The Annals of Statistics
1996, Vol. 24, No. 2, 641–646

A COUNTEREXAMPLE TO A CONJECTURE CONCERNING THE HALL–WELLNER BAND

BY KANI CHEN AND ZHILIANG YING[1]

Hong Kong University of Science and Technology and Rutgers University

Hall and Wellner proposed a natural extension of the Kolmogorov–Smirnov simultaneous confidence band for survival curve using the Kaplan–Meier estimator. They and Gill conjectured that the confidence band holds for all t up to the last observed failure time. A counterexample is given herein, showing that this may not always be true.

1. Introduction. Let $X_1, \dots, X_n$ be independent and identically distributed (i.i.d.) positive random variables with a continuous distribution F. A theorem of Donsker [Billingsley (1968), Theorem 16.4] states that $\sqrt{n}\,(\hat{F}_n^* - F)$ converges in $\mathscr{D}[0,\infty]$ to $B_0 \circ F$, where $\hat{F}_n^*(t) = n^{-1}\sum_{i=1}^n I_{\{X_i \le t\}}$ is the usual empirical distribution, $\mathscr{D}[0,\infty]$ is the space, equipped with the Skorohod topology, of functions which are left-continuous with right limits and B_0 is the Brownian bridge process, and where $\circ$ denotes functional composition. It, among other things, yields a very important result for the Kolmogorov–Smirnov statistic; that is, $\sup_t|\sqrt{n}\,(\hat{F}_n^*(t) - F(t))|$ converges to $\sup_{1 \le u \le 1}|B_0(u)|$, the distribution of which is well known and has been tabulated.

In survival analysis, the X_i are often subject to independent right-censoring so that the observed data are $\tilde{X}_i = \min\{X_i, U_i\}$ and $\delta_i = I_{\{X_i \le U_i\}}$, $i = 1, \dots, n$, where the U_i are i.i.d. positive censoring variables that are independent of the X_i and have, possibly discontinuous, distribution G. In this case, the analogue of $\hat{F}_n^*$ is the Kaplan–Meier estimator defined by

$$\hat{F}_n(t) = 1 - \prod_{s \le t}\left[1 - \frac{\Delta N_n(s)}{Y_n(s)}\right], \tag{1.1}$$

where $\Delta N_n(t) = N_n(t) - N_n(t-)$ and

$$N_n(t) = \sum_{i=1}^n \delta_i I_{\{\tilde{X}_i \le t\}}, \qquad Y_n(t) = \sum_{i=1}^n I_{\{\tilde{X}_i \ge t\}}. \tag{1.2}$$

Let $H(t) = 1 - \overline{H}(t)$ and $\overline{H}(t) = \overline{F}(t)\overline{G}(t)$. Here and in the sequel, $\overline{F}(t) = 1 - F(t)$ and $\overline{G}(t) = 1 - G(t)$. Let $\tau_H = \sup_t\{t\colon H(t) < 1\}$. Breslow and Crowley (1974) showed that, for any $\tau < \tau_H$, $\sqrt{n}\,(\hat{F}_n - F)$ converges weakly in

Received December 1994; revised April 1995.
[1]Supported by NSF Grant DMS-92-07730.
AMS 1991 *subject classifications*. Primary 62E20; secondary 62G30.
Key words and phrases. Censored survival data, Kolmogorov–Smirnov statistic, Brownian bridge, Doob's transformation, confidence band, tightness, weak convergence.

$\mathscr{D}[0, \tau]$ to a Gaussian process B_{bc}, which in general is *not* a time-rescaled Brownian bridge process. Based on Doob's transformation which relates the Brownian bridge to the Brownian motion, Hall and Wellner (1980) argued that a natural analogue of the classical Kolmogorov–Smirnov statistic for the censored data should be $\sup_{t \le X_n^*} |W_n(t)|$, where $X_n^* = \max_{i \le n}\{\delta_i \tilde{X}_i\}$ is the last observed failure time and

$$W_n(t) = \sqrt{n}\,\frac{1 - \hat{K}_n(t)}{1 - \hat{F}_n(t)}\left[\hat{F}_n(t) - F(t)\right] \tag{1.3}$$

with

$$\hat{K}_n = \frac{\hat{C}_n}{1 + \hat{C}_n} \quad \text{and} \quad \hat{C}_n(t) = \int_0^t \frac{n\,dN_n(s)}{Y_n(s)(Y_n(s) - 1)}. \tag{1.4}$$

