THE WEAK CONVERGENCE OF QUANTILES
OF THE PRODUCT-LIMIT ESTIMATOR

BY

JOAN M. SANDER

TECHNICAL REPORT NO. 5
FEBRUARY 28, 1975

PREPARED UNDER THE AUSPICES
OF
PUBLIC HEALTH SERVICE GRANT 1 RO1 GM21215-01

DIVISION OF BIOSTATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
THE WEAK CONVERGENCE OF QUANTILES
OF THE PRODUCT–LIMIT ESTIMATOR

by

JOAN M. SANDER

TECHNICAL REPORT NO. 5
FEBRUARY 28, 1975

PREPARED UNDER THE AUSPICES
OF
PUBLIC HEALTH SERVICE GRANT 1 ROI GM21215-01

DIVISION OF BIOSTATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
THE WEAK CONVERGENCE OF QUANTILES
OF THE PRODUCT-LIMIT ESTIMATOR
by
Joan M. Sander†

1. Introduction.

In 1958, Kaplan and Meier introduced the product-limit estimator for a distribution function based on incomplete observations. This estimator is useful in life testing or medical followup, for example, where one may have data from a single lifetime or survival distribution, but the data may be censored on the right, i.e., the observation of the event of interest (called a "death") is prevented by the previous occurrence of some other event (called a "loss"). Despite the resulting incompleteness of the data, one still desires an estimate of the true survival distribution.

In the single sample problem with right censored data, the product-limit estimator (henceforth called the Kaplan-Meier estimator) of the right tail of the underlying lifetime distribution is defined by

\[ 1 - \hat{F}_N(t) = \prod_{r: X(r) \leq t} \left[ \frac{N-r}{N-r+1} \right]^{\delta(r)}, \quad (1.1) \]

where \( X(1) \leq X(2) \leq \cdots \leq X(N) \) are the ordered observations (censored or uncensored), and where \( \delta(r) \) indicates whether \( X(r) \)

† This research was supported by a Fellowship from the San Mateo, California, Branch of the American Association of University Women, and by the Public Health Service Grant 1 R01 GM21215-01.
is uncensored \((\delta_r = 1)\) or censored \((\delta_r = 0)\). Hence, the product in (1.1) is just over the uncensored observations less than or equal to \(t\). The convention in case of a tie is that uncensored observations are ranked ahead of censored observations with which they are tied. If the last observation \(X_{(N)}\) is censored, then \(\hat{F}_N^0(t)\) is considered by some authors to be undefined for \(t \geq X_{(N)}\), or is defined to be one for \(t \geq X_{(N)}\) [cf., Efron (1967); Breslow and Crowley (1974)].

The estimator (1.1) is a step-function which changes its value only by jumps at the uncensored points. Moreover, it is the distribution, unrestricted as to form, which maximizes the likelihood of the observations. If there is no censoring, then the Kaplan-Meier estimator \(\hat{F}_N^0\) is simply the usual empirical distribution function.

The reader is referred to Kaplan and Meier (1958) for a more thorough discussion of the estimator. In particular, it is of interest that \(\hat{F}_N^0(t)\) is an asymptotically consistent estimator of \(F^0(t)\), and is asymptotically normally distributed.

Efron (1967) studied the Kaplan-Meier estimator in the context of the two-sample problem with randomly censored data. That is, the model assumes that the censoring variables are distributed independently of the lifetime variables. This device of random censorship is a convenient theoretical tool for studying the large sample effect of censorship. It was exploited by Breslow and Crowley (1974), who, under the assumptions of continuous distributions and random censorship, showed that \(\hat{F}_N^0(t)\), as a function of \(t\), converges weakly to a Gaussian process.
The purpose of this paper is to show that results analogous to those of Breslow and Crowley hold for quantiles of the Kaplan-Meier product-limit estimator (1.1). Precisely, the weak convergence of \( \hat{F}_N^{0-1}(p) \), properly normalized, to a Gaussian process is established.

Of special interest is the case \( p = .5 \), where the estimator \( \hat{F}_N^{0-1}(.5) \) is simply the median of the Kaplan-Meier distribution function. Currently, work is being done to compare the Kaplan-Meier median estimator with the estimator of the mean \( \mu \) of \( F^0 \) given by

\[
\hat{\mu}_N = \int_0^\infty x \, d\hat{F}_N^0(x) .
\]  

(1.2)

2. **Statement of Results.**

Let \( X_1^0, \ldots, X_N^0 \) denote the true survival times of \( N \) individuals. The random variables \( X_n^0 \) are assumed to be independently, identically distributed according to the distribution \( F^0(t) = \Pr[X_n^0 \leq t] \) with \( F^0(0) = 0 \). Let \( Y_1, \ldots, Y_N \) denote the censoring random variables. Under the model of random censorship, these are assumed to be independent of the \( X_n^0, n = 1, \ldots, N \), and to be independently, identically distributed according to the distribution \( H(t) = \Pr[Y_n \leq t] \).

