

## UNIVERSAL GAUSSIAN APPROXIMATIONS UNDER RANDOM CENSORSHIP<sup>1</sup>

BY SÁNDOR CSÖRGŐ

*University of Michigan*

*Dedicated to Endre Csáki on his sixtieth birthday*

Universal Gaussian approximations are established for empirical cumulative hazard and product-limit processes under random censorship. They hold uniformly up to some large order statistics in the sample, with the approximation rates depending on the order of these statistics, and require no assumptions on the censoring mechanism. Weak convergence results and laws of the iterated logarithm follow on the whole line if the respective processes are stopped at certain large order statistics, depending on the type of result. Some new consequences and negative results for confidence-band construction are discussed. Some new uniform consistency rates up to large order statistics are also derived and shown to be universally best possible for a wide range of tail order statistics.

**1. Main results.** Let  $X_1, X_2, \dots, X_n$  be independent random variables with distribution function  $F(x) = P\{X \leq x\}$ ,  $x \in \mathbb{R}$ . An independent sequence of independent random variables  $Y_1, Y_2, \dots, Y_n$  with distribution function  $G$  censors them on the right, so that at each stage  $n$  we can only observe  $Z_j = \min(X_j, Y_j)$  and  $\delta_j = I\{X_j \leq Y_j\}$ ,  $j = 1, \dots, n$ , where  $I\{A\}$  stands for the indicator of an event  $A$ . Let  $H$  be the distribution function of  $Z$ , where  $(Z, \delta) = (Z_1, \delta_1)$ , so that  $1 - H = (1 - F)(1 - G)$ . Let  $Z_{1,n} \leq \dots \leq Z_{n,n}$  denote the order statistics pertaining to  $Z_1, \dots, Z_n$  with the corresponding concomitants  $\delta_{1,n}, \dots, \delta_{n,n}$ , so that  $\delta_{j,n} = \delta_i$  if  $Z_{j,n} = Z_i$ . Consider

$$\Lambda(x) = \int_{-\infty}^x \frac{dF(y)}{1 - F_-(y)} = \int_{-\infty}^x \frac{d\tilde{H}(y)}{1 - H_-(y)},$$

the cumulative hazard function, and its Nelson–Aalen estimator [Nelson (1972) and Aalen (1976)]

$$\Lambda_n(x) = \int_{-\infty}^x \frac{d\tilde{H}_n(y)}{1 - H_n(y-)} = \frac{1}{n} \sum_{j=1}^n \frac{\delta_{j,n} I\{Z_{j,n} \leq x\}}{1 - H_n(Z_{j,n}-)},$$

where  $\tilde{H}(x) = P\{Z \leq x, \delta = 1\} = \int_{-\infty}^x [1 - G_-(y)] dF(y)$ ,

$$H_n(x) = \frac{1}{n} \sum_{j=1}^n I\{Z_j \leq x\}$$

---

Received August 1995; revised March 1996.

<sup>1</sup>Work partially supported by NSF Grants DMS-92-08067 and DMS-96-25732.

AMS 1991 subject classifications. Primary 62G05, 62G30; secondary 60F17, 60F15.

Key words and phrases. Random censorship, cumulative hazard, product-limit, Gaussian approximations.

and

$$\tilde{H}_n(x) = \frac{1}{n} \sum_{j=1}^n I\{Z_j \leq x, \delta_j = 1\}, \quad x \in \mathbb{R},$$

and where, for a right-continuous real function  $f$  on  $\mathbb{R}$ , the function  $f_-(x) = f(x-)$ ,  $x \in \mathbb{R}$ , denotes its left-continuous version. Finally, for a continuous  $F$ , consider

$$\hat{F}_n(x) = 1 - \prod_{j=1}^n \left[ 1 - \frac{\delta_{j,n}}{n-j+1} \right]^{I\{Z_{j,n} \leq x\}} = 1 - \prod_{j=1}^n \left[ 1 - \frac{\delta_{j,n} I\{Z_{j,n} \leq x\}}{n-j+1} \right],$$

the Kaplan–Meier (1958) product-limit estimator of  $F$ , and the asymptotic variance

$$d(x) = \int_{-\infty}^x \frac{d\tilde{H}(y)}{[1 - H_-(y)]^2} = \int_{-\infty}^x \frac{dF(y)}{[1 - F_-(y)]^2 [1 - G_-(y)]}$$

for any  $F$ , with its empirical counterpart, associated with Greenwood’s formula from 1926,

$$d_n(x) = \int_{-\infty}^x \frac{d\tilde{H}_n(y)}{[1 - H_n(y-)]^2} = \frac{1}{n} \sum_{j=1}^n \frac{\delta_{j,n} I\{Z_{j,n} \leq x\}}{[1 - H_n(Z_{j,n}-)]^2}, \quad x \in \mathbb{R}.$$

We use the integral convention that  $\int_{-\infty}^x = \int_{(-\infty, x)}$  or  $\int_{(-\infty, x]}$  according as integration is with respect to (the measure induced by) a left- or right-continuous function. Also,  $\log$  will be the natural logarithm and  $\log^u x = (\log x)^u$ ,  $x > 1$ ,  $u \in \mathbb{R}$ . For a sequence of random variables  $\xi_n$  and positive constants  $a_n$ , we say that  $\xi_n = \mathcal{O}(a_n)$  or  $\xi_n = o(a_n)$  almost surely if  $\limsup_{n \rightarrow \infty} |\xi_n|/a_n \leq C$  for a finite deterministic constant  $C > 0$  or for  $C = 0$ , respectively, and write  $\xi_n = \mathcal{O}_P(a_n)$  if  $\lim_{x \rightarrow \infty} \limsup_{n \rightarrow \infty} P\{|\xi_n| > a_n x\} = 0$ . We say that  $\{a_n\}$  is asymptotically nonincreasing if  $\lim_{n \rightarrow \infty} a_n/a_n^* = 1$  for a nonincreasing sequence  $\{a_n^*\}$  of numbers. Let  $\tau_J = \sup\{x: J(x) < 1\} \leq \infty$  for a right-continuous distribution function  $J$  and denote by  $J^{-1}(s) = \inf\{x \in \mathbb{R}: J(x) \geq s\}$ ,  $0 < s < 1$ , its left-continuous quantile function. Throughout,  $\{k_n\}$  will be a sequence of integers such that  $1 \leq k_n < n$  and the almost sure statements require the condition

- (\*)  $k_n \geq \log n$  for all  $n$  large enough and  $\{k_n/n\}$  is asymptotically nonincreasing.

Finally, a standard Wiener process  $W(s)$ ,  $0 \leq s < \infty$ , a Brownian bridge  $B(s)$ ,  $0 \leq s \leq 1$ , a standard Wiener sheet  $W(s, u)$ ,  $0 \leq s, u < \infty$ , and a Kiefer sheet  $K(s, u)$ ,  $0 \leq s \leq 1, 0 \leq u < \infty$ , are mean-zero Gaussian processes with covariance functions  $E(W(s)W(t)) = s \wedge t$ ,  $E(W(s, u)W(t, v)) = (s \wedge t)(u \wedge v)$ ,  $0 \leq s, t, u, v < \infty$ , and  $E(B(s)B(t)) = (s \wedge t) - st$ ,  $E(K(s, u)K(t, v)) = [(s \wedge t) - st](u \wedge v)$ ,  $0 \leq s, t \leq 1, 0 \leq u, v < \infty$ , where  $a \wedge b = \min(a, b)$ ,  $a, b \in \mathbb{R}$ .

The classical weak-convergence results of Breslow and Crowley (1974) for  $\Lambda_n$  and  $\hat{F}_n$ , anticipated by Efron (1967) and constituting the basis for the whole

theory and practice with randomly censored data, hold over a fixed half-line  $(-\infty, T]$ , where  $T < \tau_H$ . Of course, such a choice of  $T$  asymptotically excludes a whole fixed proportion of the data. By natural desire, however, practitioners commonly use the theory up to  $Z_{n,n}$  or to the largest uncensored observation. In a remarkable paper, Gill (1983) has shown that a version of the basic weak-convergence result does hold on the whole half-line  $[-\infty, Z_{n,n}]$ , provided that condition (2.8) below on the lightness of censorship is satisfied. Gill's condition is violated in many standard situations and is difficult to check in practice. What is a reasonable compromise, or a conservative rule of thumb, between a fixed  $T < \tau_H$  and  $Z_{n,n}$ ? Stute (1994b), posing the problem in this light and, with his technical groundwork contained in Stute (1994a), improving an earlier  $U$ -statistic approach in Burke, Csörgő and Horváth (1988), has derived important results for the almost sure representation of  $\Lambda_n$  and  $\hat{F}_n$  in terms of sums of independent and identically distributed functions with a uniform asymptotic error rate up to certain order statistics  $Z_{n-k_n, n}$ . The essence of Stute's approach is to "let the data speak for themselves." In other words, instead of assuming unverifiable conditions on the lightness of censoring, one tries to assess what may be said in a worst-case situation and thus derive results that are valid in all possible scenarios. The question then is: how far up can one go, that is, how small can  $k_n$  be chosen?

In this paper, using results of Csáki (1977) and Wellner (1978), we improve Stute's strong rates, as stated in Proposition 1 of Section 3. The problem of centering this representation of  $\Lambda_n$  at  $\Lambda$  is delicate, and to complement Stute's strong rates for that problem in Proposition 3, we use the continuous-time martingale approach of Gill (1980) to obtain rates of convergence in probability. In Propositions 2 and 4 we derive a comparable theory for the variance estimator  $d_n$ . Having all these, we then return to the primary problem of Gaussian approximations in Burke, Csörgő and Horváth (1981, 1988). These were originally designed to get ideas about the precision of the classical weak-convergence results of Breslow and Crowley (1974). With the changed scope and greatly improved ingredients, the question now is: how far out do Gaussian approximations hold universally? The question is of particular relevance now that Chen and Ying (1996) have constructed examples in which Gill's (1983) result does not hold up to the largest uncensored observation when condition (2.8) is violated and, with  $[t]$  denoting the usual integer part of  $t > 0$ , an elaboration on their construction in Remark 5.2 below shows that it does not necessarily hold up to  $Z_{n-[6.597(\log n)/\log \log \log n], n}$  either. For further introduction, motivation and a vast literature (although some of these are on black holes and many listed twice, entering the keywords "censorship," "censoring" and "censored" into the AMS MathSci database, one is given over 4800 titles), we refer to the above papers, Andersen Borgan, Gill and Keiding (1993), Gill (1994) and many of our further references.

The main results, for the respective approximations of  $\Lambda_n$  and  $\hat{F}_n$ , are in the following two theorems, where we put  $D(x) = d(x)/[1 + d(x)]$  for all  $x \leq \tau_H$  and  $K_n(\cdot) = K(\cdot, n)/\sqrt{n}$  for a Kiefer sheet  $K(\cdot, \cdot)$ . A further important notational convention used throughout is that all  $\mathcal{O}_p(\cdot)$  statements hold for

all integers  $1 \leq k_n < n$ , as  $n \rightarrow \infty$ , whereas, unless otherwise stated, all unspecified limit-superior,  $\mathcal{O}(\cdot)$  and  $o(\cdot)$  statements hold almost surely under condition (\*) as  $n \rightarrow \infty$ .

THEOREM 1. *We have*

$$(1.1) \quad \sup_{x \leq Z_{n-k_n, n}} |\Lambda_n(x) - \Lambda(x)| = \begin{cases} \mathcal{O}_P\left(\frac{1}{\sqrt{k_n}}\right), \\ \mathcal{O}\left(\frac{\sqrt{\log n}}{\sqrt{k_{2n}}}\right), \end{cases}$$

and

$$(1.2) \quad \sup_{x \leq Z_{n-k_n, n}} |d_n(x) - d(x)| = \begin{cases} \mathcal{O}_P\left(\frac{n}{k_n^{3/2}}\right), \\ \mathcal{O}\left(\frac{n\sqrt{\log n}}{k_{2n}^{3/2}}\right). \end{cases}$$

Furthermore, on suitable probability spaces, there exist a sequence  $\{W_n(\cdot)\}$  of standard Wiener processes, a standard Wiener sheet  $W(\cdot, \cdot)$ , a sequence  $\{B_n(\cdot)\}$  of Brownian bridges and a Kiefer sheet  $K(\cdot, \cdot)$  such that if  $F$  is continuous on  $\mathbb{R}$ , then

$$(1.3) \quad \sup_{x \leq Z_{n-k_n, n}} |\sqrt{n}[\Lambda_n(x) - \Lambda(x)] - W_n(d(x))| = \begin{cases} \mathcal{O}_P\left(\frac{\sqrt{n} \log n}{k_n}\right), \\ \mathcal{O}\left(\frac{\sqrt{n} \log n}{k_{2n}}\right), \end{cases}$$

$$(1.4) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \sqrt{n}[\Lambda_n(x) - \Lambda(x)] - \frac{W(d(x), n)}{\sqrt{n}} \right| = \mathcal{O}\left(\frac{\sqrt{n} \log^2 n}{k_{2n}}\right),$$

$$(1.5) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \sqrt{n} \frac{\Lambda_n(x) - \Lambda(x)}{1 + d_n(x)} - B_n(D(x)) \right| = \mathcal{O}_P\left(\frac{n}{k_n^{3/2}} + \frac{\sqrt{n} \log n}{k_n}\right)$$

and

$$(1.6) \quad \begin{aligned} & \sup_{x \leq Z_{n-k_n, n}} \left| \sqrt{n} \frac{\Lambda_n(x) - \Lambda(x)}{1 + d_n(x)} - K_n(D(x)) \right| \\ &= \mathcal{O}\left(\frac{n\sqrt{\log n}\sqrt{\log \log n}}{k_{2n}^{3/2}} + \frac{\sqrt{n} \log^2 n}{k_{2n}}\right). \end{aligned}$$

THEOREM 2. *If  $F$  is continuous on  $\mathbb{R}$ , then*

$$(1.7) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} \right| = \begin{cases} \mathcal{O}_P\left(\frac{1}{\sqrt{k_n}}\right), \\ \mathcal{O}\left(\frac{\sqrt{\log n}}{\sqrt{k_{2n}}}\right), \end{cases}$$

$$(1.8) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \frac{\hat{F}_n(x) - F(x)}{1 - \hat{F}_n(x)} \right| = \mathcal{O}_P\left(\frac{1}{\sqrt{k_n}}\right),$$

and, for the respective constructions in Theorem 1,

$$(1.9) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \sqrt{n} \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - W_n(d(x)) \right| = \begin{cases} \mathcal{O}_P\left(\frac{\sqrt{n \log n}}{k_n}\right), \\ \mathcal{O}\left(\frac{\sqrt{n \log n}}{k_{2n}}\right), \end{cases}$$

$$(1.10) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \sqrt{n} \frac{\hat{F}_n(x) - F(x)}{1 - \hat{F}_n(x)} - W_n(d(x)) \right| = \mathcal{O}_P\left(\frac{\sqrt{n \log n}}{k_n}\right),$$

$$(1.11) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \sqrt{n} \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - \frac{W(d(x), n)}{\sqrt{n}} \right| = \mathcal{O}\left(\frac{\sqrt{n \log^2 n}}{k_{2n}}\right),$$

$$(1.12) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \frac{\sqrt{n} [\hat{F}_n(x) - F(x)]}{[1 - F(x)][1 + d_n(x)]} - B_n(D(x)) \right| = \mathcal{O}_P\left(\frac{n}{k_n^{3/2}} + \frac{\sqrt{n \log n}}{k_n}\right),$$

$$(1.13) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \frac{\sqrt{n} [\hat{F}_n(x) - F(x)]}{[1 - \hat{F}_n(x)][1 + d_n(x)]} - B_n(D(x)) \right| = \mathcal{O}_P\left(\frac{n}{k_n^{3/2}} + \frac{\sqrt{n \log n}}{k_n}\right),$$

$$(1.14) \quad \begin{aligned} & \sup_{x \leq Z_{n-k_n, n}} \left| \frac{\sqrt{n} [\hat{F}_n(x) - F(x)]}{[1 - F(x)][1 + d_n(x)]} - K_n(D(x)) \right| \\ &= \mathcal{O}\left(\frac{n \sqrt{\log n} \sqrt{\log \log n}}{k_{2n}^{3/2}} + \frac{\sqrt{n \log^2 n}}{k_{2n}}\right) \end{aligned}$$

and

$$(1.15) \quad \begin{aligned} & \sup_{x \leq Z_{n-k_n, n}} \left| \frac{\sqrt{n} [\hat{F}_n(x) - F(x)]}{[1 - \hat{F}_n(x)][1 + d_n(x)]} - K_n(D(x)) \right| \\ &= \mathcal{O}\left(\frac{n \sqrt{\log n} \sqrt{\log \log n}}{k_{2n}^{3/2}} + \frac{\sqrt{n \log^2 n}}{k_{2n}}\right). \end{aligned}$$