Following Gill (1983), define the stopped process $W_n^*(t) = W_n(\min(X_n^*, t))$. Clearly, $\hat{K}_n$ and $\hat{C}_n$ are estimators of

$$K(t) = \frac{C(t)}{1 + C(t)} \quad \text{and} \quad C(t) = \int_0^t \frac{dF(s)}{\overline{H}(s-)\overline{F}(s)}.$$

The reason that W_n is asymptotically distribution-free is that, from the weak convergence result of Breslow and Crowley,

$$W_n \to_{\mathscr{D}[0, \tau]} B_0 \circ K, \tag{1.5}$$

for any $\tau < \tau_H$ [Hall and Wellner (1980)]. However, in order to justify the use of $\sup_{t \le X_n^*} |W_n(t)|$, one needs to show that the weak convergence (1.5) can be extended to the last observed failure time X_n^*. This was proved to be true by Gill (1983) under the condition

$$\int_0^{\tau_H} \frac{dF(t)}{1 - G(t-)} < \infty. \tag{1.6}$$

This condition appears to be more than necessary since Gill (1983) and Ying (1989) showed that

$$\sqrt{n}\,\frac{1 - K}{1 - F}\left(\hat{F}_n - F\right) \to_{\mathscr{D}[0, \tau_H]} B_0 \circ K \tag{1.7}$$

without (1.6). In view of (1.7), it is natural to speculate that the weak convergence may hold for W_n up to the last point without imposing (1.6); see Hall and Wellner [(1980), page 137] and Gill (1983, 1994).

Recently, Kaplan–Meier analysts have made significant progress, including the elegant results of Gill (1983), Wang (1987) and Stute and Wang (1993), on th endpoint behavior of the product-limit estimator for censored data. Yet the convergence of W_n^* remains open [Gill (1994)]. This open "problem is rather important since so far there is no theorem justifying 'common practice,' which is to compute a confidence band on a large interval whose endpoint σ is such that $Y_n(\sigma)$ is rather small" [Gill (1994), page 162].

In this note, we present a counterexample to show that W_n^* does not in general converge weakly on the whole interval.

2. A counterexample. Our construction of the counterexample is conceptually rather simple: we find an integer subsequence $n_k \uparrow \infty$ and an increasing sequence $a_k \uparrow \tau_H$ such that, for some $\varepsilon > 0$ and any $L > 0$,

$$\liminf_{k\to\infty} P\Big\{\big|W_{n_k}^*(2^{-1}(a_k + a_{k-1})) - W_{n_k}^*(a_{k-1})\big| \geq L\Big\} \geq \varepsilon, \tag{2.1}$$

violating a necessary condition for the tightness of W_n^* [Billingsley (1968), Theorem 15.3]. This shows that W_n^* is not tight and therefore does not converge weakly in $\mathscr{D}[0, \tau_H]$ [Billingsley (1968), Theorem 15.3]. Furthermore, by the triangle inequality, (2.1) implies that $\sup|W_n^*|$ does not converge to $\sup|B_0 \circ K|$ and thus the validity of the Hall–Wellner band cannot be extended to the last failure time.

Without loss of generality, we may assume that F is uniform on $[0,1]$. For definiteness, define

$$\begin{aligned} n_k &= (k+1)^{k^k}, & r_k &= (k+1)^k, \\ a_k &= \tau \frac{\sum_{i=1}^k r_i^{-1}}{\sum_{j=1}^\infty r_j^{-1}}, & p_k &= \frac{r_k}{n_k}\left(\sum_{i=1}^\infty \frac{r_i}{n_i}\right)^{-1}, \end{aligned} \tag{2.2}$$

where τ is any number in $(0,1]$. It follows that $a_k \uparrow \tau$ and $\sum p_j = 1$. Define the censoring distribution by $P(U_1 = a_k) = p_k$, $k = 1, 2, \dots$. Clearly, $\tau_H = \tau$. In addition, it can be verified easily that, as $k \to \infty$,

$$n_k^{-\varepsilon} r_k^2 n_{k-1} \to 0 \quad \text{for any } \varepsilon > 0, \tag{2.3}$$

$$\frac{p_k}{\sum_{j=k}^\infty p_j} \to 1, \tag{2.4}$$

$$\frac{n_k^{-1}}{\sum_{j=k}^\infty n_j^{-1}} \to 1. \tag{2.5}$$

Now let A_k be the event that, among the n_k observations, the largest uncensored failure time falls into the interval $(2^{-1}(a_k + a_{k-1}), a_k)$ and all other uncensored failure times are in $[0, a_{k-1})$. By symmetry,