In the model to be considered here, the \( X_n^0 \) may be censored on the right by the \( Y_n \), so that in reality one observes only the pairs \( (X_n^0, \delta_n) \), \( n = 1, \ldots, N \), where
\[ X_n = \min(X_n^0, Y_n) \quad \text{and} \quad \delta_n = \mathbb{I}_{\{X_n^0 \leq Y_n\}}. \quad (2.1) \]

Hence, the observed \( X_1, \ldots, X_N \) are independently distributed with common distribution \( F \) given by

\[ 1 - F = (1 - F^0)(1 - H). \quad (2.2) \]

The sub-distribution \( \tilde{F} \) of an uncensored observation is given by

\[ \tilde{F}(t) = \Pr[X_n \leq t, \delta_n = 1] = \int_0^t (1 - H) \, dF^0. \quad (2.3) \]

Let \( \hat{F}_N^0(t) \) be the Kaplan-Meier estimator of the distribution \( F^0(t) \) based on the pairs \((X_n, \delta_n), n = 1, \ldots, N\). The \( p^{th} \) quantile of the distribution \( F^0 \) may be estimated by the left-continuous inverse of \( \hat{F}_N^0 \), namely,

\[ \hat{F}_N^{-1}(p) = \inf\{t: \hat{F}_N^0(t) \geq p\}, \quad (2.4) \]

where the infimum of the empty set is taken to be \(+\infty\). The large sample properties of \( \hat{F}_N^{-1}(\cdot) \) studied in this report may be summarized by the following theorems, whose proofs are deferred to later sections.

**Theorem 1.** Suppose \( F^0 \) and \( H \) are continuous distributions. Let \( \alpha \) be any number such that \( 0 \leq \alpha < 1 \) and \( H \circ F^{-1}_0(\alpha) < 1 \). Assume there exists a unique \( t \) such that \( F^0(t) = p \) for each \( 0 \leq p \leq \alpha \). Then, as \( N \to \infty \),

\[ \sqrt{N} \left( \frac{\hat{F}_N^{-1}(p)-p}{\sqrt{\hat{F}_N^{-1}(p)}} \right) \xrightarrow{w} -X(p), \quad 0 \leq p \leq \alpha, \quad (2.5) \]
where \( X(\cdot) \) is a mean zero Gaussian process on \([0, \alpha]\) with covariance

\[
\text{Cov}(X(s), X(t)) = (1-s)(1-t) \int_0^s \frac{1}{(1-w)^2(1-H\circ F_0^{-1}(w))} \, dw, \quad s < t. \tag{2.6}
\]

**Theorem 2.** Under the conditions of Theorem 1, and the additional assumption that \( F_0^{-1} \) has a continuous derivative at each \( p, \ 0 \leq p \leq \alpha \), it holds that

\[
\sqrt{N}\left( F_0^{-1}_N(p) - F_0^{-1}(p) \right) \xrightarrow{w} \left( \frac{\partial F_0^{-1}}{\partial x} \right)_{x=p} (-X(p)), \ 0 \leq p \leq \alpha, \tag{2.7}
\]

as \( N \to \infty \).

**Corollary 1.** Suppose \( f^0 \) is the continuous density corresponding to the distribution \( F^0 \), and that \( f^0(F_0^{-1}(p)) \neq 0 \) for all \( 0 \leq p \leq \alpha \). Then, as \( N \to \infty \),

\[
\sqrt{N}\left( F_0^{-1}_N(p) - F_0^{-1}(p) \right) \xrightarrow{w} \frac{1}{f^0[F_0^{-1}(p)]} (-X(p)), \ 0 \leq p \leq \alpha. \tag{2.8}
\]

If there is no censoring, i.e., \( H(t) \equiv 0 \), then the Kaplan-Meier estimator of \( F^0 \) is simply the usual empirical distribution function, and the covariance function (2.6) reduces to

\[
(1-s)(1-t) \int_0^s \frac{1}{(1-w)^2} \, dw = s(1-t). \tag{2.9}
\]
Setting \( s = t = p \), the usual variance function for the \( p \)-th sample quantile is obtained, namely,

\[
\frac{1}{[f^0(\xi_p)]^2} \cdot p(1-p),
\]

(2.10)

where \( \xi_p = F^0^{-1}(p) \).

If the distribution \( H \) of censoring variables has support on all of \((0, \infty)\), then the condition \( H^0 \circ F^0^{-1} (\alpha) < 1 \) of Theorems 1 and 2 becomes vacuous and the only restriction on \( \alpha \) is that \( 0 \leq \alpha < 1 \).

However, in many applications the censoring variables will be bounded, and then the condition \( H^0 \circ F^0^{-1} (\alpha) < 1 \) is relevant. If \( b \) is taken to be the upper limit of the range of observation, i.e.,

\( b = \inf\{t: H(t) = 1\} \), then the conditions on \( \alpha \) can be rewritten as

\[
\alpha < F^0(b).
\]

(2.11)

Since \( F^0(b) \) is not usually known, care must be taken in deciding which quantiles may be used when applying the above results.

As \( p \) approaches the limit \( F^0(b) \), the variance of the limiting process \( X(p) \) will get arbitrarily large, unless \( b \) is also the limit of support of the underlying \( F^0 \).
3. Two lemmas on the Kaplan-Meier estimator.

These lemmas will be needed in the proof of Theorem 1, so they are presented first.

Lemma 1. Assume $F^0$, $H$ continuous. For the Kaplan-Meier estimator $\hat{P}_N^0(t) = 1 - \hat{P}(t)$, it holds that

$$N \times \text{(the jump size at a death time } t) \overset{P}{\Rightarrow} \frac{1}{1 - H(t)}$$

(3.1)
as $N \to \infty$ for all $t$ such that $H(t) < 1$.

Proof. Label the $N$ ages of death or loss in order of increasing magnitude, and denote them $X(1) \leq X(2) \leq \cdots \leq X(N)$. Thus

$$\hat{P}(t) = \prod_{r} \frac{N-r}{N-r+1},$$

(3.2)

where $r$ runs through those positive integers for which $X(r) \leq t$, and $X(r)$ is the age of death (not loss).