The almost sure rates in (1.1) and (1.7) improve those in Corollaries 1.2 and 1.3 of Stute (1994b); he does not have  $\mathcal{O}_P$ -bounds for these statements. All domains  $\sup_{x \leq Z_{n-k_n, n}}$  can be replaced by  $\sup_{x < Z_{n+1-k_n, n}}$  with some extra work. Since these versions have some value of curiosity only for (1.1), (1.7) and (1.8) when  $k_n \equiv 1$ , we preferred to keep the present easier unified proofs for  $\sup_{x \leq Z_{n-k_n, n}}$ . The corresponding special case of (1.1) for  $k_n \equiv 1$  and the  $\mathcal{O}_P(1)$ -bound also follows from Theorem 2.1 of Zhou (1991). We conjecture that, in general, the  $\mathcal{O}_P(1/\sqrt{k_n})$  convergence rates in (1.1), (1.7) and (1.8) are the universally best possible in the sense that they are unimprovable for some (bad) censorship situation. We also conjecture that the almost sure rates  $\mathcal{O}(\sqrt{\log \log n}/\sqrt{k_n})$  are universally possible and best in (1.1) and (1.7) for a nondecreasing  $\{k_n\}$  satisfying (\*). The latter would follow from Proposition 1 in Section 3 and the corresponding log log refinement of Stute's (1994b) result in Proposition 3 there. In fact, except for the respective extreme ranges  $1 \leq k_n \leq C \log^2 n$  and  $1 \leq k_n \leq [C \log^4 n]/\log \log n$  of small  $k_n$  left uncovered for any  $C > 0$ , we can prove the first conjecture and the universal optimality of the conjectured rate in the second by establishing the following

*CLAIM. Let  $\xi_n$  be the random variable on the left-hand side of any one of (1.1), (1.7) and (1.8). There is a censorship situation such that if  $k_n/\log^2 n \rightarrow \infty$ , then  $\sqrt{k_n} \xi_n$  does not converge to 0 in probability. If  $\xi_n$  is the random variable on the left-hand side of either (1.1) or (1.7), the sequence  $\{k_n\}$  is nondecreasing and  $[k_n \log \log n]/\log^4 n \rightarrow \infty$ , then in the same censorship situation  $\sqrt{k_n} \xi_n/\sqrt{\log \log n}$  does not converge to 0 almost surely.*

This is done at the end of Section 4 by constructing a simple concrete censorship model and applications of the respective Gaussian approximations in (1.3), (1.9), (1.10) and (1.4), (1.11). Using the lower bound from (4.3) below, for the large sequences  $k_n \equiv k_n(p) \equiv \lfloor pn \rfloor$ , where  $p \in (0, 1)$  is fixed, modulo the precise choice of the constant, the second conjecture would be asymptotically equivalent to the unimprovable bounded law of the iterated logarithm on  $(-\infty, T]$ ,  $T < \tau_H$  [cf. Csörgő and Horváth (1983)]. With regard to (1.7), note also the last major theoretical result by Stute and Wang (1993) [also exposed by Gill (1994)] concerning the almost sure behavior of  $\sup\{|\hat{F}_n(x) - F(x)|: x \leq \tau_H\}$ .

The result in (1.2) may be viewed as a far better version of Lemma 9.2 in Csörgő and Horváth (1982), which in turn improved Lemma 6.2 in Burke, Csörgő and Horváth (1981). The analogous conjectures corresponding to (1.2) are also made. In particular, if  $\mathcal{O}(n\sqrt{\log n}/k_n^{3/2})$  could be replaced by  $\mathcal{O}(n\sqrt{\log \log n}/k_n^{3/2})$  in (1.2), which would follow from Proposition 2 and the corresponding log log refinement in Proposition 4, then the factor  $\sqrt{\log n}\sqrt{\log \log n}$  would reduce to  $\log \log n$  in (1.6), (1.14), (1.15) and hence also in (2.11) below.

The results in (1.1) and (1.2) do not require the continuity of  $F$ . With the same proof, (1.3) and (1.4) also hold for possibly discontinuous  $F$  if we replace

$d(x)$  by

$$(1.16) \quad d_*(x) = \int_{-\infty}^x \frac{1}{[1 - H_-]^2} \frac{1 - F}{1 - F_-} d\tilde{H}$$

in them. Of course,  $d_*(\cdot) = d(\cdot)$  if  $F$  is continuous. The continuity of  $F$  is needed for (1.5) and (1.6) since otherwise  $d_n(\cdot)$  estimates the wrong function.

Typically, the sequence  $\{k_n\}$  will be nondecreasing, but it is not a real loss to replace  $k_{2n}$  by  $k_n$  in all the almost sure rate sequences for such a  $\{k_n\}$  since, besides having  $1 \leq k_{2n}/k_n$  for all  $n$ , we also have  $\limsup_{n \rightarrow \infty} k_{2n}/k_n \leq 2$  by condition (\*).

The basic approximations (1.3) and (1.9) contain unimprovable statements for the large sequences  $k_n \equiv k_n(p) \equiv \lfloor pn \rfloor$ , where  $p \in (0, 1)$  is fixed. For  $k_n(p)$ , neither (1.3) nor (1.9) holds for any sequence  $\{W_n\}$  of Wiener processes with the resulting rate  $\mathcal{O}((\log n)/\sqrt{n})$  changed to  $o((\log n)/\sqrt{n})$ . To see this for (1.3), if this improvement were possible, then, choosing any  $\beta > 1$  in the lower bound in (4.3) such that  $\beta p < 1$ , Proposition 5 would imply that  $\sup_{x \leq H^{-1}(1-\beta p)} |\lambda_n^*(x) - W_n(d(x))| = o((\log n)/\sqrt{n})$ . Since  $\lambda_n^*(\cdot)$  is a normalized sum of independent and identically distributed processes, it follows by the Erdős–Rényi strong law exactly as in Komlós, Major and Tusnády (1975b), for the uncensored empirical process, that this is impossible. Since  $\tilde{F}_n$  reduces to the ordinary sample distribution function in the absence of censoring (the case of a degenerate  $G$  is permissible throughout even when  $G$  degenerates at  $\infty$ , so that  $Y = \infty$ , giving the uncensored case), the corresponding statement for (1.9) follows directly from Komlós, Major and Tusnády (1975b). These facts do not, in principle, exclude improvability for smaller  $k_n$ , the real scope of this paper, but they indicate that any improvement, if there is, will be hard. Various further remarks concerning possible improvements are given in Section 5 following the proofs, particularly for (1.5), (1.6), (1.12)–(1.15), where at present the approximation rates are completely determined by that in (1.2) obtained for the variance estimator, for any sequence  $k_n \equiv \lfloor n^\alpha \rfloor$  for a fixed  $\alpha \in (0, 1)$ .

The approximation rate  $\mathcal{O}(n^{-1/2} \log n)$  corresponding to (1.3) and (1.9), and the rate  $\mathcal{O}(n^{-1/2} \log^2 n)$  corresponding to (1.4) and (1.11), for the fixed interval  $(-\infty, H^{-1}(1-p)]$ , which here results from  $k_n \equiv \lfloor pn \rfloor$  for some  $p \in (0, 1)$ , were first obtained by Burke, Csörgő and Horváth (1988) and, for (1.9) and (1.11) only, independently by Major and Rejtő (1988). They come from algebraic and exponential probability inequalities, respectively, both resulting in the same  $\mathcal{O}(n^{-1/2} \log n)$  convergence rates for the distributions of some functionals of the respective processes on  $(-\infty, H^{-1}(1-p)]$  and asymptotically exclude about  $\lfloor np \rfloor$  data points. The first Gaussian approximations in Burke, Csörgő and Horváth (1981, 1988) were in fact also the first to move out to  $\tau_H$  at some intermediate rate. Although this was not entertained at the time, the 1988 version would give, through (4.3) below,  $\mathcal{O}(n^{5/2}(\log n)/k_n^3)$  for (1.3) and for the unweighted version in (1.9), and  $\mathcal{O}(n^{7/2}(\log n)^2/k_n^4)$  for (1.4) and for the unweighted version in (1.11), which are very poor in comparison with the present results.

On the technical side, it is perhaps best to view the present paper as one close to being a possibly ultimate version of the approach of Breslow and Crowley (1974), with each ingredient taken to a limit achieved in the last two decades. Thus our basic framework is given by the theory of empirical processes, with many new developments since 1974. This is combined with Stute’s observation that discrete-time (which is the sample size) martingale methods are applicable to some degenerate  $U$ -statistics that arise in the main remainder terms. However, we also use Gill’s continuous-time martingale methods (for each sample size) which proved to be successful for problems very far out in the tail.

The next section exhibits some examples of applications of the theorems to weak convergence, the construction of confidence bands and rates of uniform consistency in the form of the law of the iterated logarithm. Section 3 contains some basic propositions that are important in the proofs and are of interest in their own right. This is followed by the proofs, and some supplementary remarks are given in the final section.

**2. Some applications.** The approximations in the two theorems may be useful even for very small  $k_n$  for which their rates blow up. One such application is the proof of the claim above; in other applications the presence of weights will pull down the stochastic order of the problem. Presently, we consider three circles of questions in subsections, in which  $k_n$  is chosen large enough to force the approximation rates to go to 0. Throughout,  $a(n)$  will denote a sequence of positive numbers such that  $1 \leq \lfloor \sqrt{n} a(n) \log n \rfloor < n - 1$  and  $1 \leq \lfloor n^{2/3} a(n) \rfloor < n - 1$  for all  $n$ ,  $a(n)/n^{1/3} \rightarrow 0$  and  $a(n) \rightarrow \infty$  arbitrarily slowly. Recalling the remark at (1.16), we assume in this section that  $F$  is continuous.

**2.1. Weak convergence.** Let  $\mathcal{D}[-\infty, \tau_H]$  be the nonseparable metric space of functions  $f$  defined on  $(-\infty, \tau_H)$  that are right-continuous, have left-hand limits,  $f(-\infty) = \lim_{x \downarrow -\infty} f(x) = 0$  and  $f(\tau_H) = f(\tau_H-)$  is finite, where for  $f_1, f_2 \in \mathcal{D}[-\infty, \tau_H]$  the distance is  $\sup_{-\infty \leq x \leq \tau_H} |f_1(x) - f_2(x)|$ , and let  $\rightarrow_{\mathcal{D}}$  denote convergence in distribution, as  $n \rightarrow \infty$ , with respect to the  $\sigma$ -algebra generated by the open balls for this metric. For  $f \in \mathcal{D}[-\infty, \tau_H]$  and  $T \in (-\infty, \tau_H)$ , let  $\llbracket f \rrbracket^T$  denote the function  $f(x \wedge T)$ ,  $x \in [-\infty, \tau_H]$ , stopped at  $T$ . Let  $W(\cdot)$  be a standard Wiener process and  $B(\cdot)$  be a Brownian bridge. Then (1.3), (1.5), (1.9), (1.10), (1.12), (1.13) and (4.3) below immediately imply the following three groups of theorems, the convergence within each group holding jointly:

$$(2.1) \quad \sqrt{n} \llbracket \Lambda_n(\cdot) - \Lambda(\cdot) \rrbracket^{Z_{n-\lfloor \sqrt{n} a(n) \log n \rfloor, n}} \rightarrow_{\mathcal{D}} W(d(\cdot)) \quad \text{in } \mathcal{D}[-\infty, \tau_H],$$

$$(2.2) \quad \sqrt{n} \left[ \left[ \frac{\hat{F}_n(\cdot) - F(\cdot)}{1 - F(\cdot)} \right] \right]^{Z_{n-\lfloor \sqrt{n} a(n) \log n \rfloor, n}} \rightarrow_{\mathcal{D}} W(d(\cdot)) \quad \text{in } \mathcal{D}[-\infty, \tau_H],$$



$$(2.3) \quad \sqrt{n} \left[ \frac{\hat{F}_n(\cdot) - F(\cdot)}{1 - \hat{F}_n(\cdot)} \right]^{Z_{n - \lfloor \sqrt{n} a(n) \log n \rfloor, n}} \rightarrow_{\mathcal{D}} W(d(\cdot)) \quad \text{in } \mathcal{D}[-\infty, \tau_H]$$

if and only if  $d(\tau_H) < \infty$ ;

$$(2.4) \quad \sqrt{n} \left[ \hat{F}_n(\cdot) - F(\cdot) \right]^{Z_{n - \lfloor \sqrt{n} a(n) \log n \rfloor, n}} \rightarrow_{\mathcal{D}} [1 - F(\cdot)] W(d(\cdot)) \quad \text{in } \mathcal{D}[-\infty, \tau_H]$$

if and only if  $v_{\tau_H} = \sup_{x < \tau_H} v(x) < \infty$  for  $v(x) = [1 - F(x)]^2 d(x)$ ; and, always,

$$(2.5) \quad \sqrt{n} \left[ \frac{\Lambda_n(\cdot) - \Lambda(\cdot)}{1 + d_n^{\leq}(\cdot)} \right]^{Z_{n - \lfloor n^{2/3} a(n) \rfloor, n}} \rightarrow_{\mathcal{D}} B(D(\cdot)) \quad \text{in } \mathcal{D}[-\infty, \tau_H],$$

$$(2.6) \quad \sqrt{n} \left[ \frac{\hat{F}_n(\cdot) - F(\cdot)}{[1 - F(\cdot)][1 + d_n^{\leq}(\cdot)]} \right]^{Z_{n - \lfloor n^{2/3} a(n) \rfloor, n}} \rightarrow_{\mathcal{D}} B(D(\cdot)) \quad \text{in } \mathcal{D}[-\infty, \tau_H],$$

$$(2.7) \quad \sqrt{n} \left[ \frac{\hat{F}_n(\cdot) - F(\cdot)}{[1 - \hat{F}_n(\cdot)][1 + d_n^{\leq}(\cdot)]} \right]^{Z_{n - \lfloor n^{2/3} a(n) \rfloor, n}} \rightarrow_{\mathcal{D}} B(D(\cdot)) \quad \text{in } \mathcal{D}[-\infty, \tau_H],$$

where  $d_n^{\leq}(\cdot)$  is either  $d_n(\cdot)$  or the Hall–Wellner (1980) estimator  $\hat{d}_n(\cdot)$  defined as

$$\begin{aligned} \hat{d}_n(x) &= \int_{-\infty}^x \frac{d\tilde{H}_n(y)}{[1 - H_n(y-)][1 - H_n(y)]} \\ &= n \sum_{j=1}^n \frac{\delta_{j,n} I\{Z_{j,n} \leq x\}}{(n-j+1)(n-j)} \geq d_n(x), \quad x < Z_{n,n}. \end{aligned}$$

The versions with  $\hat{d}_n$  follow from those with the original  $d_n$  because, letting  $\xi_n$  be any one of the left-hand sides of (1.1), (1.7) or (1.8) again, by the fourth statement of Lemma 3 below we have  $\sqrt{n} \xi_n \sup_{x \leq Z_{n-k_n, n}} [\hat{d}_n(x) - d_n(x)] = \mathcal{O}_P(n^{3/2}/k_n^{5/2})$ , which for the present  $k_n \equiv \lfloor n^{2/3} a(n) \rfloor$  is  $o_P(n^{-1/6})$ . Of course, if  $d(\tau_H) < \infty$ , then all ten statements hold jointly, and if  $v_{\tau_H} < \infty$ , then the last seven statements hold jointly.