$$\begin{aligned} P(A_k) &= n_k P(A_k, X_1 \text{ is the largest uncensored failure time}) \\ &= n_k P\big(2^{-1}(a_k + a_{k-1}) < X_1 < a_k,\, X_1 \leq U_1\big) \\ &\quad \times [1 - P(a_{k-1} \leq X_2 \leq U_2)]^{n_k - 1} \\ &= n_k \frac{a_k - a_{k-1}}{2} \sum_{j=k}^\infty p_j \left[1 - \sum_{l=k}^\infty (a_l - a_{l-1}) \sum_{i=l}^\infty p_i\right]^{n_k - 1} \end{aligned}$$

$$= (1+o(1))\frac{n_k\tau r_k^{-1}}{2\sum_{j=1}^{\infty} r_j^{-1}} p_k\left[1-(1+o(1))\sum_{l=k}^{\infty}\frac{\tau r_l^{-1}p_l}{\sum_{i=1}^{\infty} r_i^{-1}}\right]^{n_k-1} \quad \text{[by (2.4)]}$$

(2.6)

$$= (1+o(1))\frac{\xi}{2}\left[1-(1+o(1))\xi\sum_{l=k}^{\infty} n_l^{-1}\right]^{n_k-1}$$

$$= (1+o(1))\frac{\xi}{2}\left[1-(1+o(1))\frac{\xi}{n_k}\right]^{n_k-1} \quad \text{[by (2.5)]}$$

$$= (1+o(1))\frac{\xi}{2}\exp(-\xi),$$

where $\xi = \tau/(\sum_{j=1}^{\infty} r_j^{-1}\sum_{l=1}^{\infty} r_l/n_l)$.

Next we show that $|W_{n_k}^*(2^{-1}(a_k+a_{k-1})) - W_{n_k}^*(a_{k-1})|$ tends to ∞ on A_k, thus proving (2.1). Since on A_k the largest uncensored observation is no smaller than $(a_k+a_{k-1})/2$ and there is no uncensored observation in $(a_{k-1},(a_k+a_{k-1})/2)$, it follows that, again on A_k,

(2.7)

$$\begin{aligned}\left|W_{n_k}^*\left(\frac{a_k+a_{k-1}}{2}\right) - W_{n_k}^*(a_{k-1})\right| &= n_k^{1/2}\frac{1-\hat{K}_{n_k}(a_{k-1})}{1-\hat{F}_{n_k}(a_{k-1})}\\ &\quad\times\left[F\left(\frac{a_k+a_{k-1}}{2}\right) - F(a_{k-1})\right]\\ &\geq \left[1-\hat{K}_{n_k}(a_{k-1})\right] n_k^{1/2} 2^{-1}(a_k - a_{k-1})\\ &= \frac{n_k^{1/2}\tau}{2r_k\left[1+\hat{C}_{n_k}(a_{k-1})\right]\sum_{j=1}^{\infty} r_j^{-1}},\end{aligned}$$

recalling that $\hat{C}_n(t) = \int_0^t n\,dN_n(u)/[Y_n(u)(Y_n(u)-1)]$. On A_k, $t \leq a_{k-1}$ and $\Delta N_{n_k}(t) \neq 0$ imply that $Y_{n_k}(t) \geq 2$, or $Y_{n_k}(t) - 1 \geq Y_{n_k}(t)/2$, so $\hat{C}_{n_k}(t) \leq 2\int_0^t n_k\, dN_{n_k}(u)/Y_{n_k}^2(u)$. Thus, to show $n_k^{1/2}/[r_k(1+\hat{C}_{n_k}(a_{k-1}))] \to_P \infty$ on A_k, it suffices to show that

(2.8) $$\frac{r_k}{n_k^{1/2}}\left[1+2\int_0^{a_{k-1}}\frac{n_k\,dN_{n_k}(t)}{Y_{n_k}^2(t)}\right] \to_P 0, \quad \text{or} \quad r_k n_k^{1/2}\int_0^{a_{k-1}}\frac{dN_{n_k}(t)}{Y_{n_k}^2(t)} \to_P 0.$$