Let $t$ be a death time (and consequently a jump point of $\hat{P}_N$). Then

$$\hat{P}(t) = \prod_{r} \frac{N-r}{N-r+1} = \left( \prod_{r<n} \frac{N-r}{N-r+1} \right) \left( \frac{N-n}{N-n+1} \right)$$

(3.3)

$$= \prod_{r<n} \left( \frac{N-r}{N-r+1} \right) \left( 1 - \frac{1}{N-n+1} \right)$$

$$= \hat{P}(t - 0) \cdot \left( 1 - \frac{1}{N-n+1} \right),$$
where \( t \) is taken to be equal to \( X_{(n)} \), a death time among the \( N \) observations. Consequently, \( N \times \text{(jump size at } t) \) is equal to

\[
\hat{P}(t-0) \cdot \frac{N}{N-n+1}.
\]

But since \( \hat{P} \) is a consistent estimator of \( P = 1 - F^0 \), it holds that \( \hat{P}(t-0) \overset{P}{\to} P(t-0) = 1 - F^0(t-0) = 1 - F^0(t) \) as \( N \to \infty \) by the assumed continuity of \( F^0 \). In addition, \( (N-n+1)/N \) equals the binomial fraction of the \( N \) observations which are neither censored nor dead by time \( t \). Hence,

\[
\frac{N-n+1}{N} \xrightarrow{a.s.} [1-F^0(t)] [1-H(t)] \quad \text{as } N \to \infty,
\]

and so,

\[
\hat{P}(t-0) \cdot \frac{N}{N-n+1} \overset{P}{\to} \frac{[1-F^0(t)]}{[1-F^0(t)] [1-H(t)]} \xrightarrow{P} \frac{1}{1-H(t)}, \quad \text{q.e.d.}
\]

**Lemma 2.** Let the ordered observations be \( X(1) < X(2) < \cdots < X(N) \), and let \( t_1 < t_2 < \cdots < t_k \) be all the ordered times of death (not loss) among the \( X_1, \ldots, X_N \). Then, for any \( j, \quad 0 \leq j \leq k-1 \), it holds that

\[
\frac{\hat{F}_N^0(t_{j+1})}{\hat{F}_N^0(t_{j})} > \frac{\hat{F}_N^0(t_{i+1})}{\hat{F}_N^0(t_{i})} \quad \text{for all } i \leq j,
\]

for all \( i \leq j \).
with strict inequality holding if at least one censored observation falls between $t_{j+1}$ and $t_j$. (Set $t_s = 0$ for $s \leq 0$, so that $\hat{F}_N(t_s) = 0$ for $s \leq 0$.)

**Proof.** By induction. Suppose $j = 0$. Then (3.7) reduces to

$$\hat{F}_N(t_1) - \hat{F}_N(t_0) > \hat{F}_N(t_1) - \hat{F}_N(t_{1-1}) \quad \text{for} \quad i \leq 0,$$

(3.8)

i.e., $\hat{F}_N(t_1) > 0$ which is clearly true by definition of $\hat{F}_N$.

Suppose (3.7) is true for $j = 0, 1, \ldots, k - 1$ ($k \geq 1$). It must be shown that (3.7) holds for $j = k$ as well. By the inductive hypothesis, it is enough to show that

$$\hat{F}_N(t_{k+1}) - \hat{F}_N(t_k) \geq \hat{F}_N(t_k) - \hat{F}_N(t_{k-1}).$$

(3.9)

Since $\hat{P}(t) = 1 - \hat{F}_N(t)$, (3.9) is equivalent to

$$\hat{P}(t_{k-1}) - \hat{P}(t_k) \leq \hat{P}(t_k) - \hat{P}(t_{k+1}).$$

(3.10)

From the definition of $t_1, \ldots, t_{k_N}$, it holds that

$$t_1 = X(j_1), \quad t_2 = X(j_2), \quad \ldots, \quad t_{k_N} = X(j_{k_N})$$

(3.11)

for some $1 \leq j_1 < j_2 < \cdots < j_{k_N} \leq N$. Therefore,
\[ \hat{P}(t_{k-1}) - \hat{P}(t_k) = \hat{P}(x_{(j_{k-1})}) - \hat{P}(x_{(j_k)}) \]

\[ = \prod_{r=1}^{j_{k-1}} \frac{N-r}{N-r+1} - \prod_{r=1}^{j_k} \frac{N-r}{N-r+1} = \prod_{r=1}^{k-1} \left( \frac{N-j_r}{N-j_r+1} \right) \left( 1 - \frac{N-j_k}{N-j_k+1} \right) \]

\[ = \prod_{r=1}^{k-1} \left( \frac{N-j_r}{N-j_r+1} \right) \left( \frac{1}{N-j_k+1} \right), \quad (3.12) \]

while

\[ \hat{P}(t_k) - \hat{P}(t_{k+1}) = \prod_{r=1}^{k} \frac{N-j_r}{N-j_r+1} - \prod_{r=1}^{k+1} \frac{N-j_r}{N-j_r+1} \]

\[ = \left( \prod_{r=1}^{k-1} \frac{N-j_r}{N-j_r+1} \right) \left( \frac{N-j_k}{N-j_k+1} \right) \left( 1 - \frac{N-j_{k+1}}{N-j_{k+1}+1} \right) \]

\[ = \left( \prod_{r=1}^{k-1} \frac{N-j_r}{N-j_r+1} \right) \left( \frac{N-j_k}{N-j_k+1} \right) \left( \frac{1}{N-j_{k+1}+1} \right). \quad (3.13) \]

From (3.12) and (3.13), it follows that the inequality in (3.10) holds if and only if

\[ \frac{1}{N-j_k+1} \leq \frac{N-j_k}{N-j_k+1} \cdot \frac{1}{N-j_{k+1}+1}, \quad (3.14) \]

and (3.14) holds if and only if

\[ 1 \leq \frac{N-j_k}{N-j_{k+1}+1}. \quad (3.15) \]

10
Now \( j_k < j_{k+1} \). Therefore, \( N - j_k > N - j_{k+1} \), and so
\[ N - j_k \geq N - j_{k+1} + 1. \]
This verifies the inequality (3.15), and by induction (3.7) holds for all \( j \).