Part (ii) of Gill's (1983) Theorem 1.2 says that, understanding the stopped process as its left-hand limit at that point, one can replace  $Z_{n - \lfloor n^{2/3} a(n) \rfloor, n}$  by  $Z_{n,n}$  in (2.7) with  $d_n^{\leq} = \hat{d}_n$ , and, as among other things Ying (1989) pointed out, one can also replace  $Z_{n - \lfloor \sqrt{n} a(n) \log n \rfloor, n}$  by  $Z_{n,n}$  in (2.4), provided that, for  $g(x) = \int_{-\infty}^x dF/[1 - G_-]$ ,  $x \leq \tau_H$ ,

$$(2.8) \quad g(\tau_H) = \int_{-\infty}^{\tau_H} \frac{dF}{1 - G_-} = \int_{-\infty}^{\tau_H} \frac{[1 - F]^2}{[1 - H_-]^2} d\tilde{H} < \infty.$$

Of course,  $v(x) \leq g(x) \leq d(x)$  for all  $x \leq \tau_H$ ; even  $g(\tau_H) < \infty$  implies that  $\lim_{x \uparrow \tau_H} v(x) = 0$ . In the uncensored case,  $d(\tau_H) = d(\tau_F) = \infty$ . If  $\tau_G > \tau_F = \tau_H$ , then  $d(\tau_H) = \infty$  and  $g(\tau_H) < \infty$ . If  $\tau_G = \tau_F$ , then  $d(\tau_H) = \infty$  and  $g(\tau_H)$  may or may not be finite, and if  $g(\tau_H) = \infty$ , then  $v_{\tau_H}$  may or may not be finite. If  $\tau_H = \tau_G < \tau_F$ , then  $d(\tau_H) < \infty$  and  $v_{\tau_H} < \infty$  are equivalent and both  $v_{\tau_H} < \infty$

and  $v_{\tau_H} = \infty$  are possible. In Remark 5.2 below we show that (2.5)–(2.7) may not hold when  $Z_{n-\lfloor n^{2/3}a(n)\rfloor, n}$  is replaced by  $Z_{n-\lfloor 6.597(\log n)/\log \log n\rfloor, n}$ .

2.2. *Laws of the iterated logarithm.* Consider the Finkelstein set

$$\mathcal{F} = \{f|f: [0, 1] \mapsto \mathbb{R}, f(0) = 0 = f(1), \int_0^1 [f'(s)]^2 ds \leq 1\}$$

of absolutely continuous functions  $f$  with derivatives  $f'$  and the Strassen-type set

$$\mathcal{L} = \{f|f: [0, \infty) \mapsto \mathbb{R}, f(0) = 0, \int_0^\infty [f'(s)]^2 ds \leq 1\}.$$

Changing  $Z_{n-\lfloor \sqrt{n}a(n)\log n\rfloor, n}$  to  $Z_{n-\lfloor \sqrt{n}a(n)\log^2 n\rfloor, n}$  and dividing by  $\sqrt{2 \log \log n}$ , denote by  $\zeta_n(\cdot)$  any one of the resulting two processes from (2.1) and (2.2). Also, changing  $Z_{n-\lfloor n^{2/3}a(n)\rfloor, n}$  to  $Z_{n-\lfloor n^{2/3}a(n)\sqrt{\log n}\sqrt{\log \log n}\rfloor, n}$  and dividing by  $\sqrt{2 \log \log n}$ , denote by  $\eta_n(\cdot)$  any one of the resulting six processes from (2.5)–(2.7). Let  $\hookrightarrow$  denote almost sure relative compactness in  $\mathcal{D}[-\infty, \tau_H]$  with a limit set on the right. Let  $\mathcal{L}$  be the set in (1.30) of Gu and Lai (1990). Then (1.4), (1.6), (1.11), (1.14), (1.15) and the corresponding log log laws for the approximating processes [e.g., in M. Csörgő and Révész (1981)] immediately imply via (4.3) the following three groups of theorems:  $\{\zeta_n(\cdot)\} \hookrightarrow \{\ell: \ell = f(d), f \in \mathcal{L}\}$  if and only if  $d(\tau_H) < \infty$ , in which case

$$(2.9) \quad \limsup_{n \rightarrow \infty} \sup_{x \leq Z_{n-\lfloor \sqrt{n}a(n)\log^2 n\rfloor, n}} |\zeta_n(x)| = \sqrt{d(\tau_H)};$$

$\{\sqrt{n} \|\hat{F}_n(\cdot) - F(\cdot)\|^{Z_{n-\lfloor \sqrt{n}a(n)\log^2 n\rfloor, n}} / \sqrt{2 \log \log n}\} \hookrightarrow \mathcal{L}$  if and only if  $v_{\tau_H} < \infty$ , in which case

$$(2.10) \quad \limsup_{n \rightarrow \infty} \sup_{x \leq Z_{n-\lfloor \sqrt{n}a(n)\log^2 n\rfloor, n}} \sqrt{n} |\hat{F}_n(x) - F(x)| / \sqrt{2 \log \log n} = \sqrt{v_{\tau_H}};$$

and, always,  $\{\eta_n(\cdot)\} \hookrightarrow \{\ell: \ell = f(D), f \in \mathcal{F}\}$ , so that

$$(2.11) \quad \limsup_{n \rightarrow \infty} \sup_{x \leq Z_{n-\lfloor n^{2/3}a(n)\sqrt{\log n}\sqrt{\log \log n}\rfloor, n}} |\eta_n(x)| = \sup_{x < \tau_H} \sqrt{D(x)[1 - D(x)]}.$$

Gu and Lai (1990), through a strong approximation of the type in Proposition 5 below, obtained the sufficiency half of the two statements at (2.10) with the better  $Z_{n-\lfloor cn^\delta\rfloor, n}$ , where  $c > 0$  and  $1/3 < \delta < 1/2$  are constants. However, this is achieved under two sets of conditions on the censoring, one of which is (2.8) if  $F(\tau_H) < 1$ . A related nice result is by Gu (1991), up to  $Z_{n, n}$  under a more restrictive condition.

2.3. *Confidence bands.* The literature for bands on  $(-\infty, T]$ , for some  $T < \tau_H$ , is sizable; the reader is referred to the systematic analyses by Nair (1984) and Csörgő and Horváth (1986). Also, we mention here bands for  $F$  only, the corresponding statements for  $\Lambda$  being analogous and simpler. Let  $\alpha \in (0, 1)$  be fixed.

For the two choices of  $d_n^\triangleleft$  as  $d_n$  or  $\hat{d}_n$ , let  $(r_n^\triangleleft)^2 = d_n^\triangleleft(Z_{n-\lfloor n^{2/3}a(n)\rfloor}, n)$ . If  $w_\alpha > 0$  is such that  $P\{\sup_{0 \leq t \leq 1} |W(t)| > w_\alpha\} = \alpha$ , then, extending the Aalen–Nair bands, (2.2) and (2.3), or the original approximations, (1.2) and (4.2) below, yield

$$P\left\{\frac{\hat{F}_n(x) - w_\alpha n^{-1/2} r_n^\triangleleft}{1 - w_\alpha n^{-1/2} r_n^\triangleleft} \leq F(x) \leq \frac{\hat{F}_n(x) + w_\alpha n^{-1/2} r_n^\triangleleft}{1 + w_\alpha n^{-1/2} r_n^\triangleleft}, \right. \\ \left. x \leq Z_{n-\lfloor n^{2/3}a(n)\rfloor}, n\right\} \rightarrow 1 - \alpha$$

and, writing  $[a \mp b] = [a - b, a + b]$  for  $a, b \geq 0$ ,

$$P\{F(x) \in [\hat{F}_n(x) \mp w_\alpha n^{-1/2} r_n^\triangleleft \{1 - \hat{F}_n(x)\}], x \leq Z_{n-\lfloor n^{2/3}a(n)\rfloor}, n\} \rightarrow 1 - \alpha$$

as  $n \rightarrow \infty$ , provided  $d(\tau_H) < \infty$ . Since the respective upper and lower contours of the first and second versions are uniformly better for all  $n$  large enough and the source statements hold jointly, one may want to entertain the corresponding *mixed* band [cf. Csörgő and Horváth (1986)]. However, the condition  $d(\tau_H) < \infty$  is necessary. It is violated in many common situations, even with light censoring, and is difficult to check.

Although in general asymptotically conservative, the extended Hall–Wellner bands do not share this drawback: if  $y_\alpha > 0$  is such that  $P\{\sup_{0 \leq t \leq 1} |B(t)| > y_\alpha\} = \alpha$ , then, by (2.6) and (2.7), (1.2) and (4.2), in any censorship situation we have

$$P\left\{\frac{\hat{F}_n(x) - y_\alpha h_n^\triangleleft(x)}{1 - y_\alpha h_n^\triangleleft(x)} \leq F(x) \leq \frac{\hat{F}_n(x) + y_\alpha h_n^\triangleleft(x)}{1 + y_\alpha h_n^\triangleleft(x)}, \right. \\ \left. x \leq Z_{n-\lfloor n^{2/3}a(n)\rfloor}, n\right\} \rightarrow p_\alpha(\tau_H),$$

where temporarily  $h_n^\triangleleft(x) = n^{-1/2}[1 + d_n^\triangleleft(x)]$ , and

$$P\{F(x) \in [\hat{F}_n(x) \mp y_\alpha n^{-1/2} \{1 + d_n^\triangleleft(x)\} \{1 - \hat{F}_n(x)\}], \\ x \leq Z_{n-\lfloor n^{2/3}a(n)\rfloor}, n\} \rightarrow p_\alpha(\tau_H),$$

where  $p_\alpha(\tau_H) = P\{\sup_{0 \leq t \leq D(\tau_H)} |B(t)| \leq y_\alpha\} \geq 1 - \alpha$ . Here  $D(\tau_H) = 1$ , or equivalently,  $p_\alpha(\tau_H) = 1 - \alpha$ , if and only if  $d(\tau_H) = \infty$ . The original Hall–Wellner (1980) band is the second version with the choice  $d_n^\triangleleft(\cdot) = \hat{d}_n(\cdot)$ . This band, which by Gill’s theorem holds up to  $Z_{n,n}$  if (2.8) is satisfied, has the attractive feature that it reduces to the common Kolmogorov band in the absence of censorship. This is because, then,  $\{1 + \hat{d}_n(x)\} \{1 - \hat{F}_n(x)\} \equiv 1$ , which is nothing but the identity (4.8) below. However, using  $d_n$  instead of  $\hat{d}_n$  results

in uniformly narrower bands in both versions. Choosing the contours from all versions freely, there are in fact sixteen asymptotically equivalent bands here, and by arguments in Csörgő and Horváth (1986) it is also permitted to fabricate the narrowest possible from them. In the absence of censorship, a variant reduces to a uniform improvement over the Kolmogorov band in that paper. Letting again  $r_n^2 = d_n(Z_{n-\lfloor n^{2/3}a(n) \rfloor, n})$  and  $t_n = r_n^2/[1 + r_n^2]$ , which by (1.2) is a consistent estimator of  $D(\tau_H)$ , for a finite-sample correction of conservatism one may want to use in practice the number  $y_\alpha(n)$  in place of  $y_\alpha$ , calculated from  $P\{\sup_{0 \leq t \leq t_n} |B(t)| \leq y_\alpha(n)\} = 1 - \alpha$ . See the table, references and the discussion in Hall and Wellner (1980). Again, the bands in the present paragraph may not hold up to  $Z_{n-\lfloor 6.597(\log n)/\log \log n \rfloor, n}$ .

It is, of course, impossible, on the basis of asymptotics, to give precise practical advice valid in all situations. If one takes the natural basis of the logarithm literally, for example, then the first  $n$  for which the “large” sequence  $n^{2/3}$  overtakes the “small” sequence  $\sqrt{n} \log n$  is  $n = 24, 128, 092$ . Common sense suggests that estimating (a possibly infinite) variance can only make things worse, so that  $n^{2/3}$  should indeed be larger than (a constant multiple of)  $\sqrt{n} \log n$ . Thus it is tempting to replace  $\log n$  and  $a(n)$  by 1 and, on the basis of the weak convergence results, propose the heuristic rule of thumb: to use uniform Gaussian approximations safely, delete about  $\lfloor n^{2/3} \rfloor$  or  $\lfloor \sqrt{n} \rfloor$  top observations from a sample of  $n$ , depending on whether a variance is estimated or not. Extended simulation studies would be needed to substantiate this. Further discussion is given in Remarks 5.1 and 5.2.

For an ultimate understanding of the asymptotic behavior of the basic processes, one can envisage results in which those of the present paper, perhaps with some improvements, up to  $Z_{n-k_n, n}$  and with no condition on censorship, would represent one extreme special case, while those up to  $Z_{n, n}$  under Gill’s condition (2.8) on the lightness of censoring, or perhaps only under  $v_{\tau_H} < \infty$ , would be the other extreme special case. Assuming for the sake of argument that (2.7) is optimal, one would like to have the result up to  $Z_{n-k_n^{(\alpha)}, n}$  under the condition  $\int_{-\infty}^{\tau_H} [1 - G_-]^{-\alpha} dF < \infty$ , or perhaps only under the weaker condition  $\sup_{x < \tau_H} [1 - F(x)]^{2\alpha} \int_{-\infty}^{\tau_H} [1 - H_-]^{-2\alpha} d\tilde{H} < \infty$ , for some  $\alpha \in [0, 1]$ , such that  $k_n^{(0)} \equiv \lfloor n^{2/3}a(n) \rfloor$  and  $k_n^{(1)} \equiv 0$ . What is  $k_n^{(\alpha)}$ ?

**3. Basic propositions.** With the notation of the previous sections, for  $x < \tau_H$ , set

$$\Lambda_n^*(x) = \int_{-\infty}^x \frac{d\tilde{H}_n(y)}{1 - H_-(y)} + \int_{-\infty}^x \frac{H_n(y-) - H_-(y)}{[1 - H_-(y)]^2} d\tilde{H}(y)$$

and

$$\begin{aligned} d_n^*(x) &= d_n^\diamond(x) + 2d_n^{\diamond\diamond}(x) \\ &= \int_{-\infty}^x \frac{d\tilde{H}_n(y)}{[1 - H_-(y)]^2} + 2 \int_{-\infty}^x \frac{H_n(y-) - H_-(y)}{[1 - H_-(y)]^3} d\tilde{H}(y). \end{aligned}$$

The almost sure statement in Proposition 1 is a refinement of Stute's (1994b) Theorem 1.1. Throughout,  $\{c_n\}$  denotes any nondecreasing sequence of positive numbers such that  $\sum_{k=1}^{\infty} (k c_{2^k})^{-1} < \infty$ . The natural examples for the "smallest" such sequences  $\{c_n\}$ , for all  $n$  large enough, are  $c_n = (\log \log n)^{1+\varepsilon}$ ,  $c_n = (\log \log n)(\log \log \log n)^{1+\varepsilon}$ ,  $c_n = (\log \log n)(\log \log \log n)(\log \log \log \log n)^{1+\varepsilon}$  and so on, where  $\varepsilon > 0$  is as small as we wish. The  $\mathcal{O}_P$ -bound is Theorem 1.5 in Stute (1994b); a short proof for this is also given. The details for Proposition 1 help a shorter rendering of the proof of Proposition 2.