Since $N_{n_k}(t)$ has compensator $\int_0^t Y_{n_k}(s)(1-s)^{-1}\,ds$, we have

$$\begin{aligned}E\int_0^{a_{k-1}}\frac{dN_{n_k}(t)}{Y_{n_k}^2(t)} &= E\int_0^{a_{k-1}}\frac{I_{(Y_{n_k}(t)\geq 1)}}{Y_{n_k}(t)}\frac{dt}{1-t}\\ &= \int_0^{a_{k-1}}\sum_{i=1}^{n_k}\frac{1}{i}\binom{n_k}{i}\overline{H}^i(t-)H^{n_k-i}(t-)\frac{dt}{1-t}\end{aligned}$$

$$
\begin{aligned}
&\le \int_0^{a_{k-1}} \sum_{i=1}^{n_k} \frac{2}{i+1}\binom{n_k}{i} \overline{H}^i(t-)H^{n_k-i}(t-)\frac{dt}{1-t}\\
&= \frac{2}{n_k+1}\int_0^{a_{k-1}} \sum_{i=1}^{n_k} \binom{n_k+1}{i+1} \overline{H}^{i+1}(t-)\\
&\quad \times H^{n_k+1-(i+1)}(t-)\frac{dt}{\overline{H}(t-)(1-t)}\\
&\le \frac{2}{n_k+1}\int_0^{a_{k-1}} \frac{dt}{\overline{H}(t-)(1-t)}\\
&\le \frac{2}{(n_k+1)\overline{G}(a_{k-1}-)(1-a_{k-1})}\\
&\le \frac{2}{(n_k+1)p_{k-1}(a_k-a_{k-1})}\\
&= \frac{2r_k n_{k-1}}{\xi(n_k+1)r_{k-1}}.
\end{aligned}
$$

Therefore,

$$
E\left[r_k n_k^{1/2}\int_0^{a_{k-1}} \frac{dN_{n_k}(t)}{Y_{n_k}^2(t)}\right] \le \frac{2r_k^2 n_{k-1}}{\xi\sqrt{n}_k r_{k-1}},
$$

which converges to 0 by (2.3). Hence (2.8) holds. From (2.7) and (2.8) we conclude (2.1), proving that W_n^* cannot be tight.

3. Remarks. By taking a monotone transformation, the failure-time distribution F in the counterexample can be any continuous distribution function. The censoring distribution G should be changed accordingly.

A similar counterexample may be produced with a continuous censoring distribution. This can be done by spreading mass p_k at a_k evenly to interval $[a_k - \varepsilon_0/n_k, a_k + \varepsilon_0/n_k]$, with ε_0 being sufficiently small. It is easy to see that (2.1) still holds.

If we regard the left-hand side of (1.7) as a *normalized* Kaplan–Meier process, then W_n may be viewed as a *Studentized* Kaplan–Meier process. Thus our counterexample reveals that the Studentized Kaplan–Meier process does not in general converge on the whole interval, even though the normalized process does.

Acknowledgment. The authors would like to thank the referees for their careful reading and constructive comments.

REFERENCES

BILLINGSLEY, P. (1968). *Weak Convergence of Probability Measures*. Wiley, New York.

BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product-limit estimates under random censorship. *Ann. Statist.* **2** 437–453.

DONSKER, M. D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov–Smirnov theorem. *Ann. Mat. Statist.* **23** 277–281.

GILL, R. D. (1983). Large sample behaviour of the product-limit estimator on the whole line. *Ann. Statist.* **11** 49–58.

GILL, R. D. (1994). Lectures on survival analysis. *Ecole d'Eté de Probabilités de Saint-Flour XXII. Lecture Notes in Math.* **1581** 115–141. Springer, Berlin.

HALL, W. J. and WELLNER, J. A. (1980). Confidence bands for a survival curve from censored data. *Biometrika* **67** 133–143.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

STUTE, W. and WANG, J.-L. (1993). The strong law under random censorship. *Ann. Statist.* **21** 1591–1607.

WANG, J. G. (1987). A note on the uniform consistency of the Kaplan–Meier estimator. *Ann. Statist.* **15** 1313–1316.

YING, Z. (1989). A note on the asymptotic properties of the product-limit estimator on the whole line. *Statist. Probab. Lett.* **7** 311–314.

DEPARTMENT OF MATHEMATICS
HKUST
CLEAR WATER BAY
KOWLOON, HONG KONG

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
HILL CENTER, BUSCH CAMPUS
PISCATAWAY, NEW JERSEY 08855