[
P\left(F^{-1}(p)\right)_{D_{y,n} - C_n} | > \varepsilon \right) = P\left(\left|\xi_n^*(\xi_n^*) - \xi^*(\xi)\right| > \varepsilon\right)
\]

\[
\leq P\left(\left|\xi_n^*(\xi_n^*) - \xi^*(\xi_n^*)\right| > \varepsilon/3\right) + P\left(\left|\xi_n^*(\xi) - \xi^*(\xi)\right| > \varepsilon/3\right)
\]

\[
+ P\left(\left|\xi_n^*(\xi_n^*) - \xi^*(\xi)\right| > \varepsilon/3\right) + \varepsilon \text{ for all } n
\]

sufficiently large since \( \rho(\xi_n^*, \xi^*) \to 0 \) a.s.. By Chebyshev's inequality,

\[
P\left(\left|\xi_n^*(\xi_n^*) - \xi^*(\xi)\right| > \varepsilon/3\right) \leq 9\varepsilon^{-2}E\left[\xi_n^*(\xi_n^*) - \xi^*(\xi)\right]^2
\]

\[
= 9\varepsilon^{-2}\left\{\int \frac{\xi_n^2}{\xi_n^*} dF - \frac{\xi_n}{\xi_n^*} \int \left[F^2\right]^{-1} dF + \frac{\xi_n}{\xi_n^*} \int \left[F^2\right]^{-1} dF - 2\left(\frac{\xi_n}{\xi_n^*}\right) \int \left[F^2\right]^{-1} dF\right\}
\]

\[
\to 0 \text{ as } n \to \infty.
\]

Thus, \( D_{y,n} - C_n \to 0 \) as \( n \to \infty \), and for all \( \varepsilon > 0 \),

\[
P(\mathbb{A}_n \leq y, C_n \geq y + \varepsilon) = P(D_{y,n} \leq y, C_n \geq y + \varepsilon) \to 0 \text{ as } n \to \infty \text{ and}
\]

\[
P(\mathbb{A}_n > y, C_n \geq y - \varepsilon) = P(D_{y,n} > y, C_n \geq y - \varepsilon) \to 0 \text{ as } n \to \infty.
\]

This completes the proof of Theorem 1.
References


Asymptotic Normality of L-Statistics with Randomly Censored Data

Harry Joe and Frank Proschan

Florida State University
Department of Statistics
Tallahassee, FL 32306

Air Force Office of Scientific Research
The U.S. Air Force
Bolling Air Force Base, DC 20332

April, 1982

Distribution unlimited

linear combination of order statistics, randomly censored data, Kaplan-Meier product-limit estimator, asymptotic normality

An asymptotic normality theorem is proved under the random censoring model for statistics of the form $T_n = \sum_{j=1}^{k} c_j F_{n}^{-1}(\alpha_j) + \int_{0}^{\infty} F_{n}(x)dx$, where $F_{n}$ is the Kaplan and Meier (1958) product-limit estimator. When there is no censoring, the theorem reduces to an asymptotic normality theorem for L-statistics (linear combinations of order statistics).