PROPOSITION 1. *We have*

$$\sup_{x \leq Z_{n-k_n, n}} |\Lambda_n(x) - \Lambda_n^*(x)| = \begin{cases} \mathcal{O}_P\left(\frac{1}{k_n}\right) & [\text{Stute(1994b)}], \\ o\left(\frac{\sqrt{c_n \log n}}{k_{2n}}\right). \end{cases}$$

PROPOSITION 2. *We have*

$$\sup_{x \leq Z_{n-k_n, n}} |d_n(x) - d_n^*(x)| = \begin{cases} \mathcal{O}_P\left(\frac{n}{k_n^2}\right), \\ o\left(\frac{n\sqrt{c_n \log n}}{k_{2n}^2}\right). \end{cases}$$

Even though the almost sure statements will follow by purely empirical-process methodology, to appreciate the structure of the next two propositions relative to each other, we right away introduce ( $n^{-1/2}$  times) the basic martingale [Gill (1980); for a fine exposition see Shorack and Wellner (1986), Chapter 7, where for Section 3 see Wang (1987) and finally Stute and Wang (1993)]. Since transformations of the original lifetime data, which may result in negative  $X_1, X_2, \dots$ , are sometimes useful, we do not restrict considerations to  $F$  and  $G$  supported on  $[0, \infty)$ . Accordingly, we use the martingale approach in its form extended to the whole line with starting point  $-\infty$ , as in Gu and Lai (1990). For a given  $x \in \mathbb{R}$ , let  $\mathcal{F}_x$  be the complete  $\sigma$ -algebra generated by  $\{I\{Z_k \leq x\}, \delta_k I\{Z_k \leq x\}, Z_k I\{Z_k \leq x\}, k = 1, 2, \dots\}$ . Set  $M_n(x) = \tilde{H}_n(x) - \int_{-\infty}^x [1 - H_n(y-)] d\Lambda(y)$ ,  $x \in \mathbb{R}$ . Then  $\{(M_n(x), \mathcal{F}_x): x \in [-\infty, \tau_H]\}$  is a square-integrable martingale, for each  $n$ , with predictable variation process

$$\begin{aligned} \langle M_n \rangle(x) &= \frac{1}{n} \int_{-\infty}^x [1 - H_n(y-)] \frac{1 - F(y)}{1 - F_-(y)} d\Lambda(y) \\ &= \frac{1}{n} \int_{-\infty}^x \frac{1 - H_n(y-)}{1 - H_-(y)} \frac{1 - F(y)}{1 - F_-(y)} d\tilde{H}(y), \end{aligned}$$

where, of course,  $0 \leq [1 - F(y)]/[1 - F_-(y)] \leq 1$ ,  $y < \tau_H$ . Simple calculations show that

$$L_n(x) := \Lambda_n^*(x) - \Lambda(x) = \int_{-\infty}^x \frac{dM_n}{1 - H_-}$$

and

$$D_n^\diamond(x) := d_n^*(x) - d_n^{\diamond\diamond}(x) - d(x) = d_n^\diamond(x) + d_n^{\diamond\diamond}(x) - d(x) = \int_{-\infty}^x \frac{dM_n}{[1 - H_-]^2}$$

for all  $x < \tau_H$ . Both  $(L_n(\cdot), \mathcal{F})$  and  $(D_n^\diamond(\cdot), \mathcal{F})$  are square-integrable martingales again [cf. Shorack and Wellner (1986), pages 890–891] for every  $n$  over  $[-\infty, \tau_H)$ , and their respective predictable variation processes for  $x < \tau_H$  are given by

$$\langle L_n \rangle(x) = \int_{-\infty}^x \frac{d\langle M_n \rangle}{[1 - H_-]^2} = \frac{1}{n} \int_{-\infty}^x \frac{1 - H_n(y-)}{[1 - H_-(y)]^3} \frac{1 - F(y)}{1 - F_-(y)} d\tilde{H}(y)$$

and

$$\langle D_n^\diamond \rangle(x) = \int_{-\infty}^x \frac{d\langle M_n \rangle}{[1 - H_-]^4} = \frac{1}{n} \int_{-\infty}^x \frac{1 - H_n(y-)}{[1 - H_-(y)]^5} \frac{1 - F(y)}{1 - F_-(y)} d\tilde{H}(y).$$

Although the final conclusions in the next two propositions are analogous, as are the previous two, the one in Proposition 4 is arrived at along a more circuitous course, the details of which facilitate an easier rendering of the proof and better understanding. The almost sure statement in Proposition 3 is Lemma 2.8 in Stute (1994b).

PROPOSITION 3. *We have*

$$\sup_{x \leq Z_{n-k_n, n}} |L_n(x)| = \sup_{x \leq Z_{n-k_n, n}} |\Lambda_n^*(x) - \Lambda(x)| = \begin{cases} \mathcal{O}_P\left(\frac{1}{\sqrt{k_n}}\right), \\ \mathcal{O}\left(\frac{\sqrt{\log n}}{\sqrt{k_n}}\right) \end{cases} \quad [\text{Stute}(1994b)].$$

PROPOSITION 4. *We have*

$$\sup_{x \leq Z_{n-k_n, n}} |d_n^\diamond(x) - d(x)| = \mathcal{O}\left(\frac{n\sqrt{\log n}}{k_n^{3/2}}\right),$$

$$\sup_{x \leq Z_{n-k_n, n}} |d_n^{\diamond\diamond}(x)| = \begin{cases} \mathcal{O}_P\left(\frac{n}{k_n^{3/2}}\right), \\ \mathcal{O}\left(\frac{n\sqrt{\log \log n}}{k_n^{3/2}}\right) \end{cases}$$

and

$$\sup_{x \leq Z_{n-k_n, n}} |D_n^\diamond(x)| = \sup_{x \leq Z_{n-k_n, n}} |d_n^\diamond(x) + d_n^{\diamond\diamond}(x) - d(x)| = \mathcal{O}_P\left(\frac{n}{k_n^{3/2}}\right),$$

so that

$$\sup_{x \leq Z_{n-k_n, n}} |d_n^*(x) - d(x)| = \begin{cases} \mathcal{O}_P\left(\frac{n}{k_n^{3/2}}\right), \\ \mathcal{O}\left(\frac{n\sqrt{\log n}}{k_n^{3/2}}\right). \end{cases}$$

To have a counterpart of Proposition 1 for  $\hat{F}_n$  as well, interesting in its own right, we formulate one more proposition. [Such a representation is implicit in Burke, Csörgő and Horváth (1981) for  $\Lambda_n$ ; in fact, on some increasing intervals, the first explicit form for  $\hat{F}_n$  on  $(-\infty, T]$ ,  $T < \tau_H$ , was given by Lo and Singh (1985).] It is not used for the proof of Theorem 2 and will be verified after that proof from the same source. Introduce the normalized cumulative hazard process  $\lambda_n(x) = \sqrt{n}[\Lambda_n(x) - \Lambda(x)]$  and its approximating process  $\lambda_n^*(x) = \sqrt{n}L_n(x) = \sqrt{n}[\Lambda_n^*(x) - \Lambda(x)]$ ,  $x < \tau_H$ . The almost sure half of the second statement improves Stute's (1994b) Theorem 1.4 in two respects. In the second and third assertions, due to the continuity of  $F$ , the function  $d_*(\cdot)$  of (1.16) reduces to  $d(\cdot)$ .

PROPOSITION 5. *We have*

$$\sup_{x \leq Z_{n-k_n, n}} |\lambda_n(x) - \lambda_n^*(x)| = \begin{cases} \mathcal{O}_P\left(\frac{\sqrt{n}}{k_n}\right), \\ \mathcal{O}\left(\frac{\sqrt{nc_n \log n}}{k_{2n}}\right), \end{cases}$$

and if  $F$  is continuous on  $\mathbb{R}$ , then also

$$\sup_{x \leq Z_{n-k_n, n}} \left| \sqrt{n} \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - \lambda_n^*(x) \right| = \begin{cases} \mathcal{O}_P\left(\frac{\sqrt{n}}{k_n}\right), \\ \mathcal{O}\left(\frac{\sqrt{n \log n}}{k_{2n}}\right) \end{cases}$$

and

$$\sup_{x \leq Z_{n-k_n, n}} \left| \sqrt{n} \frac{\hat{F}_n(x) - F(x)}{1 - \hat{F}_n(x)} - \lambda_n^*(x) \right| = \mathcal{O}_P\left(\frac{\sqrt{n}}{k_n}\right),$$

where  $\lambda_n^*(x) = n^{-1/2} \sum_{j=1}^n l_j(x)$  is a normalized sum of independent and identically distributed processes  $l_1(x), l_2(x), \dots$  with  $E(l_j(x)) = 0$  and  $E(l_j^2(x)) = d_*(x)$ , given by

$$l_j(x) = \frac{\delta_j I\{Z_j \leq x\} - \tilde{H}(x)}{1 - H_-(x)} - \int_{-\infty}^x \frac{\delta_j I\{Z_j \leq y\} - \tilde{H}(y)}{[1 - H_-(y)]^2} dH_-(y) \\ + \int_{-\infty}^x \frac{I\{Z_j < y\} - H_-(y)}{[1 - H_-(y)]^2} d\tilde{H}(y),$$

such that  $\{(l_j(x), \mathcal{F}_x): x \in [-\infty, \tau_H]\}$  is a martingale,  $j = 1, 2, \dots$

**4. Proofs.** Lemma 1 below gathers some facts from the theory of empirical processes needed for the proof of the propositions, whereas Lemmas 2 and 3 are required for the proofs of the theorems. The almost sure statements in Lemma 2 are the exceptions from our notational convention; they hold for all  $1 \leq k_n \leq n$ . Introduce the notation

$$\begin{aligned}
 J_n(p; k_n) &= \sup_{x \leq H^{-1}(1-p(k_n/n))} \frac{1 - H_n(x-)}{1 - H_-(x)}, \\
 E_n(k_n) &= \sup_{x \leq H^{-1}(1-(k_n/7n))} \frac{|H_n(x-) - H_-(x)|}{\sqrt{1 - H_-(x)}}, \\
 A_n^{(\gamma)}(p) &= A_n^{(\gamma)}(p; k_n) \\
 &= \sup_{x \leq H^{-1}(1-p(k_n/n))} \frac{|H_n(x-) - H_-(x)|}{[1 - H_-(x)]^{1-\gamma}} \\
 &\leq A_n^{(0)}(p; k_n) = A_n^{(0)}(p),
 \end{aligned}$$

where  $0 < p < 1$  and  $0 \leq \gamma < 1/2$ , and put  $J_n(p) = J_n(p; k_n)$ ,  $J_n = J_n(k_n) = J_n(1/7; k_n)$ ,  $E_n = E_n(k_n)$  and  $A_n(p) = A_n(p; k_n) = A_n^{(0)}(p; k_n)$  for later use.

LEMMA 1 [Wellner (1978), Mason (1985) and Csáki (1977)]. *For any fixed  $p \in (0, 1)$  and  $\gamma \in [0, 1/2)$ , we have  $J_n(p; k_n) = \mathcal{O}_P(1)$ ,  $\limsup_{n \rightarrow \infty} J_n(p; k_n) \leq 1$ ,*

$$A_n^{(\gamma)}(p; k_n) = \mathcal{O}_P\left(\frac{1}{\sqrt{n}} \left(\frac{n}{k_n}\right)^{(1-2\gamma)/2}\right)$$

and

$$\limsup_{n \rightarrow \infty} \sqrt{\frac{n}{\log \log n}} E_n \leq 2.$$

LEMMA 2. *For any integers  $1 \leq k_n \leq n$ , any fixed  $p \in (0, 1)$ , any standard Wiener processes  $\{W_n(\cdot)\}$ , any standard Wiener sheet  $W(\cdot, \cdot)$  and Kiefer sheet  $K(\cdot, \cdot)$ , we have*

$$\begin{aligned}
 W_n^*(p) = W_n^*(p; k_n) &= \sup_{x \leq H^{-1}(1-p(k_n/n))} |W_n(d(x))| = \mathcal{O}_P\left(\frac{\sqrt{n}}{\sqrt{k_n}}\right), \\
 W_n^\diamond = W_n^\diamond(k_n) &= \sup_{x \leq H^{-1}(1-(k_n/7n))} \frac{|W(d(x), n)|}{\sqrt{n}} = \mathcal{O}\left(\frac{\sqrt{n \log \log n}}{\sqrt{k_n}}\right)
 \end{aligned}$$

and

$$K_n^\diamond = K_n^\diamond(k_n) = \sup_{x \leq H^{-1}(1-(k_n/7n))} \frac{|K(D(x), n)|}{\sqrt{n}} = \mathcal{O}(\sqrt{\log \log n}).$$



LEMMA 3 [Gill (1980, 1983) for arbitrary  $F$ ]. *Suppose that  $F$  is continuous. Then  $\sup_{x \leq Z_{n-1, n}} |[1 - F(x)]/[1 - \hat{F}_n(x)]| = \mathcal{O}_p(1)$  and  $[1 - \hat{F}_n(x)][1 + \hat{d}_n(x)] \geq 1$  for all  $x \in (-\infty, Z_{n-1, n}]$ . Furthermore,  $\sup_{x \leq Z_{n-k_n, n}} [1 + \hat{d}_n(x)]/[1 + d_n(x)] \leq 1 + k_n^{-1}$  and  $\sup_{x \leq Z_{n-k_n, n}} [\hat{d}_n(x) - d_n(x)] \leq n/k_n^2$  for all integers  $1 \leq k_n < n$ .*

We state a few more useful facts before the proofs. Note that  $J_-(J^{-1}(s)) \leq s$  for all  $s \in (0, 1)$  for any distribution function  $J$ , and Lemma 2.1 of Stute (1994b) states

$$(4.1) \quad \int_{-\infty}^{J^{-1}(s)} \frac{dJ(x)}{[1 - J_-(x)]^\beta} \leq \frac{\beta}{(\beta - 1)(1 - s)^{\beta-1}}, \quad 0 < s < 1, \text{ for all } \beta > 1.$$

Next, considering the functions  $\psi(x) = 2[(1+x)\log(1+x) - x]/x^2$  and  $\psi^*(x) = x\psi(x)$ ,  $x > 0$ , set  $C_p = (1-p)\psi^*((1-p)/p)/2$  for  $p \in (0, 1)$  and  $C_\beta^* = (\beta - 1)\psi^*((\beta - 1)/\beta)/2$  for  $\beta > 1$ . It is important that  $\lim_{p \downarrow 0} C_p = \infty = \lim_{\beta \uparrow \infty} C_\beta^*$ . Corresponding to Lemmas 2.4 and 2.3 in Stute (1994b) [and Lemma 7.1 in Csörgő and Horváth (1982)], we claim, for all  $1 \leq k_n < n$ ,  $0 < p < 1$  and  $\beta > 1$ ,

$$(4.2) \quad \begin{aligned} P\left\{Z_{n-k_n, n} > H^{-1}\left(1 - p \frac{k_n}{n}\right)\right\} &< e^{-C_p}, \\ P\left\{Z_{n-k_n, n} < H^{-1}\left(1 - \beta \frac{k_n}{n}\right)\right\} &< e^{-C_\beta^*}, \end{aligned}$$

where  $H^{-1}(s) = -\infty$  for  $s < 0$ , and, with an almost surely finite random integer  $n_0$ , if  $k_n \geq \log n$  for all  $n$  large enough, then, for all  $n \geq n_0$ ,

$$(4.3) \quad H^{-1}\left(1 - \beta \frac{k_n}{n}\right) \leq Z_{n-k_n, n} = H_n^{-1}\left(1 - \frac{k_n}{n}\right) \leq H^{-1}\left(1 - \frac{k_n}{7n}\right),$$

the lower bound holding for any  $\beta > 1$  for which  $\sum_{n=1}^{\infty} e^{-C_\beta^* k_n} < \infty$ . The choice  $\beta = 5$  suffices for all  $\{k_n\}$  in (4.3). (No eventually meaningful lower bound of this type is possible and needed for the uninteresting large sequences  $k_n$  for which  $k_n/n \rightarrow 1$ .)

Indeed, letting  $U_1, U_2, \dots$  be independent variables uniformly distributed on  $[0, 1]$ , denoting by  $\Gamma_n(s)$ ,  $0 \leq s \leq 1$ , the sample distribution function of the first  $n$  of these and using Bennett's one-sided inequality [Sharack and Wellner (1986), pages 851–853, where an exponent 2 is missing in the statement, or Pollard (1984), pages 192–194] along with the fact that  $\psi^*$  is increasing on  $(0, \infty)$ , we see that

$$P\left\{\frac{Z_{n-k_n, n}}{H^{-1}(1 - p(k_n/n))} > 1\right\} \leq P\left\{\Gamma_n\left(p \frac{k_n}{n}\right) > \frac{k_n}{n}\right\} < e^{-C_p k_n} = n^{-C_p k_n / \log n}$$

for all  $1 \leq k_n < n$ . Now the upper bound in (4.2) follows upon replacing  $k_n$  by 1, whereas that in (4.3) follows by the Borel–Cantelli lemma upon noting that  $C_{1/7} > 1.08$ . For the lower bounds in both, one has for every  $\beta > 1$ , again by Bennett’s inequality,

$$P\left\{\frac{Z_{n-k_n, n}}{H^{-1}(1 - \beta(k_n/n))} < 1\right\} \leq P\left\{\Gamma_n\left(\beta \frac{k_n}{n}\right) < \frac{k_n}{n}\right\} < e^{-C_\beta^* k_n} = n^{-C_\beta^* k_n / \log n},$$

where  $C_5^* > 1.29$ .