For the case \( j = k \) in (3.7), expressions (3.12) - (3.15) show that strict inequality holds in (3.7) if and only if
\[ j_{k+1} > j_k + 1. \]
This occurs when at least one censored observation falls between \( t_k \) and \( t_{k+1} \). q.e.d.

4. Proof of Theorem 1.

It will be convenient to define a new set of random variables by setting
\[
U^0_n = F^0(Y^0_n), \quad V^0_n = F^0(Y^0_n), \quad \text{and} \\
U^0_n = F^0(X^0_n) = \min(U^0_n, V^0_n), \quad n = 1, \ldots, N. \tag{4.1}
\]

Since \( \Pr[U^0_n \leq t] = \Pr[X^0_n \leq F^{-1}_0(t)] = F^0(F^{-1}_0(t)) = t, \quad 0 \leq t \leq 1, \)
and \( \Pr[V^0_n \leq t] = \Pr[Y^0_n \leq F^{-1}_0(t)] = H(F^{-1}_0(t)), \quad 0 \leq t \leq 1, \)
it holds that \( U^0_1, \ldots, U^0_N \) are independently and identically distributed according to the uniform distribution \( G^0 = F^0 F^{-1}_0 = I \) on \([0,1] \); \( V_1, \ldots, V_N \) are independently and identically distributed on \([0,1] \) according to the distribution \( H F^{-1}_0 \), and are independent of the \( U^0_n, n = 1, \ldots, N \); and \( U_1, \ldots, U_N \) are independent and identically distributed on \([0,1] \) according to the distribution \( G \), where
\[
1 - G = (1-G^0) (1-H F^{-1}_0) = (1-I) (1-H F^{-1}_0).
\]
Let \( \hat{H}_N^0 \) be the Kaplan-Meier estimator of the distribution \( G^0 \) based on \( U_1, \ldots, U_N \). Applying Theorem 5 of Breslow and Crowley (1974), we know that the random function \( X_N(t) = \sqrt{N}(\hat{H}_N^0(t) - t) \), for \( 0 < t < \alpha \), converges weakly to a mean zero Gaussian process \( X(t) \) with covariance function

\[
\text{Cov}(X(s), X(t)) = (s-t) \int_0^s \frac{1}{(1-G)(1-G^0)} \, dG^0 \quad (4.2)
\]

\[
= (s-t) \int_0^s \frac{1}{(1-w)(1-HF^0) - 1} \, dw, \quad s \leq t,
\]

provided that \( \alpha < 1 \) and \( G(\alpha) < 1 \). Notice that

\[
0 < 1 - G(\alpha) = (1-\alpha) (1-HF^0(\alpha)) \quad (4.3)
\]

if and only if

\[
\alpha < 1 \quad \text{and} \quad HF^0(\alpha) < 1. \quad (4.4)
\]

Breslow and Crowley prove convergence in the supremum metric on \([0, \alpha] \), but this is equivalent to convergence in the Skorohod metric \( d_\alpha \) on \( D[0, \alpha] \), the space of functions on \([0, \alpha] \) having at most jump discontinuities, since the limiting random function is (uniformly) continuous on \([0, \alpha] \). [cf., Billingsley (1968), p. 112 and pp. 150-151.]

By the assumed continuity of \( H \) and \( F^0 \), and condition (4.4) on \( \alpha \), there exists \( T_\alpha > \alpha \) such that \( T_\alpha \) satisfies \( T_\alpha < 1 \) and \( HF^0(T_\alpha) < 1 \). Define the set
\[ \Lambda^\alpha_N = \{ \omega: \hat{U}^{-1}_N(\alpha) \text{ defined for all } p \leq \alpha \} \]

and \[ \hat{U}^{-1}_N(\alpha) \leq T_\alpha, \] (4.5)

where to say \[ \hat{U}^{-1}_N(\alpha) \text{ defined} \] means there exists \( t, 0 \leq t < \infty, \)
such that \[ \hat{U}^{-1}_N(t) \geq p. \]

**Lemma 3.** The set \( \Lambda^\alpha_N \) satisfies

\[ P\{\Lambda^\alpha_N\} \rightarrow 1 \text{ as } N \rightarrow \infty. \] (4.6)

**Proof.** The monotonicity of \( \hat{U}^{-1}_N \) permits \( \Lambda^\alpha_N \) to be rewritten as

\[ \Lambda^\alpha_N = \{ \omega: \hat{U}^{-1}_N(\alpha) \text{ defined, and } \hat{U}^{-1}_N(\alpha) \leq T_\alpha \} \]

\[ = \{ \omega: \hat{U}^{-1}_N(\alpha) \leq T_\alpha \} \]

\[ = \{ \omega: \alpha \leq \hat{U}_N(\alpha) \}. \] (4.7)

By the consistency of the Kaplan-Meier estimator [cf., Kaplan and Meier (1958); Efron (1967); Breslow and Crowley (1974)],

\[ \hat{U}^0_N(T_\alpha) \rightarrow U^0(T_\alpha) = T_\alpha, \] (4.8)

i.e.,

\[ P\{|\hat{U}^0_N(T_\alpha) - T_\alpha| \leq \varepsilon\} \rightarrow 1 \text{ for every positive } \varepsilon. \] (4.9)