PROOF OF LEMMA 1. Supposing, for the purposes of the present lemma, that  $Z_1 = H^{-1}(U_1), Z_2 = H^{-1}(U_2), \dots$  for the  $U_1, U_2, \dots$  as above, we have

$$J_n(p) \leq \sup_{0 \leq s \leq 1-p(k_n/n)} \frac{1 - \Gamma_n(s)}{1 - s},$$

$$A_n^{(\gamma)}(p) \leq \sup_{0 \leq s \leq 1-p(k_n/n)} \frac{|\Gamma_n(s) - s|}{[1 - s]^{1-\gamma}} \leq \sup_{0 \leq s \leq 1-p(k_n/n)} \frac{|\Gamma_n(s) - s|}{s^\gamma [1 - s]^{1-\gamma}}$$

in the simplified notation, and, for all  $n$  large enough under (\*),

$$E_n \leq \sup_{0 \leq s \leq 1-(k_n/7n)} \frac{|\Gamma_n(s) - s|}{\sqrt{1 - s}}$$

$$\leq \max \left\{ \sup_{0 \leq s \leq 1} \frac{|\Gamma_n(s) - s|}{\sqrt{6/7}}, \sup_{\log \log n/n \leq s \leq 1-(\log \log n/n)} \frac{|\Gamma_n(s) - s|}{\sqrt{s(1 - s)}} \right\}.$$

The statements for  $J_n(p)$  thus follow from Lemma 1 and Theorem 1 of Wellner (1978) [Shorack and Wellner (1986), pages 415 and 424], respectively. Since  $pk_n/n \leq p$ , when  $k_n \rightarrow \infty$  the statement for  $A_n^{(\gamma)}(p)$  follows directly from the fourth Rényi-type limit theorem in Theorem 1 of Mason (1985), being an extension of the left-sided version of Theorem 4.5.1 in M. Csörgő, S. Csörgő, Horváth and Mason (1986) for the case  $\gamma = 0$ . If  $\{k_n\}$  is bounded, then in fact the better result  $A_n(p) = A_n^{(0)}(p) = \mathcal{O}_P(1/\sqrt{n})$  follows from the simple proof of Theorem 4.5.1 in M. Csörgő, S. Csörgő, Horváth and Mason (1986), that is, from the weighted approximation results there. (This better form for a bounded  $\{k_n\}$  is not significant in the applications of the lemma below.) The statement for  $E_n$  follows by the ordinary Smirnov–Chung log log law and by Theorem 3.2 of Csáki (1977) [Shorack and Wellner (1986), pages 504 and 609, respectively]. □

PROOF OF PROPOSITION 1. For every  $x \leq Z_{n, n}$  Stute (1994b) obtains

$$(4.4) \quad \Lambda_n(x) - \Lambda_n^*(x) = -\frac{R_n(x)}{n^2} - R_{n1}(x) + R_{n2}(x) + R_{n3}(x) - R_{n4}(x),$$

where

$$\begin{aligned} R_{n1}(x) &= \frac{1}{n} \int_{-\infty}^x \frac{d\tilde{H}_n(y)}{[1 - H_-(y)]^2}, \\ R_{n2}(x) &= \int_{-\infty}^x \frac{[H_n(y-) - H_-(y)]^2}{[1 - H_-(y)]^2 [1 - H_n(y-)]} d\tilde{H}_n(y), \\ R_{n3}(x) &= \frac{1}{n} \int_{-\infty}^x \frac{d\tilde{H}_n(y)}{1 - H_-(y)}, \quad R_{n4}(x) = \frac{1}{n} \int_{-\infty}^x \frac{H_n(y-) - H_-(y)}{[1 - H_-(y)]^2} d\tilde{H}_n(y), \end{aligned}$$

and  $R_n(x)$  is a  $U$ -statistic minus its Hájek projection, so itself a degenerate  $U$ -statistic for which, for all positive integers  $l$  and  $n$  and for all  $p \in (0, 1)$ ,

$$(4.5) \quad E\left(\sup_{x \leq T_l(p)} |R_n(x)|^2\right) \leq \frac{C_1}{p^2} n^2 \frac{l^2}{k_l^2}, \quad \text{where } T_l(p) = H^{-1}\left(1 - p \frac{k_l}{l}\right),$$

for a universal constant  $C_1 > 0$ . Bringing  $k_{2n}$  in for the first time, this implies

$$(4.6) \quad \sup_{x \leq Z_{n-k_n, n}} |R_n(x)/n^2| = \begin{cases} \mathcal{O}_P\left(\frac{1}{k_n}\right), \\ o\left(\frac{\sqrt{c_n \log n}}{k_{2n}}\right). \end{cases}$$

The bound in probability is immediate by taking  $l = n$  in (4.5) and using (4.2) and Chebyshev's inequality. Putting  $T_l = T_l(1/7)$  in (4.5) and using (4.3), the almost sure rate follows by Stute's (1994b) proof of his Lemma 2.5 [cf. also Serfling (1980), page 189]. This is based on the fact that  $\{\sup_{x \leq T_{2^{j+1}}} |R_n(x)|/[n(n-1)]: 2^j \leq n \leq 2^{j+1}\}$  is a reverse submartingale for each  $j = 1, 2, \dots$ , so Doob's inequality can be used in each block of the union  $\cup_{j=2}^{\infty} \{n: 2^j \leq n < 2^{j+1}\} = \{4, 5, \dots\}$ , followed by an application of (4.5) with  $l = 2^{j+1}$  and  $n = 2^j$  for the  $j$ th block. Referring to (\*), when  $\{k_n/n\}$  is itself nonincreasing, for any  $\varepsilon > 0$  and any nondecreasing positive sequence  $\{a_n\}$ , we obtain

$$(4.7) \quad \sum_{j=2}^{\infty} P\left\{\max_{2^j \leq n \leq 2^{j+1}} \sup_{x \leq T_{2^{j+1}}} \frac{|R_n(x)|}{n^2} \frac{k_{2n}}{a_n} \geq \varepsilon\right\} \leq 1568 \frac{C_1}{\varepsilon^2} \sum_{j=2}^{\infty} \frac{1}{a_{2^j}^2}.$$

The choice  $a_n \equiv \sqrt{c_n \log n}$  and the Borel–Cantelli lemma give the result. (Since some of the other terms  $R_{nj}$ ,  $j = 1, 2, 3, 4$ , appeared to be bigger in his proof, Stute states this only with  $\log^\alpha n$  replacing  $\sqrt{c_n \log n}$  for  $\alpha > 1/2$ .)

Assuming (\*), we now turn to improving Stute's almost sure bounds for  $R_{nj}^* = \sup_{x \leq Z_{n-k_n, n}} |R_{nj}(x)|$ ,  $j = 1, 2, 3, 4$ , showing that it is the one in (4.6) that dominates.

We use (4.3) freely when convenient. In  $R_{n1}^*$  and  $R_{n2}^*$  we can and do change  $d\tilde{H}_n$  to  $dH_n$ , since the increments of  $\tilde{H}_n$  are not greater than the corresponding increments of  $H_n$ . Similarly, we change  $d\tilde{H}$  to  $dH$  in  $R_{n4}^*$  after taking the absolute value under the integral. When dealing with  $R_{n1}^*$ , we factor out  $J_n^2$

from under the integral and use the second statement of Lemma 1 with  $p = 1/7$ , and then (4.1) for what remains, with  $J = H_n$  and  $\beta = 2$ . From  $R_{n2}^*$  we factor out  $E_n^2 J_n$  and use both the second and the fourth statements of Lemma 1, and again (4.1) for  $J = H_n$  and  $\beta = 2$ . Also,  $R_{n3}^* \leq n^{-1}[1 - H_-(Z_{n-k_n, n})]^{-1}$  and (4.3) may be used directly. Finally, from  $R_{n4}^*$  we factor out  $E_n$  and use (4.1) for  $J = H$  and  $\beta = 3/2$ . Thus we get

$$\begin{aligned} \limsup_{n \rightarrow \infty} k_n R_{n1}^* &\leq 14, \\ \limsup_{n \rightarrow \infty} \frac{k_n}{\log \log n} R_{n2}^* &\leq 56, \\ \limsup_{n \rightarrow \infty} n \frac{\sqrt{k_n}}{\sqrt{\log \log n}} R_{n4}^* &\leq 3\sqrt{14} \end{aligned}$$

and  $\limsup_{n \rightarrow \infty} k_n R_{n3}^* \leq 7$ . Using now (\*), the strong version of the proposition follows.

To prove also the weak version [Stute (1994b) probably had different proofs for the  $R_{nj}^*$ ; he does not give details], choose any  $\gamma \in (0, 1/2)$ ,  $p \in (0, 1)$  and  $1 \leq k_n < n$ . On the event  $Q_n(p) = \{Z_{n-k_n, n} \leq H^{-1}(1 - pk_n/n)\}$ , we have  $R_{n1}^* \leq 2J_n^2(p)/k_n$ ,  $R_{n3}^* \leq p^{-1}/k_n$ ,

$$R_{n4}^* \leq \frac{A_n^{(\gamma)}(p)}{n} \frac{1 + \gamma}{\gamma p^\gamma} \left(\frac{n}{k_n}\right)^\gamma$$

and

$$R_{n2}^* \leq J_n^{2\gamma}(p)[A_n^{(\gamma)}(p)]^2 \frac{1 + 2\gamma}{2\gamma} \left(\frac{n}{k_n}\right)^{2\gamma}.$$

Hence, by the first and the third statements of Lemma 1,  $R_{n1}^* + R_{n3}^* = \mathcal{O}_P(1/k_n)$ ,  $R_{n4}^* = \mathcal{O}_P(\sqrt{k_n}/[nk_n]) = \mathcal{O}_P(1/k_n)$  and  $R_{n2}^* = \mathcal{O}_P(n^{-1}[n/k_n]^{1-2\gamma}[n/k_n]^{2\gamma}) = \mathcal{O}_P(1/k_n)$  on  $Q_n(p)$  for each fixed  $p \in (0, 1)$ . (Note that it is important that we could choose  $\gamma > 0$ .) These, (4.6) and an application of (4.2) yield the weak version of the proposition.  $\square$

PROOF OF PROPOSITION 2. For all  $y \leq Z_{n, n}$  enjoyable algebra gives

$$\begin{aligned} &\frac{1}{[1 - H_n(y-)]^2} \\ &= 2 \frac{(1/n) - [1 - H_n(y-)]}{[1 - H_n(y-)]^3} + \frac{3}{[1 - H_n(y-)]^2} - \frac{2}{n} \frac{1}{[1 - H_n(y-)]^3} \\ &\quad + 2 \frac{[H_n(y-) - H_-(y)]^2}{[1 - H_-(y)]^3 [1 - H_n(y-)]} + \frac{[H_n(y-) - H_-(y)]^2}{[1 - H_-(y)]^2 [1 - H_n(y-)]^2}. \end{aligned}$$

Integration with respect to  $\tilde{H}_n$  gives a representation for  $d_n$ , and making the latter equation play the role of (4.3) in Stute (1994b) and imitating his

procedure for the present situation, we arrive at the following analog of (4.4):

$$d_n(x) - d_n^*(x) = -2 \frac{R_n^\circ(x)}{n^2} - 2R_{n1}^\circ(x) + R_{n2}^\circ(x) + 2R_{n3}^\circ(x) - 2R_{n4}^\circ(x),$$

$$x \leq Z_{n,n},$$

where  $R_n^\circ(x)$  is again a  $U$ -statistic minus its Hájek projection, so itself a degenerate  $U$ -statistic for which, paralleling (4.5), for all positive integers  $l$  and  $n$  and for all  $p \in (0, 1)$ ,

$$E \left( \sup_{x \leq T_l(p)} |R_n^\circ(x)|^2 \right) \leq \frac{C_2}{p^4} n^2 \frac{l^4}{k_l^4}$$

for some universal constant  $C_2 > 0$ , and  $R_{n1}^\circ(x) = n^{-1} \int_{-\infty}^x d\tilde{H}_n(y) / [1 - H_-(y)]^3$ ,

$$R_{n2}^\circ(x) = \int_{-\infty}^x \frac{[H_n(y-) - H_-(y)]^2 d\tilde{H}_n(y)}{[1 - H_-(y)]^3 [1 - H_n(y-)]} \\ + \int_{-\infty}^x \frac{[H_n(y-) - H_-(y)]^2 d\tilde{H}_n(y)}{[1 - H_-(y)]^2 [1 - H_n(y-)]^2},$$

$$R_{n3}^\circ(x) = \frac{1}{n} \int_{-\infty}^x \frac{d\tilde{H}_n(y)}{[1 - H_-(y)]^2}$$

and

$$R_{n4}^\circ(x) = \frac{1}{n} \int_{-\infty}^x \frac{H_n(y-) - H_-(y)}{[1 - H_-(y)]^3} d\tilde{H}_n(y).$$

With  $T_l = T_l(1/7)$ ,  $l = 1, 2, \dots$ , an arbitrary  $\varepsilon > 0$  and any nondecreasing positive sequence  $\{a_n\}$ , when  $\{k_n/n\}$  is itself nonincreasing the analog (4.7) is

$$\sum_{j=2}^{\infty} P \left\{ \max_{2^j \leq n \leq 2^{j+1}} \sup_{x \leq T_{2^j+1}} \frac{|R_n^\circ(x)|}{n^2} \frac{k_{2n}^2}{na_n} \geq \varepsilon \right\} \leq 1229312 \frac{C_2}{\varepsilon^2} \sum_{j=2}^{\infty} \frac{1}{a_{2^j}^2}.$$

Thus, using (4.3) for the almost sure statement under condition (\*), and setting  $l = n$  and using (4.2) for the probability bound for all  $1 \leq k_n < n$ , we obtain

$$\sup_{x \leq Z_{n-k_n, n}} \left| \frac{R_n^\circ(x)}{n^2} \right| = \begin{cases} \mathcal{O}_P \left( \frac{n}{k_n^2} \right), \\ o \left( \frac{n \sqrt{c_n \log n}}{k_{2n}^2} \right). \end{cases}$$

Also, setting  $R_{nj}^* = \sup_{x \leq Z_{n-k_n, n}} |R_{nj}^\circ(x)|$ ,  $j = 1, \dots, 4$ , and using the second and third statements of Lemma 1, by obvious modifications of the arguments above, we obtain  $R_{n1}^* = \mathcal{O}(n/k_n^2)$ ,  $R_{n2}^* = \mathcal{O}([n \log \log n]/k_n^2)$ ,  $R_{n3}^* = \mathcal{O}(1/k_n)$  and  $R_{n4}^* = \mathcal{O}(\sqrt{\log \log n}/k_n^{3/2})$ . Thus the strong half of the proposition is proved.

For the weak half we can now use the third statement of Lemma 1 with  $\gamma = 0$ . Combined with the first statement there and (4.1), on the event  $Q_n(p)$  we obtain  $R_{n1}^* \leq [3/2]J_n^3(p)n/k_n^2 = \mathcal{O}_P(n/k_n^2)$ ,  $R_{n3}^* \leq 2J_n^2(p)/k_n = \mathcal{O}_P(1/k_n)$ ,  $R_{n4}^* \leq 2p^{-1}A_n(p)/k_n = \mathcal{O}_P(1/k_n^{3/2})$  and  $R_{n2}^* \leq 2A_n^2(p)J_n(p)n/k_n + 2A_n^2n/k_n = \mathcal{O}_P(n/k_n^2)$  for each fixed  $p \in (0, 1)$ . Collecting the bounds, an application of (4.2) completes the proof.  $\square$

PROOF OF PROPOSITION 3. The strong statement here is Stute’s (1994b) result with a direct empirical-process proof; for a version see the second part of the next proof.

To establish the probability bound, put  $T_n(p) = H^{-1}(1 - pk_n/n)$ ,  $p \in (0, 1)$ , as before. By Lengart’s inequality [Shorack and Wellner (1986), page 893], for any  $C > 0$  and  $y > 0$  we have

$$P\left\{ \sup_{x \leq T_n(p)} |L_n(x)| \geq \frac{C}{\sqrt{k_n}} \right\} \leq \frac{yk_n}{C^2} + P\{\langle L_n \rangle(T_n(p)) \geq y\},$$

with the predictable quadratic variation  $\langle L_n \rangle$  given before the statement of the proposition. Factoring out  $J_n(p)$ , changing  $d\tilde{H}$  to  $dH$  and using (4.1) with  $J = H$  and  $\beta = 2$ , we obtain  $\langle L_n \rangle(T_n(p)) \leq 2J_n(p)/[pk_n]$ , so the choice  $y = 2K/[pk_n]$  and (4.2) yield

$$p_n(C) = P\left\{ \sup_{x \leq Z_{n-k_n, n}} |L_n(x)| \geq \frac{C}{\sqrt{k_n}} \right\} \leq e^{-Cp} + \frac{2K}{pC^2} + P\{J_n(p) \geq K\}$$

for every  $C > 0$ ,  $p \in (0, 1)$  and  $K > 0$ . Therefore,

$$\lim_{C \rightarrow \infty} \limsup_{n \rightarrow \infty} p_n(C) \leq e^{-Cp} + \limsup_{n \rightarrow \infty} P\{J_n(p) \geq K\}$$

for every  $p \in (0, 1)$  and  $K > 0$ . Letting first  $K \rightarrow \infty$  and using the first statement of Lemma 1, and then  $p \downarrow 0$ , we see that

$$\lim_{C \rightarrow \infty} \limsup_{n \rightarrow \infty} p_n(C) = 0. \quad \square$$

PROOF OF PROPOSITION 4. Since  $d_n^{\circ\circ}(x) = nR_{n4}^\circ(x)$  for all  $x \leq Z_{n, n}$ , both versions of the second statement are already established in the proof of Proposition 2.

To begin with the third statement, we now have  $\langle D_n^\circ \rangle(T_n(p)) \leq 4J_n(p)n^2/[3p^3k_n^3]$  by an application of (4.1) with  $J = H$  and  $\beta = 4$ . Hence an obvious version of the last martingale-theoretic argument yields

$$p_n^\circ(C) = P\left\{ \sup_{x \leq Z_{n-k_n, n}} |D_n^\circ(x)| \geq C \frac{n}{k_n^{3/2}} \right\} \leq e^{-Cp} + \frac{4K}{3p^3C^2} + P\{J_n(p) \geq K\}$$

for every  $C > 0$ ,  $p \in (0, 1)$  and  $K > 0$ . Thus  $\lim_{C \rightarrow \infty} \limsup_{n \rightarrow \infty} p_n^\circ(C) = 0$  as above.

By (4.3), it suffices to prove the first statement, which is an analog of the strong version of Proposition 3, with  $T_n = T_n(1/7) = H^{-1}(1 - k_n/7n)$  replacing  $Z_{n-k_n, n}$ . For  $0 \leq y \leq d(T_n)$  define  $d^{-1}(y) = \inf\{0 \leq x \leq T_n : d(x) \geq$

$y\}$ . Setting  $\varepsilon_n = n\sqrt{\log n/k_n^{3/2}}$ , let  $x_{r(n),n} = T_n$  and  $x_{l,n} = d^{-1}(l\varepsilon_n)$ ,  $l = 1, 2, \dots, r(n) - 1$ , where the integer  $r(n)$  is chosen such that  $d(x_{r(n)-1,n}) \leq d(T_n) \leq d(x_{r(n)-1,n} + \varepsilon_n)$ . Since  $0 \leq d(x) \leq 14n/k_n$  for  $x \leq T_n$  by (4.1), we see that  $r(n) < 14\sqrt{n}$  for  $n \geq 3$ , and since both  $d$  and  $d_n^\circ$  are nondecreasing functions, standard reasoning shows that

$$|d_n^\circ(x) - d(x)| \leq \max \left( \max_{1 \leq l \leq r(n)} |d_n^\circ(x_{l,n}) - d(x_{l,n})|, \max_{1 \leq l \leq r(n)} |d_n^\circ(x_{l,n} -) - d(x_{l,n} -)| \right) + \varepsilon_n$$

for any  $x \leq T_n$ . These two inequalities in turn imply for any constant  $C > 0$  that

$$q_n(C) = P \left\{ \sup_{x \leq T_n} |d_n^\circ(x) - d(x)| > (C + 1)\varepsilon_n \right\} < 28\sqrt{n} \sup_{x \leq T_n} P\{|d_n^\circ(x) - d(x)| > C\varepsilon_n\}.$$

Now introduce the independent and identically distributed random variables  $V_j(x) = I\{Z_j \leq x\} \delta_j / [1 - H_-(Z_j)]^2$  with mean  $d(x)$  and second moment  $\int_{-\infty}^x d\tilde{H}/[1 - H_-]^4$ , so that  $d_n^\circ(x) - d(x) = n^{-1} \sum_{j=1}^n [V_j(x) - d(x)]$ ,  $x \leq T_n$ . Since, again by (4.1),

$$\sum_{j=1}^n \text{Var}(V_j(x) - d(x)) \leq \frac{1372}{3} \frac{n^4}{k_n^3}$$

and

$$|V_j(x) - d(x)| \leq 49 \frac{n^2}{k_n^2}$$

for all  $x \leq T_n$ , we see by an application of the two-sided Bennett inequality that  $q_n(C) < 56\sqrt{n}/n^{\psi_\circ(C)}$ , where we also used  $k_n \geq \log n$  from (\*) and the fact that the function  $\psi$  is decreasing, and where  $\psi_\circ(C) = 3C^2\psi(3C/28)/2744 \rightarrow \infty$  as  $C \rightarrow \infty$ . Fixing a  $C > 0$  for which  $\psi_\circ(C) > 3/2$ , the Borel–Cantelli lemma implies the first statement.  $\square$

PROOF OF LEMMA 2. We have  $W_n^* \leq \sup\{|W_n(t)|: 0 \leq t \leq 2p^{-1}n/k_n\}$  by (4.1), so  $P\{W_n^* > x\} \leq 2P\{|W(2p^{-1}n/k_n)| > x\}$ ,  $x > 0$ , for a standard Wiener process  $W$ . This implies the first statement. Similarly,  $W_n^\circ \leq \sup\{|W(t, n)|: 0 \leq t \leq 14n/k_n\}/\sqrt{n}$ , so the second statement follows from Theorem 1.12.3 in M. Csörgő and Révész (1981). Finally,  $K_n^\circ \leq \sup\{|K(t, n)|: 0 \leq t \leq 14n/[14n + k_n]\}/\sqrt{n} \leq \sup\{|K(t, n)|: 0 \leq t \leq 1\}/\sqrt{n}$  since the function  $D(\cdot)$  is nondecreasing, so the third statement follows from Corollary 1.15.1 in M. Csörgő and Révész (1981).  $\square$

PROOF OF LEMMA 3. The first statement follows from Gill’s (1980) inequality [Shorack and Wellner (1986), pages 317 and 318] which extends to arbitrary  $F$  and  $G$  by writing the martingale-theoretic identities for  $H_n, \hat{F}_n$  and

the product-limit estimator  $\hat{G}_n$  of  $G$  on the whole line  $[-\infty, \tau_H]$ . [See Corollary 1 of Yang (1992) for an inequality related to one of Gill's.] The second is the important sample version due to Gill (1983) of the corresponding observation  $[1 - F(\cdot)][1 + d(\cdot)] \geq 1$  of Hall and Wellner (1980). Again, formally Gill has this for  $F$  and  $G$  supported on  $[0, \infty)$ . The general case follows as that for the first statement. An instructive elementary derivation for a continuous  $F$ , for which tied uncensored observations occur with probability 0 and Gill's definition of  $\hat{F}_n$  reduces to ours, is as follows. It suffices to show the inequality for  $x = Z_{l,n}$ ,  $l = 1, \dots, n - 1$ . It is trivial to check this for  $l = 1$ . Supposing that it holds for some  $l < n - 1$  and using the identity

$$(4.8) \quad \left[ 1 + n \sum_{j=1}^l \frac{1}{(n-j)(n-j+1)} \right] \prod_{j=1}^l \left\{ 1 + \frac{1}{n-j+1} \right\} = 1,$$

it follows for  $x = Z_{l+1,n}$ . The identity itself is the key to the reduction property of Hall and Wellner's (1980) confidence band mentioned in Section 2 above.

Since the ratio  $[1 + \hat{d}_n(x)]/[1 + d_n(x)]$  may change only at  $x = Z_{l,n}$ ,  $l = 1, \dots, n - k_n$ , and if this happens the jump of  $\hat{d}_n$  is greater than that of  $d_n$ , the supremum in the third statement is bounded by

$$\left[ 1 + n \sum_{j=1}^{n-k_n} (n-j)^{-1}(n-j+1)^{-1} \right] / \left[ 1 + n \sum_{j=1}^{n-k_n} (n-j+1)^{-2} \right],$$

which is not greater than  $1 + k_n^{-1}$ . This is the third assertion. The fourth follows from the simple bound  $\sup_{x \leq Z_{n-k_n,n}} [\hat{d}_n(x) - d_n(x)] \leq n \sum_{i=1}^{n-k_n} (n-i)^{-3} = n \sum_{j=k_n}^{n-1} j^{-3}$ .  $\square$

PROOF OF THEOREM 1. In view of (\*) for the strong half, Propositions 1 and 3 directly imply assertion (1.1) as do Propositions 2 and 4 for (1.2).

Recalling the notation of Proposition 5 and introducing the empirical processes  $\alpha_n(x) = \sqrt{n} [H_n(x) - H(x)]$  and  $\tilde{\alpha}_n(x) = \sqrt{n} [\tilde{H}_n(x) - \tilde{H}(x)]$ ,  $x \in \mathbb{R}$ , we have

$$(4.9) \quad \begin{aligned} \lambda_n^*(x) &= \int_{-\infty}^x \frac{d\tilde{\alpha}_n(y)}{1 - H_-(y)} + \int_{-\infty}^x \frac{\alpha_n(y-)}{[1 - H_-(y)]^2} d\tilde{H}(y) \\ &= \frac{\tilde{\alpha}_n(x)}{1 - H_-(x)} - \int_{-\infty}^x \frac{\tilde{\alpha}_n(y)}{[1 - H_-(y)]^2} dH_-(y) \\ &\quad + \int_{-\infty}^x \frac{\alpha_n(y-)}{[1 - H_-(y)]^2} d\tilde{H}(y) \end{aligned}$$

upon integrating by parts. Changing both  $dH_-$  and  $d\tilde{H}$  to  $dH$ , we have

$$\begin{aligned} R_{n5} &= \frac{1}{1 - H_-(Z_{n-k_n,n})} + \int_{-\infty}^{Z_{n-k_n,n}} \frac{dH_-(y)}{[1 - H_-(y)]^2} + \int_{-\infty}^{Z_{n-k_n,n}} \frac{d\tilde{H}(y)}{[1 - H_-(y)]^2} \\ &\leq \frac{1}{1 - H_-(Z_{n-k_n,n})} + 2 \int_{-\infty}^{Z_{n-k_n,n}} \frac{dH(y)}{[1 - H_-(y)]^2}. \end{aligned}$$



Hence, using (4.1)–(4.3), we obtain

$$(4.10) \quad \limsup_{n \rightarrow \infty} \frac{k_n}{n} R_{n5} \leq 35 \quad \text{if } (*) \text{ holds} \quad \text{and}$$

$$R_{n5} = \mathcal{O}_P\left(\frac{n}{k_n}\right) \quad \text{for all } 1 \leq k_n < n.$$

Now let  $K_0(s, u)$ ,  $0 \leq s \leq 1$ ,  $0 \leq u < \infty$ , be a Kiefer process, so that  $K_0(\cdot, n)/\sqrt{n}$  is a Brownian bridge for each  $n \in \mathbb{N}$ . Define  $\tilde{B}_n(x) = K_0(\tilde{H}(x), n)/\sqrt{n}$  and  $B_n^*(x) = \tilde{B}_n(x) + \bar{B}_n(x)$ , where  $\bar{B}_n(x) = [K_0(\bar{H}(x) + \tilde{H}(\infty), n) - K_0(\tilde{H}(\infty), n)]/\sqrt{n}$  and where  $\bar{H}(x) = P\{Z \leq x, \delta = 0\} = \int_{-\infty}^x [1 - F(y)] dG(y)$  is the censored analog of  $\tilde{H}$ , so that  $\tilde{H}(x) + \bar{H}(x) = H(x)$ ,  $x \in \mathbb{R}$ . The underlying probability space can be constructed to be rich enough to carry all the above processes such that

$$(4.11) \quad \sup_{x \in \mathbb{R}} |\tilde{\alpha}_n(x) - \tilde{B}_n(x)| = \mathcal{O}\left(\frac{\log^2 n}{\sqrt{n}}\right),$$

$$\sup_{x \in \mathbb{R}} |\alpha_n(x-) - B_n^*(x)| = \mathcal{O}\left(\frac{\log^2 n}{\sqrt{n}}\right)$$

hold jointly, almost surely. [See Sections 2, 3 and 9 in Burke, Csörgő and Horváth (1981) for more general versions of these statements. Also, Bonvalot and Castelle (1991) give some precise details for the Kiefer-process approximation of an empirical process due to Komlós, Major and Tusnády (1975a).] By easy calculation we have  $E(\tilde{B}_n(x)\tilde{B}_m(y)) = [(\tilde{H}(x) \wedge \tilde{H}(y)) - \tilde{H}(x)\tilde{H}(y)]q_{nm}$ ,  $E(B_n^*(x)B_m^*(y)) = [(H(x) \wedge H(y)) - H(x)H(y)]q_{nm}$ ,  $E(\tilde{B}_n(x)B_m^*(y)) = [(\tilde{H}(x) \wedge \tilde{H}(y)) - \tilde{H}(x)H(y)]q_{nm}$ ,  $x, y \in \mathbb{R}$ ,  $n, m \in \mathbb{N}$ , where  $q_{nm} = (n \wedge m)/\sqrt{nm}$ , matching the joint covariance structure of the two empirical processes  $\tilde{\alpha}(\cdot)$  and  $\alpha(\cdot)$ .

Assuming now that we are on this rich space, for  $x \in \mathbb{R}$  consider the process

$$G_n^*(x) = \frac{\tilde{B}_n(x)}{1 - H_-(x)} - \int_{-\infty}^x \frac{\tilde{B}_n(y)}{[1 - H_-(y)]^2} dH_-(y) + \int_{-\infty}^x \frac{B_n^*(y)}{[1 - H_-(y)]^2} d\tilde{H}(y).$$

Then  $\sup_{x \leq Z_{n-k_n, n}} |\lambda_n^*(x) - G_n^*(x)| = \mathcal{O}((\log^2 n)/\sqrt{n}) R_{n5}$ , and so, applying (4.10), we obtain

$$(4.12) \quad \sup_{x \leq Z_{n-k_n, n}} |\lambda_n^*(x) - G_n^*(x)| = \mathcal{O}(\sqrt{n}(\log^2 n)/k_n).$$

This relation and the first statement of Proposition 5 (which is nothing but Proposition 1), along with (\*), then yield

$$\sup_{x \leq Z_{n-k_n, n}} |\lambda_n(x) - G_n^*(x)| = \mathcal{O}(\sqrt{n}(\log^2 n)/k_{2n}).$$

Recalling that the sequence  $\{G_n^*(\cdot)\}_{n=1}^\infty$  is constructed from a single Kiefer process, it may be viewed as a bivariate Gaussian process with mean 0, and elementary calculation [cf. Breslow and Crowley (1974)] gives  $E(G_n^*(x)G_m^*(y)) =$

$(d(x) \wedge d(y))q_{nm}$ ,  $x, y \in \mathbb{R}$ ,  $n, m \in \mathbb{N}$ . This means that the process  $\{G_n^*(x): x \in \mathbb{R}, n \in \mathbb{N}\}$  is equal in distribution to the process  $\{W(d(x), n)/\sqrt{n}: x \in \mathbb{R}, n \in \mathbb{N}\}$  of the theorem. Changing the underlying probability space once more if necessary, (1.4) follows.