But,

\[ P\{|\hat{U}^0_N(T_\alpha) - T_\alpha| \leq \varepsilon\} \leq P\{T_\alpha - \varepsilon \leq \hat{U}^0_N(T_\alpha)\}, \] (4.10)
and since $\alpha < T_\alpha$, $\varepsilon > 0$ arbitrary, it follows from (4.9) and (4.10) that

$$P\{\alpha < \hat{U}^0_N(T_\alpha)\} \rightarrow 1 \text{ as } N \rightarrow \infty.$$ \hspace{1cm} (4.11) \hspace{1cm} \text{q.e.d.}

Now, on the set $\Lambda^\alpha_N$ for all $p \leq \alpha$, we have the representation [cf., Shorack (1973)]

$$Y_N(p) = \sqrt{N}(\hat{U}^0_N(p) - p)$$

$$= -\sqrt{N}(\hat{U}^0_N(U^0_N(p)) - \hat{U}^0_N(p)) + \sqrt{N}(\hat{U}^0_N(U^0_N(p)) - p)$$

$$= -X_N(U^0_N(p)) + \sqrt{N}(\hat{U}^0_N(U^0_N(p)) - p).$$ \hspace{1cm} (4.12)

For all functions $x, y$ on $[0,1]$, any $\gamma < 1$, define the metric

$$\rho_\gamma(x, y) = \sup_{0 < t < \gamma} |x(t) - y(t)|.$$ \hspace{1cm} (4.13)

**Lemma 4.** For every positive $\varepsilon$,

$$P\{\omega \in \Lambda^\alpha_N \text{ such that } \sqrt{N} \rho_\gamma(U^0_N(U^0_N(\cdot)), I(\cdot)) \geq \varepsilon\} \rightarrow 0$$

as $N \rightarrow \infty$. \hspace{1cm} (4.14)

**Proof.** On the set $\Lambda^\alpha_N$,

$$\sqrt{N} \rho_\gamma(U^0_N(U^0_N(\cdot)), I(\cdot)) = \sqrt{N} \sup_{0 < p \leq \alpha} |U^0_N(U^0_N(p)) - p|$$

$$= \sqrt{N} \sup_{0 < p \leq \alpha} \left(U^0_N(U^0_N(p)) - p\right).$$ \hspace{1cm} (4.15)
where the absolute value signs have been dropped since
\[ \hat{U}_N^0(p) = \inf \{ t : \hat{U}_N^0(t) \geq p \}. \]
Now, (4.15) is
\[
\leq \sqrt{N} \max_{j=1, \ldots, m_N} (\hat{U}_N^0(t_{j+1}) - \hat{U}_N^0(t_j)),
\] (4.16)
where \( t_1 < t_2 < \cdots < t_{m_N} \leq T \) are all the times of death observed in \([0, T]\). Moreover, by Lemma 2, the largest jump of the Kaplan-Meier estimator is the last one. Hence (4.16) is equal to
\[
\frac{1}{\sqrt{N}} N (\hat{U}_N^0(t_{m_N}^N) - \hat{U}_N^0(t_{m_N-1}^N)),
\] (4.17)
which by Lemma 1 is
\[
\leq \frac{1}{\sqrt{N}} \left( \frac{1}{1 - \text{HOF}^0(T_\alpha)} + o_p(1) \right).
\] (4.18)

This last expression converges in probability to zero on \( \Lambda_N^\alpha \) as \( N \to \infty \) since \( 1 - \text{HOF}^0(T_\alpha) > 0 \) by assumption. Hence (4.14) holds and the lemma is proven.

**Lemma 5.** For every positive \( \varepsilon \),
\[
P \{ \rho \hat{\Gamma}_N^{0-1} \geq \varepsilon \} \to 0 \quad \text{as} \quad N \to \infty.
\] (4.19)

**Proof.** On the set \( \Lambda_N^\alpha \).
\[
\rho_\alpha(\hat{U}_N^{0-1}, I) = \sup_{0 \leq p < \alpha} |\hat{U}_N^{0-1}(p) - p|
\]

\[
\leq \sup_{0 < t < T_\alpha} |\hat{U}_N^0(t) - t| \quad \text{by symmetry (see Figure 1 below)}
\]

\[
= N^{-1/2} \sup_{0 < t < T_\alpha} |X_N(t)|
\]

\[
p \to 0 \quad \text{(4.20)}
\]

since

\[
\sup_{0 < t < T_\alpha} |X_N(t)| \xrightarrow{w} \sup_{0 < t < T_\alpha} |X(t)|, \quad \text{and} \quad N^{-1/2} \to 0. \quad \text{(4.21)}
\]

Thus,

\[
P\{\rho_\alpha(\hat{U}_N^{0-1}, I) > \varepsilon\}
\]

\[
\leq P(\omega \in \Lambda_N^\alpha: \rho_\alpha(\hat{U}_N^{0-1}, I) > \varepsilon) + P(\omega \notin \Lambda_N^\alpha) \quad \text{(4.22)}
\]

\[
p \to 0 \text{ as } N \to \infty. \quad \text{q.e.d.}
\]

**Figure 1.** Example of Kaplan-Meier estimator \(\hat{U}_N^0(t)\) for \(N = 4, m_N = 3\).
Since \( X_N(\cdot) \Rightarrow X(\cdot) \) relative to \((D[0,T_\alpha], d_T^\alpha)\), we may introduce Skorohod equivalent random elements \( X_N^*(\cdot), N = 1, 2, \ldots, \) and \( X^*(\cdot) \) on \([0,T_\alpha]\) satisfying

\[
\mathcal{L}(X_N) = \mathcal{L}(X_N^*), \quad N = 1, 2, \ldots,
\]

\[
\mathcal{L}(X) = \mathcal{L}(X^*),
\]

(4.23)

and

\[
d_T^{\alpha}(X_N^*, X^*) \to 0 \text{ a.s. as } N \to \infty.
\]

(4.24)

[cf., Pyke (1969); Billingsley (1971).]