In the conceptually simpler case of (1.3), we begin with a suitable sequence  $\{W_n^0(s): 0 \leq s \leq 1\}$  of Brownian bridges replacing  $K_0(s, n)/\sqrt{n}$ ,  $0 \leq s \leq 1$ , above. [See Mason and van Zwet (1987) and Bretagnolle and Masart (1989) for two alternative sets of fine details giving the Komlós, Major and Tusnády (1975a) approximation of an empirical process by a sequence of Brownian bridges.] This means that we *redefine*  $\tilde{B}_n(x) = W_n^0(\tilde{H}(x))$  and  $\bar{B}_n(x) = W_n^0(\bar{H}(x) + \tilde{H}(\infty)) - W_n^0(\tilde{H}(\infty))$ , put  $B_n^*(x) = \tilde{B}_n(x) + \bar{B}_n(x)$  and redefine  $G_n^*(x)$  in terms of the new  $\tilde{B}_n(x)$  and  $B_n^*(x)$ ,  $x \in \mathbb{R}$ . Then we have (4.11) with  $\log^2 n$  reduced to  $\log n$  in both statements, and obtain (4.12) with  $\log^2 n$  reduced to  $\log n$ . This and the first statement of Proposition 5 then give

$$\sup_{x \leq Z_{n-k_n, n}} |\lambda_n(x) - G_n^*(x)| = \begin{cases} \mathcal{O}_P(\sqrt{n}(\log n)/k_n), & \text{for all } 1 \leq k_n < n, \\ \mathcal{O}(\sqrt{n}(\log n)/k_{2n}), & \text{if } (*) \text{ holds.} \end{cases}$$

Since, for each  $n$ , the process  $\{G_n^*(x): x \in \mathbb{R}\}$  is equal in distribution to the process  $\{W(d(x)): x \in \mathbb{R}\}$ , the two assertions in (1.3) also follow.

Writing  $T_n(p) = H^{-1}(1 - pk_n/n)$  as before and setting momentarily  $a_n(\cdot) = \lambda_n(\cdot) = \sqrt{n}[\Lambda_n(\cdot) - \Lambda(\cdot)]$  and  $b_n(\cdot) = W_n(d(\cdot))$ , on the event  $Q_n(p) = \{Z_{n-k_n, n} \leq T_n(p)\}$ ,

$$(4.13) \quad \begin{aligned} & \sup_{x \leq Z_{n-k_n, n}} \left| \frac{a_n(x)}{1 + d_n(x)} - \frac{b_n(x)}{1 + d(x)} \right| \\ & \leq \sup_{x \leq T_n(p)} \frac{|b_n(x)|}{1 + d(x)} \sup_{x \leq Z_{n-k_n, n}} |d_n(x) - d(x)| \\ & \quad + \sup_{x \leq Z_{n-k_n, n}} |a_n(x) - b_n(x)| \end{aligned}$$

for all  $1 \leq k_n < n$  and  $p \in (0, 1)$ . Since, for each  $n$ , the process  $W_n(d(\cdot))/[1 + d(\cdot)]$  is equal in distribution to  $B(D(\cdot))$ , where  $B(\cdot)$  is a Brownian bridge on  $[0, 1]$ , implying also that the first factor in the first term of the upper bound in (4.13) is  $\mathcal{O}_P(1)$ , the statement in (1.5) follows, generally on a new probability space, from (1.2) and (1.3) after using (4.2). Similarly, if now  $b_n(\cdot) = W(d(\cdot), n)/\sqrt{n}$ , then, with  $T_n = T_n(1/7)$  replacing  $T_n(p)$ , (4.13) holds under  $(*)$  almost surely by (4.3). Hence, taking the third statement of Lemma 2 into account, (1.6) follows from (1.2) and (1.4).  $\square$

PROOF OF THEOREM 2. Setting  $\ell_n(x) = -\log[1 - \hat{F}_n(x)] - \Lambda_n(x)$ ,  $x < Z_{n, n}$ , by the continuity of  $F$  and the classical expansion of Breslow and Crowley (1974),

$$\left| \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - h(x) \right| \leq |[\Lambda_n(x) - \Lambda(x)] - h(x)| + |R_{n6}(x)|$$

for any function  $h: \mathbb{R} \mapsto \mathbb{R}$ , where

$$R_{n6}(x) = \exp(|\Lambda_n(x) - \Lambda(x)|) \left[ \frac{1}{2} |\Lambda_n(x) - \Lambda(x)|^2 + |\ell_n(x)| \exp(|\ell_n(x)|) \right].$$

Since, the first inequality also going back to Breslow and Crowley (1974),

$$\sup_{x \leq Z_{n-k_n, n}} |\ell_n(x)| = \sup_{x \leq H_n^{-1}(1-(k_n/n))} |\ell_n(x)| \leq \frac{2}{n} \int_{-\infty}^{H_n^{-1}(1-(k_n/n))} \frac{d\tilde{H}_n(y)}{[1 - H_n(y-)]^2} \leq \frac{2}{n} \frac{2n}{k_n}$$

by (4.1), using now (1.1) and (\*) for the strong statement, we obtain

$$(4.14) \quad \sup_{x \leq Z_{n-k_n, n}} \left| \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - h(x) \right| \leq \sup_{x \leq Z_{n-k_n, n}} [|\Lambda_n(x) - \Lambda(x)| - h(x)] + R_{n6}^*,$$

where

$$R_{n6}^* = \sup_{x \leq Z_{n-k_n, n}} |R_{n6}(x)| = \begin{cases} \mathcal{O}_P\left(\frac{1}{k_n}\right), \\ \mathcal{O}\left(\frac{\log n}{k_{2n}}\right). \end{cases}$$

Choosing  $h(\cdot) = 0$  here, (1.1) implies (1.7). Assertion (1.8) follows from (1.7) and the first statement of Lemma 1. Also, choosing  $h(\cdot) = W_n(d(\cdot))/\sqrt{n}$  and  $h(\cdot) = W(d(\cdot), n)/n$  and multiplying over by  $\sqrt{n}$ , we see that (1.3) implies (1.9) and (1.4) implies (1.11).

Next, if we let  $a_n = \sqrt{n}[\hat{F}_n - F]/[1 - F]$  and keep  $b_n = W_n(d)$ , we see by (4.13), as in the proof of (1.5) above, that (1.2) and (1.9) imply (1.12). Similarly, keeping this  $a_n$ , letting  $b_n = W(d, n)/\sqrt{n}$  and referring to the third statement of Lemma 2 as for (1.6) above, we see by (4.13) that (1.2) and (1.11) imply (1.14).

Also, returning to  $a_n = \sqrt{n}[\hat{F}_n - F]/[1 - F]$ ,  $b_n = W_n(d)$ , on  $Q_n(p)$  we have

$$(4.15) \quad \begin{aligned} & \sup_{x \leq Z_{n-k_n, n}} \left| \sqrt{n} \frac{\hat{F}_n(x) - F(x)}{1 - \hat{F}_n(x)} - b_n(x) \right| \\ & \leq \sup_{x \leq T_n(p)} |b_n(x)| \sup_{x \leq Z_{n-k_n, n}} \frac{|\hat{F}_n(x) - F(x)|}{1 - \hat{F}_n(x)} \\ & \quad + \sup_{x \leq Z_{n-k_n, n}} |a_n(x) - b_n(x)| \left[ 1 + \sup_{x \leq Z_{n-k_n, n}} \frac{|\hat{F}_n(x) - F(x)|}{1 - \hat{F}_n(x)} \right]. \end{aligned}$$

Using (1.8) and the first statement of Lemma 2 for  $b_n$ , (1.10) follows from (1.9) via (4.2).

Finally, let  $a_n^* = \sqrt{n}[\hat{F}_n - F]$  and  $b_n^* = b_n[1 - F]$ , where  $b_n$  is either  $W_n(d)$  or  $W(d, n)/\sqrt{n}$ . Then  $[1 - \hat{F}_n(x)][1 + d_n(x)] \geq 1/2$  for  $x \leq Z_{n-k_n, n}$  by the second and third statements of Lemma 3 and, as a counterpart of (4.13), on

$Q_n(p)$  we obtain

$$\begin{aligned}
 & \sup_{x \leq Z_{n-k_n, n}} \left| \frac{\sqrt{n}[\hat{F}_n(x) - F(x)]}{[1 - \hat{F}_n(x)][1 + d_n(x)]} - \frac{b_n(x)}{1 + d(x)} \right| \\
 & \leq 2 \sup_{x \leq Z_{n-k_n, n}} |a_n^*(x) - b_n^*(x)| \\
 (4.16) \quad & + \sup_{x \leq T_n(p)} \frac{|b_n(x)|}{1 + d(x)} \sup_{x \leq Z_{n-k_n, n}} \frac{|d_n(x) - d(x)|}{1 + d_n(x)} \\
 & + 2 \sup_{x \leq T_n(p)} |b_n(x)| \sup_{x \leq Z_{n-k_n, n}} |\hat{F}_n(x) - F(x)|.
 \end{aligned}$$

Depending on the two choices of  $b_n(\cdot)$ , the first term is either  $\mathcal{O}_P([\sqrt{n} \log n]/k_n)$  or  $\mathcal{O}([\sqrt{n} \log^2 n]/k_{2n})$ , respectively, by (1.9) or (1.11). The third term is either  $\mathcal{O}_P(\sqrt{n}/k_n)$  or  $\mathcal{O}(\sqrt{\log n} \sqrt{n \log \log n}/k_{2n})$  by (1.7) and either the first or the second statement of Lemma 2, and (\*) in the second case. Replacing  $1/[1 + d_n(\cdot)]$  by 1, the second term is either  $\mathcal{O}_P(n/k_n^{3/2})$  or  $\mathcal{O}(n \sqrt{\log n} \sqrt{\log \log n}/k_{2n}^{3/2})$  by (1.2) and by either the fact that the first factor is  $\mathcal{O}_P(1)$  or by the third statement of Lemma 2, using of course that  $n^{-1/2}W(d(\cdot), n)/[1 + d(\cdot)]$  is equal in distribution to  $K_n(D(\cdot)) = n^{-1/2}K(D(\cdot), n)$ . For the second case we choose  $p = 1/7$ , when the inequality in (4.16) is almost sure by (4.2), whereas for the first case we let  $p \downarrow 0$  and make a final application of (4.2). Changing again the probability spaces if needed, we see that (1.13) and (1.15) follow from the unweighted versions of (1.9) and (1.11), respectively.  $\square$

**PROOF OF PROPOSITION 5.** The first statement is from Proposition 1. Putting  $h = \Lambda^* - \Lambda$  in (4.14) and multiplying over by  $\sqrt{n}$ , the second follows from the first. The third follows from the second by (1.7) and (1.8). The form of the  $l_j(\cdot)$  is clearest from a glance at (4.9), while the martingale property is trivial from that of  $l_1(\cdot) = L_1(\cdot)$ .  $\square$

**PROOF OF THE CLAIM.** Clearly, we only have to prove the statements for a  $\{k_n\}$  such that  $n/k_n \geq 10$  for all  $n$  large enough. Assuming this, consider the first claim, in which  $k_n^{-1} \log^2 n \rightarrow 0$ . Suppose that  $\sqrt{k_n} \xi_n = o_P(1)$  is universally possible, contrary to the claim. Then, multiplying (1.3), (1.9) and (1.10) over by  $\sqrt{k_n}/\sqrt{n}$ , the triangle inequality for the supremum norm forces

$$\begin{aligned}
 \eta_n &= \frac{\sqrt{k_n}}{\sqrt{n}} \sup_{x \leq H^{-1}(1-5(k_n/n))} |W_n(d(x))| \\
 (4.17) \quad & \leq \frac{\sqrt{k_n}}{\sqrt{n}} \sup_{x \leq Z_{n-k_n, n}} |W_n(d(x))| = o_P(1)
 \end{aligned}$$

for the sequence  $\{W_n(\cdot)\}$  of standard Wiener processes of the given construction, where the inequality holds almost surely by (4.3) for all  $n$  large enough.

Consider now the simple proportional hazards model of censorship [cf. Csörgő (1988) and the references therein] in which, for some  $\alpha \in (0, 1)$ , we have  $1 - F(x) = (1 - x)^\alpha$ ,  $1 - G(x) = (1 - x)^{1-\alpha}$ , so that  $H(x) = x$ ,  $\tilde{H}(x) = \alpha x$  and  $d(x) = \alpha[(1 - x)^{-1} - 1]$ ,  $0 \leq x \leq 1$ . Since  $H$  corresponds to the uniform distribution on  $[0, 1]$ ,  $H^{-1}(s) = s$ ,  $0 \leq s \leq 1$ , and simple calculation shows that, for all  $n$  large enough,

$$\zeta_n = \frac{\sqrt{k_n}}{\sqrt{n}} \sup \left\{ |W_n(t)| : 0 \leq t \leq \frac{\alpha}{10} \frac{n}{k_n} \right\} \leq \eta_n.$$

However, for any sequence  $\{W_n(\cdot)\}$  and a Wiener process  $W(\cdot)$ , the random variable  $\zeta_n$  is equal in distribution to  $\sqrt{\alpha} \sup\{|W(s)| : 0 \leq s \leq 1\} / \sqrt{10}$ , which contradicts (4.17).

The second part of the above proof is in fact nothing but checking that the first statement of Lemma 2 is optimal in the special model considered. Using (1.4) and (1.11) as starting points, the proof of the second claim reduces to showing that the second statement of Lemma 2 is also optimal in this model. This can again be verified by referring to the log log law for a Wiener sheet used in the proof of that statement.  $\square$

**5. Supplementary remarks.** The following are issues arising from the proofs and the results themselves, with a view toward possible refinements or complements.

REMARK 1. It is well known that weak and strong approximations with rates, on a common probability space, may not be optimal for mere weak convergence or log log laws under best conditions because of the necessary presence of some logarithmic factors. We conjecture that  $\lfloor \sqrt{n} a(n) \log n \rfloor$  can be replaced by  $\lfloor \sqrt{n} a(n) \rfloor$  in (2.1)–(2.4) and  $\lfloor \sqrt{n} a(n) \log^2 n \rfloor$  can also be replaced by  $\lfloor \sqrt{n} a(n) \rfloor$  in (2.9) and (2.10). This will probably follow from Proposition 5 directly through martingale central limit theorems or, via (4.2) or (4.3), central limit theorems and log log laws for triangular arrays in non-separable Banach spaces.

REMARK 2. Since  $D_n^{\leftarrow} = d_n^{\leftarrow} / (1 + d_n^{\leftarrow})$  and  $D = d / (1 + d)$  are well-defined distribution functions literally on the whole  $\mathbb{R}$  and the limiting process  $B(D(\cdot))$  makes sense in any censorship scenario, it is inevitable to think that (2.5)–(2.7) and the corresponding confidence bands stand a good chance to hold on the whole  $(-\infty, Z_{n,n})$  without (2.8) or in fact without any condition on censoring. Hall and Wellner (1980) conjectured this up to the largest uncensored observation  $X_n^*$  with  $d_n^{\leftarrow} = \hat{d}_n$  and Gill (1983, 1994) posed the problem repeatedly. However, Chen and Ying (1996) have shown that (2.5)–(2.7) and the corresponding confidence-band statements do *not* hold in general when  $Z_{n - \lfloor n^{2/3} a(n) \rfloor, n}$  is replaced by  $X_n^*$ . (I received their preprint while revising the present paper, half a year after I submitted it.) Formally, they claim this only for (2.7) with  $d_n^{\leftarrow} = \hat{d}_n$ . However, not only does their argument give the same

for (2.6), but in fact it reduces to showing that none of (2.6) and (2.7) holds up to  $X_n^*$  with  $d_n^a = d_n$ . It is also easy to see from their proof (at least when their parameter  $\tau$  is less than 1) that the same negative result holds for (2.5) as well. We now modify their idea to produce a concrete  $k_n^- \rightarrow \infty$  such that (2.5)–(2.7) and the corresponding convergence in distribution of the supremum functionals also break down when  $Z_{n-\lfloor n^{2/3}a(n) \rfloor, n}$  is replaced by  $Z_{n-k_n^-, n}$ .