Define

\[
\hat{U}_N^0(p) = \frac{X_N^*(p)}{\sqrt{N}} + p,
\]

\[
\hat{U}_N^{0*^{-1}}(p) = \inf\{t: \hat{U}_N^0(t) \geq p\},
\]

\[
Y_N^*(p) = \sqrt{N}(\hat{U}_N^{0*^{-1}}(p) - p),
\]

\[
\Lambda_N^\alpha = \{\omega: \hat{U}_N^{0*^{-1}}(p) \leq T_\alpha \text{ for all } p \leq \alpha\}.
\]

(4.25)

By (4.23) and Lemma 3, the set \( \Lambda_N^\alpha \) satisfies

\[
P(\Lambda_N^\alpha^*) = P(\Lambda_N^0) \to 1 \text{ as } N \to \infty.
\]

(4.26)
By (4.23) and Lemma 4, for every positive \( \varepsilon \),

\[
P\{ \omega \in \Lambda^\alpha_N \cap \mathbb{N} \cap \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), I(\cdot) \right) \geq \varepsilon \} \to 0
\]
as \( N \to \infty \). \hspace{1cm} (4.27)

And by (4.23) and Lemma 5, for every positive \( \varepsilon \),

\[
P\{ \omega : \rho_{\alpha} \left( \hat{U}^0_N \left( \cdot \right), I \right) \geq \varepsilon \} \to 0 \quad \text{as} \quad N \to \infty . \hspace{1cm} (4.28)
\]

Now, \( \mathcal{L}(Y^*_N) = \mathcal{L}(X^*_N) \), \( N = 1, 2, \ldots \), so to prove the weak convergence of \( Y^*_N(\cdot) \) to \(-X(\cdot)\) on \([0, \alpha]\) it is sufficient to prove that \( Y^*_N(\cdot) \overset{w}{\to} -X(\cdot) \) on \([0, \alpha]\). Furthermore, because \(-X(\cdot)\) is (uniformly) continuous a.s. on \([0, T^*_\alpha]\), it is enough [cf., Billingsley (1968), Theorem 4.1] to show that

\[
P\{ \rho_{\alpha} (Y^*_N, -X^*) \geq \varepsilon \} \to 0
\]
as \( N \to \infty \) for arbitrary positive \( \varepsilon \).

By analogy to expression (4.12), on the set \( \Lambda^\alpha_N \), for all \( p \leq \alpha \), we have the representation

\[
Y^*_N(p) = -X^*_N \left( \hat{U}^0_N \left( p \right) \right) + \sqrt{N} \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), X(\cdot) \right) + \sqrt{N} \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), I(\cdot) \right)
\]

Subtracting \(-X^*(p)\) from both sides of (4.30), it follows that

\[
\rho_{\alpha}(Y^*_N, -X^*) = \sup_{0 \leq p \leq \alpha} \left| Y^*_N(p) - (-X^*(p)) \right|
\]

\[
\leq \rho_{\alpha} \left( X^*_N \left( \cdot \right), X^*(\cdot) \right) + \sqrt{N} \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), I(\cdot) \right)
\]

\[
+ \sqrt{N} \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), I(\cdot) \right)
\]

\[
\leq \rho_{\alpha} \left( X^*_N \left( \cdot \right), X^*(\cdot) \right) + \sqrt{N} \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), I(\cdot) \right)
\]

\[
\leq \rho_{\alpha} \left( X^*_N \left( \cdot \right), X^*(\cdot) \right) + \sqrt{N} \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), I(\cdot) \right)
\]

\[
+ \sqrt{N} \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), I(\cdot) \right)
\]

\[
\leq \rho_{\alpha} \left( X^*_N \left( \cdot \right), X^*(\cdot) \right) + \sqrt{N} \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), I(\cdot) \right)
\]

\[
+ \sqrt{N} \rho_{\alpha} \left( \hat{U}^0_N \left( \hat{U}^0_N \left( \cdot \right) \right), I(\cdot) \right)
\]
\[
\leq \rho_\alpha \left( X_N^*(U_N^0 \ast^{-1}, (\cdot)), X^*(U_N^0 \ast^{-1}, (\cdot)) \right) + \rho_\alpha \left( X^*(U_N^0 \ast^{-1}, (\cdot)), X^*(\cdot) \right)
\]
\[
+ \sqrt{N} \rho_\alpha \left( U_N^0 \ast^{-1}, U_N^0 \ast^{-1}, (\cdot)), I(\cdot) \right)
\]
\[
\leq \rho_T^\alpha (X_N^*, X^*) + \rho_\alpha \left( X^*(U_N^0 \ast^{-1}, (\cdot)), X^*(\cdot) \right)
\]
\[
+ \sqrt{N} \rho_\alpha \left( U_N^0 \ast^{-1}, U_N^0 \ast^{-1}, (\cdot)), I(\cdot) \right).
\] (4.31)

Since \( X^*(\cdot) \) is (uniformly) continuous on \([0, T_\alpha]\), \( d_\alpha (X_N^*, X^*) = \rho_T^\alpha (X_N^*, X^*) \), and so by (4.24) it holds that \( \rho_T^\alpha (X_N^*, X^*) \to 0 \) a.s. as \( N \to \infty \), and therefore

\[
P\{ \omega \in \Lambda_N^\alpha : \rho_T^\alpha (X_N^*, X^*) \geq \epsilon \} \to 0
\] (4.32)

for all positive \( \epsilon \). In addition, the uniform continuity of \( X^* \) together with (4.28) imply that

\[
P\{ \omega \in \Lambda_N^\alpha : \rho_\alpha \left( X^*(U_N^0 \ast^{-1}, (\cdot)), X^*(\cdot) \right) \geq \epsilon \} \to 0
\] (4.33)

for each positive \( \epsilon \) as \( N \to \infty \).

Combining (4.27) and (4.31) - (4.33), it follows that

\[
P\{ \omega \in \Lambda_N^\alpha : \rho_\alpha (Y_N^*, -X^*) \geq \epsilon \}
\]
\[
\leq P\{ \omega \in \Lambda_N^\alpha : \rho_T^\alpha (X_N^*, X^*) + \rho_\alpha \left( X^*(U_N^0 \ast^{-1}, (\cdot)), X^*(\cdot) \right)
\]
\[
+ \sqrt{N} \rho_\alpha \left( U_N^0 \ast^{-1}, U_N^0 \ast^{-1}, (\cdot)), I(\cdot) \right) \geq \epsilon \}.
\]

19
\[
\leq \Pr[\omega \in \Lambda_N^\alpha : \rho_T(X^*_N, X^*) \geq \varepsilon/4] \\
+ \Pr[\omega \in \Lambda_N^\alpha : \rho_\alpha \left( \hat{\Delta}_N^{0}(-0)^{-1}, X^*(\cdot) \right) \geq \varepsilon/4] \\
+ \Pr[\omega \in \Lambda_N^\alpha : \sqrt{N} \rho_\alpha \left( \hat{\Delta}_N^{0}(-0)^{-1}, I(\cdot) \right) \geq \varepsilon/2] \\
\rightarrow 0 \text{ as } N \rightarrow \infty.
\] (4.34)

(The last inequality in (4.34) holds since for any two random variables \( R, S \), the relation \( \Pr[R+S \geq \varepsilon] = \Pr[R+S \geq \varepsilon, R \geq \varepsilon/2] + \Pr[R \geq \varepsilon/2] + \Pr[S \geq \varepsilon/2] \) holds. Apply this relation twice.) Finally, by (4.26) and (4.34), for arbitrary positive \( \varepsilon \), it holds that

\[
\Pr[\omega : \rho_\alpha(Y^*_N, -X^*) \geq \varepsilon] \\
\leq \Pr[\omega : \omega \in \Lambda_N^\alpha \text{ and } \rho_\alpha(Y^*_N, -X^*) \geq \varepsilon] + \Pr[\omega : \omega \notin \Lambda_N^\alpha] \\
\rightarrow 0 \text{ as } N \rightarrow \infty.
\] (4.35)

Therefore (4.29) holds, and we have proven that

\[
Y_N^w(p) = \sqrt{N} \hat{\Delta}_N^{0}(-0)^{-1}(p - p) \underset{w}{\rightarrow} -X(p), \quad 0 \leq p \leq \alpha,
\] (4.36)

where \( X(\cdot) \) is a mean zero Gaussian process on \([0, \alpha]\) with covariance

\[
\text{Cov}(X(s), X(t)) = (1-s)(1-t) \int_0^s \frac{1}{(1-w)^2(1-H_0^{w-1}(w))} \, dw, \quad s \leq t.
\] (4.37)
Since
\[ \hat{U}_N^{-1}(p) = F^0(F_N^{-1}(p)), \]  \hspace{1cm} (4.38)

the result (4.36) may be rewritten as
\[ \sqrt{N} \left( F^0(F_N^{-1}(p)) - p \right) \xrightarrow{w} -X(p), \quad 0 \leq p \leq \alpha, \]  \hspace{1cm} (4.39)

which is Theorem 1.

5. Proof of Theorem 2.

To prove the weak convergence of the random process
\[ \sqrt{N} \left( F_N^{-1}(\cdot) - F^0(\cdot) \right), \]  the one must assume a little more.

Specifically, we now assume \( F^0(\cdot) \) continuous with a continuous derivative at each \( p, 0 \leq p \leq \alpha \). Then, with
\[ Y_N(p) = \sqrt{N} \left( F^0(F_N^{-1}(p)) - p \right) = \sqrt{N}(\hat{U}_N^{-1}(p) - p), \]  \hspace{1cm} (5.1)

for \( \hat{U}_N^{-1}(p) \neq p \), write
\[ \sqrt{N} \left( F_N^{-1}(p) - F^0(\gamma) \right) = Y_N(p) \cdot \left( \frac{\partial F^0(\gamma)}{\partial x} \right)_{x=p} \]  \hspace{1cm} (5.2)

\[ + Y_N(p) \cdot \left[ \frac{F_N^{-1}(p) - F^0(\gamma)}{\hat{U}_N^{-1}(p) - p} - \left( \frac{\partial F^0(\gamma)}{\partial x} \right)_{x=p} \right]. \]

Lemma 6. Under the conditions of Theorem 1, and the additional assumption that \( F^0 \) has a continuous derivative at each \( p \),
\[ 0 \leq p \leq \alpha, \text{ it holds that} \]
\[ Y_N(p) \cdot \left( \frac{\partial F^{-1}}{\partial x} \right)_{x=p} \overset{w}{=} (-X(p)) \cdot \left( \frac{\partial F^{-1}}{\partial x} \right)_{x=p}, \quad 0 \leq p \leq \alpha, \]
\[ \text{as } N \to \infty. \quad (5.3) \]

**Proof.** As in the proof of Theorem 1, let \( Y_N^*(\cdot), \quad N = 1, 2, \ldots, \)
and \(-X^*(\cdot)\) be the Skorohod equivalent random functions on \([0, \alpha]\) satisfying \( \mathcal{L}(Y_N^*) = \mathcal{L}(Y_N), \quad N = 1, 2, \ldots, \)
\( \mathcal{L}(-X^*) = \mathcal{L}(-X), \)
and \( d_\alpha(Y_N^*, -X^*) \to 0 \) a.s. as \( N \to \infty, \) where \( d_\alpha \) is the Skorohod metric on \([0, \alpha].\) By the continuity a.s. of \(-X^*(\cdot),\) it also holds that \( \rho_\alpha(Y_N^*, -X^*) \to 0 \) a.s. Thus, for every positive \( \varepsilon, \)

\[ P \left\{ \sup_{0 \leq p \leq \alpha} \left| \left( \frac{\partial F^{-1}}{\partial x} \right)_{x=p} \cdot Y_N^*(p) - \left( \frac{\partial F^{-1}}{\partial x} \right)_{x=p} \cdot (-X^*(p)) \right| \geq \varepsilon \right\} \]
\[ \leq P \left\{ \sup_{0 \leq p \leq \alpha} \left| \frac{\partial F^{-1}}{\partial x} \right|_{x=p} \cdot \sup_{0 \leq p \leq \alpha} \left| Y_N^*(p) - (-X^*(p)) \right| \geq \varepsilon \right\} \]
\[ = P \left\{ \sup_{0 \leq p \leq \alpha} \left| Y_N^*(p) - (-X^*(p)) \right| \geq \frac{\varepsilon}{M} \right\}, \quad (5.4) \]

since the continuity of \( \frac{\partial F^{-1}}{\partial x} \) on compact \([0, \alpha]\) implies

\[ \sup_{0 \leq p \leq \alpha} \left| \left( \frac{\partial F^{-1}}{\partial x} \right)_{x=p} \right| \equiv M < \infty. \quad (5.5) \]

The last expression in (5.4) converges to zero as \( N \to \infty \) since
\( \rho_\alpha(Y_N^*, -X^*) \to 0 \) a.s. Thus, by an easy application of Billingsley's
Theorem 4.1 (1968; p. 25), it holds that

\[ Y^*_N(p) \cdot \left( \frac{\partial F^{-1}}{\partial x} \right)_{x=p} \overset{W}{\Rightarrow} -X^*(p) \cdot \left( \frac{\partial F^{-1}}{\partial x} \right)_{x=p}, \quad 0 \leq p \leq \alpha, \quad (5.6) \]

and so the corresponding result holds for the original (unstarred) random functions. \[ \text{q.e.d.} \]

Lemma 7. Under the conditions of Lemma 6, for arbitrary positive \( \varepsilon \),

\[
P \left\{ \sup_{0 < p < \alpha} \left| Y^*_N(p) \cdot \left[ \frac{\hat{F}^{-1}_N(p) - F^{-1}(p)}{\hat{U}^{-1}_N(p) - p} - \left( \frac{\partial F^{-1}}{\partial x} \right)_{x=p} \right] \right| \geq \varepsilon \right\} \to 0 \quad \text{as} \quad N \to \infty.
\] \[
(5.7)
\]

Proof. For convenience, define

\[
A_N(p) = \frac{\hat{F}^{-1}_N(p) - F^{-1}(p)}{\hat{U}^{-1}_N(p) - p}, \quad 0 \leq p \leq \alpha.
\] \[
(5.8)
\]

Recalling that \( \hat{F}^{-1}_N(p) = F^{-1}\left( \hat{U}^{-1}_N(p) \right) \), we see that \( A_N(p) \) is just the differential quotient

\[
\frac{F^{-1}_N(\hat{U}^{-1}_N(p)) - F^{-1}(p)}{\hat{U}^{-1}_N(p) - p}.
\] \[
(5.9)
\]
By Lemma 5, \( p_\alpha (\alpha^N_0, 1) \to 0 \), and so, by the assumed (uniform) continuity of the derivative of \( F_0^{-1} \) at each \( p \), \( 0 \leq p \leq \alpha \), it follows that

\[
\sup_{0 \leq p \leq \alpha} |A_N(p) - \left( \frac{\partial F_0^{-1}}{\partial x} \right)_{x=p} | \overset{p}{\to} 0
\]  

(5.10)

as \( N \to \infty \). Furthermore, as \( N \to \infty \),

\[
\sup_{0 \leq p \leq \alpha} |Y_N(p)| \overset{N}{\to} \sup_{0 \leq p \leq \alpha} |-X(p)|
\]  

(5.11)

Hence, by (5.10) and (5.11), it follows that

\[
P \left\{ \sup_{0 \leq p \leq \alpha} |Y_N(p)| \cdot \left| A_N(p) - \left( \frac{\partial F_0^{-1}}{\partial x} \right)_{x=p} \right| \geq \varepsilon \right\} 
\leq P \left\{ \sup_{0 \leq p \leq \alpha} |Y_N(p)| \cdot \sup_{0 \leq p \leq \alpha} |A_N(p) - \left( \frac{\partial F_0^{-1}}{\partial x} \right)_{x=p} | \geq \varepsilon \right\}
\]  

(5.12)

\[
\to 0 \quad \text{as} \quad N \to \infty .
\]

q.e.d.

Applying Lemmas 6 and 7 and Theorem 4.1 of Billingsley (1968) to the identity (5.2) completes the proof of Theorem 2.

The assumptions of Corollary 1 imply that \( \left( \frac{\partial F_0^{-1}}{\partial x} \right)_{x=p} \) exists, and equals \( f^0[F_0^{-1}(p)] \). [cf., Hájek and Šidák (1967), p. 34.] Thus, (2.8) follows directly from (2.7).

q.e.d.

24
REFERENCES