Chen and Ying’s (1996) clever construction is defined by

$$\begin{aligned} n_l &= (l + 1)^l, & r_l &= (l + 1)^l, \\ a_l &= \frac{\tau}{K_Y} \sum_{i=1}^l \frac{1}{r_i}, & a_l^* &= \frac{a_{l-1} + a_l}{2}, \\ p_l &= \frac{1}{C_Y} \frac{r_l}{n_l}, & l &\in \mathbb{N}, \end{aligned}$$

where  $K_Y = \sum_{i=1}^\infty r_i^{-1} \in (0.62, 0.63)$ ,  $C_Y = \sum_{i=1}^\infty r_i n_i^{-1} \in (1.11, 1.12)$ , and choosing  $F(x) = x$ ,  $x \in [0, 1]$  and  $P\{Y = a_l\} = p_l$ ,  $l \in \mathbb{N}$ ; we put  $\tau = \tau_H \in (0, 1)$ , they allow  $\tau = 1$ . Shrinking their  $A_l$ , we introduce  $A_l^*$  as the event that among the  $n_l$  observations,  $X_n^* \in (a_l^*, a_l)$ , there are  $k_{n_l} + 1$  censored observations not less than  $a_l$ , and all the other uncensored observations among the remaining  $n_l - k_{n_l} - 2$  observations are less than  $a_{l-1}$ . Then, writing  $u_l \sim v_l$  if  $u_l/v_l \rightarrow 1$ , by the first part of Chen and Ying’s proof,

$$\begin{aligned} P\{A_l^*\} &= n_l P\{a_l^* < X < a_l, X \leq Y\} [1 - P\{a_{l-1} \leq X \leq Y\}]^{n_l - k_{n_l} - 2} p_l^* \\ &\sim \xi e^{-\xi} p_l^*/2 \end{aligned}$$

as  $l \rightarrow \infty$ , provided  $k_{n_l}/n_l \rightarrow 0$ , where  $\xi = \tau/(C_Y K_Y)$  and

$$p_l^* = \binom{n_l - 1}{k_{n_l} + 1} [P\{a_l \leq Y \leq X\}]^{k_{n_l} + 1} > \binom{n_l - 1}{k_{n_l} + 1} \left[ (1 - \tau) \sum_{i=l}^\infty p_i \right]^{k_{n_l} + 1} =: q_l^*,$$

using that  $1 - a_i > 1 - \tau$ ,  $i \geq l$ . Now let  $k_{n_l} = k_{n_l}(\tau) = \lceil x_l \rceil = \min\{j \in \mathbb{N} : j \geq x_l\}$  for all  $l \in \mathbb{N}$ , where  $x_l$  is the unique solution  $x$  of the equation  $x(x+1)^{1/2(x+1)} = e(1 - \tau)(n_l - 1) \sum_{i=l}^\infty p_i =: b_l$ . Note that  $b_l \sim e(1 - \tau)r_l/C_Y$ , so that  $k_{n_l}(\tau) \sim c_\tau l^l \sim c_\tau (\log n_l) / \log \log \log n_l$  as  $l \rightarrow \infty$ , where  $c_\tau = e^2(1 - \tau)/C_Y$ . Since

$$\begin{aligned} \binom{n_l - 1}{k_{n_l} + 1} &\sim \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{k_{n_l} + 1}} \left( \frac{n_l - 1}{k_{n_l} + 1} \right)^{k_{n_l} + 1} \left( \frac{n_l - 1}{n_l - k_{n_l} - 2} \right)^{n_l - k_{n_l} - 2} \\ &\sim \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{k_{n_l} + 1}} \left( \frac{n_l - 1}{k_{n_l} + 1} \right)^{k_{n_l} + 1} e^{k_{n_l} + 1} \end{aligned}$$

by Stirling’s formula and an expansion of  $\log(1 - (k_{n_l} + 1)(n_l - 1)^{-1})$ , we see that

$$q_l^* \sim \frac{1}{\sqrt{2\pi}} \left[ 1 + \frac{d_l - 1}{k_{n_l} + 1} \right]^{k_{n_l} + 1},$$

with

$$v_l \leq d_l = x_l \frac{(x_l + 1)^{1/2(x_l + 1)}}{(k_{n_l} + 1)^{1/2(k_{n_l} + 1)}} - k_{n_l} \leq u_l,$$

where, by further analysis,  $v_l \rightarrow -1$  and  $u_l \rightarrow 0$  as  $l \rightarrow \infty$ . Hence  $\lim_{j \rightarrow \infty} P\{A_{l_j}^*\} = p^*$  along a subsequence  $\{l_j\}_{j=1}^\infty \subset \mathbb{N}$ , where  $p^* \geq \xi e^{-\xi} e^{-\eta} / \sqrt{8\pi}$  for some  $\eta \in [1, 2]$ . Performing now the second part of Chen and Ying’s proof on the event  $A_{l_j}^*$ , using that  $k_{n_{l_j}}(\tau) \geq \lfloor (1 - \varepsilon)c_\tau(\log n_{l_j}) / \log \log \log n_{l_j} \rfloor$  for any fixed  $\varepsilon \in (0, 1)$  if  $j$  is large enough and that  $\lim_{\tau \downarrow 0} c_\tau > 6.597$ , we finally conclude that (2.5)–(2.7) and the corresponding confidence-band statements are *not* true in the present censorship model when  $Z_{n - \lfloor n^{2/3} a(n) \rfloor, n}$  is replaced by  $Z_{n - k_n^-, n}$ , where  $k_n^- \equiv \lfloor 6.597(\log n) / \log \log \log n \rfloor$ .

Asymptotically, there is a wide gap between  $k_n^-$  and  $k_n^+ \equiv \lfloor n^{2/3} \rfloor$ . Though, for fun, note also that  $k_{500}^- = 68$ ,  $k_{500}^+ = 62$ ;  $k_{750}^- = 68$ ,  $k_{750}^+ = 82$ ; and  $k_{1000}^- = 69$ ,  $k_{1000}^+ = 100$ .

REMARK 3. We get out to  $Z_{n - \lfloor n^{2/3} a(n) \rfloor, n}$  in (2.5)–(2.7) by spoiling the rate of the primary approximations in (1.3) and (1.9) by that of the distance for  $d_n - d$  in (1.2) when going to (1.5), (1.12) and (1.13). In this, dropping the weight function  $1 + d_n$  in (4.13) and (4.16) seems luxurious indeed. Working with  $d_n$  and letting  $g_n^\circ = \sqrt{n} [1 - F]^{-1} [1 + d]^{-1} [\hat{F}_n - F]$ , which, loosely speaking, by Gill’s (1983) Theorem 1.2 does converge in distribution to  $B(D)$  over the whole  $(-\infty, Z_{n, n})$  without any condition, a direct version of (4.16) is

$$\begin{aligned} & \sup_{x < Z_{n - k_n^-, n}} \left| \frac{\sqrt{n} [\hat{F}_n(x) - F(x)]}{[1 - \hat{F}_n(x)][1 + d_n(x)]} - B(D(x)) \right| \\ & \leq \sup_{x < Z_{n - k_n^-, n}} R_{n\tau}(x) \frac{|\hat{F}_n(x) - F(x)|}{1 - \hat{F}_n(x)} |g_n^\circ(x)| \\ & \quad + \sup_{x < Z_{n - k_n^-, n}} \frac{|d_n(x) - d(x)|}{[1 + d_n(x)][1 + d(x)]} \frac{\sqrt{n} |\hat{F}_n(x) - F(x)|}{1 - F(x)} \\ & \quad + \sup_{x < Z_{n - k_n^-, n}} |g_n^\circ(x) - B(D(x))|, \end{aligned}$$

where  $R_{n7}(x) = [1 - D_n(x)]/[1 - D(x)] = [1 + d(x)]/[1 + d_n(x)]$  and, below,  $R_{n7}^*(k_n) = \sup_{x < Z_{n-k_n, n}} R_{n7}(x)$ . In a Skorohod construction for Gill's present result, following its stopped-process statement in the formalism of Section 2.1, the third term goes to 0, in fact with  $k_n \equiv 0$ . The first term is  $\mathcal{O}_P(k_n^{-1/2})R_{n7}^*(k_n)$  by the same reason and (1.8). The second term may be viewed as  $\mathcal{O}_P(\sqrt{n}/\sqrt{k_n})R_{n8}^*(k_n)$  by (1.7), where  $R_{n8}^*(k_n) = \sup_{x < Z_{n-k_n, n}} |D_n(x) - D(x)|$ , and then the problem is the orders of  $R_{n7}^*(k_n)$  and  $R_{n8}^*(k_n)$  (maybe with some other estimator of  $D$ , different from  $D_n$ ). This term may also be looked upon as the supremum of  $R_{n7}(x)g_n^\circ(x)R_{n9}(x)$ , where  $R_{n9}(x) = |d_n(x) - d(x)|/[1 + d(x)]$  with its  $R_{n9}^*(k_n) = \sup_{x < Z_{n-k_n, n}} R_{n9}(x)$ . Then the question is the stochastic orders of  $R_{n7}^*(k_n)$  and  $R_{n9}^*(k_n)$ : a ratio problem and a weighted version of Propositions 2 and 4 for  $d_n$  or  $\hat{d}_n$ .

REMARK 4. We use the original Komlós–Major–Tusnányi embedding of the empirical process. Having some suitable versions of the propositions, the weighted approximations of M. Csörgő, S. Csörgő, Horváth and Mason (1986) and Mason and van Zwet (1987) would probably be the right tools to the question asked at the end of Section 2. For the universal problems in the present paper they appear to yield worse results than the Komlós–Major–Tusnányi approximations.

REMARK 5. If  $F$  is continuous, then the martingale jump process  $M_n(x) - M_n(x-)$  for  $M_n$  before Proposition 1 is uniformly bounded by  $1/n$ , but the bound in general is only 1. So if  $F$  is continuous, the exact same almost sure rates of Propositions 3 and 4, in the latter case for  $D_n^\circ$ , can be derived by using the exponential martingale inequality [Shorack and Wellner (1986), page 899]. The present elementary proof does not require the continuity of  $F$ .

REMARK 6. We do not have strong counterparts of (1.8), (1.10) and the third statement of Proposition 5 under (\*) because we are unaware of a strong ratio theory, paralleling that for the complete sample case by Wellner (1978), implying the almost sure boundedness of  $\sup_{x \leq H^{-1}(1-k_n/7n)} [1 - \hat{F}_n(x)]/[1 - F(x)]$ ; refer to (4.15). Whether or not this is true universally, or, if not, that it blows up at a controlled rate, we do not know. This is probably the most interesting ratio problem for  $\hat{F}_n$  to decide. [Note that the derivation of (1.15) avoids (1.10); this is made possible by Gill's second observation in Lemma 3.]

REMARK 7. The Gaussian approximation for the variance process  $\sqrt{n}[d_n(\cdot) - d(\cdot)]$  should be straightforward if tedious from Propositions 2 and 4 and either from the proof of Theorem 1 or the martingale representation of  $D_n^\circ$ , or the combination of the two, if the need arises. As in the proof of the claim above, an analog of (1.3) for this process could be helpful to establish at least part of the conjecture that the  $\mathcal{O}_P(n/k_n^{3/2})$  rate in (1.2) is universally best possible. Since, in the present approach, this determines the



rate in (1.5), (1.12) and (1.13), the problem may become important, and the range of  $\{k_n\}$  for which such a proof is applicable will most likely contain the sequence  $k_n \equiv \lfloor n^{2/3} \alpha(n) \rfloor$ .

REMARK 8. Note that in the proportional hazards model in the proof of the claim above, we have  $d(1) = d(\tau_H) = \infty$  and Gill's condition  $g(1) < \infty$  in (2.8) is satisfied if  $\alpha > 1/2$  but  $g(1) = \infty$  if  $\alpha \leq 1/2$ , whereas  $v(1) < \infty$  if  $\alpha \geq 1/2$  but  $v(1) = \infty$  if  $\alpha < 1/2$ . This is not to advocate the statistical use of this otherwise amusing submodel, the data in which may contain more information than the corresponding uncensored data for some estimation problems [cf. Csörgő and Mielniczuk (1988)]. Indeed, the *only* known data set that appears to follow the model is the Channing House data [cf. Csörgő (1988, 1989) and Henze (1993)] with the entry dates ignored as in Efron (1981).

## REFERENCES

- AALEN, O. O. (1976). Nonparametric inference in connection with multiple decrement models. *Scand. J. Statist.* **3** 15–27.
- ANDERSEN, P. K., BORGAN, Ø., GILL, R. D. and KEIDING, N. (1993). *Statistical Models Based on Counting Processes*. Springer, New York.
- BONVALOT, F. and CASTELLE, N. (1991). Strong approximation of uniform empirical process by Kiefer process. Preprint 91-41, Univ. Paris-Sud, Orsay, France.
- 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.
- BRETAGNOLLE, J. and MASSART, P. (1989). Hungarian constructions from the nonasymptotic viewpoint. *Ann. Probab.* **17** 239–256.
- BURKE, M. D., CSÖRGŐ, S. and HORVÁTH, L. (1981). Strong approximations of some biometric estimates under random censorship. *Z. Wahrsch. Verw. Gebiete* **56** 87–112.
- BURKE, M. D., CSÖRGŐ, S. and HORVÁTH, L. (1988). Correction to and improvement of “Strong approximations of some biometric estimates under random censorship.” *Probab. Theory Related Fields* **79** 51–57.
- CHEN, K. and YING, Z. (1996). A counterexample to a conjecture concerning the Hall–Wellner band. *Ann. Statist.* **24** 641–646.
- CSÁKI, E. (1977). The law of the iterated logarithm for normalized empirical distribution function. *Z. Wahrsch. Verw. Gebiete* **38** 147–167.
- CSÖRGŐ, M., CSÖRGŐ, S., HORVÁTH, L. and MASON, D. M. (1986). Weighted empirical and quantile processes. *Ann. Probab.* **14** 31–85.
- CSÖRGŐ, M. and RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press, New York.
- CSÖRGŐ, S. (1988). Estimation in the proportional hazards model of random censorship. *Statistics* **19** 437–463.
- CSÖRGŐ, S. (1989). Testing for the proportional hazards model of random censorship. In *Proceedings of the Fourth Prague Symposium on Asymptotic Statistics (1988)* (P. Mandl and M. Hušková, eds.) 41–53. Charles Univ. Press, Prague.
- CSÖRGŐ, S. and HORVÁTH, L. (1982). Statistical inference from censored samples. *Alkalmaz. Mat. Lapok* **8** 1–89. (In Hungarian.)
- CSÖRGŐ, S. and HORVÁTH, L. (1983). The rate of strong uniform consistency for the product-limit estimator. *Z. Wahrsch. Verw. Gebiete* **62** 411–426.

- CSÖRGÓ, S. and HORVÁTH, L. (1986). Confidence bands from censored samples. *Canad. J. Statist.* **14** 131–144.
- CSÖRGÓ, S. and MIELNICZUK, J. (1988). Density estimation in the simple proportional hazards model. *Statist. Probab. Lett.* **6** 419–426.
- EFRON, B. (1967). The two sample problem with censored data. *Proc. Fifth Berkeley Symp. Math Statist. Probab.* **4** 831–853. Univ. California Press, Berkeley.
- EFRON, B. (1981). Censored data and the bootstrap. *J. Amer. Statist. Assoc.* **76** 312–319.
- GILL, R. D. (1980). *Censoring and Stochastic Integrals. Math. Center Tracts* **124**. Math. Centrum, Amsterdam.
- GILL, R. D. (1983). Large sample behavior 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** 115–241. Springer, Berlin.
- GU, M. G. (1991). The Chung–Smirnov law for the product-limit estimator under random censorship. *Chinese Ann. Math. Ser. B* **12** 96–105.
- GU, M. G. and LAI, T. L. (1990). Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *Ann. Probab.* **18** 160–189.
- HALL, W. J. and WELLNER, J. A. (1980). Confidence bands for a survival curve from censored data. *Biometrika* **67** 133–143.
- HENZE, N. (1993). A quick omnibus test for the proportional hazards model of random censorship. *Statistics* **24** 253–263.
- KAPLAN, E. L. and MEIER, P. (1958). Non-parametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481, 562–563.
- KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975a). An approximation of partial sums of independent rv's and the sample df. *Z. Wahrsch. Verw. Gebiete* **32** 111–131.
- KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975b). Weak convergence and embedding. In *Limit Theorems of Probability Theory. Colloquia Mathematica Societatis János Bolyai* (P. Révész, ed.) **11** 149–165. North-Holland, Amsterdam.
- LO, S.-H. and SINGH, K. (1985). The product-limit estimator and the bootstrap: some asymptotic representations. *Probab. Theory Related Fields* **71** 455–465.
- MAJOR, P. and REJTŐ, L. (1988). Strong embedding of the estimator of the distribution function under random censorship. *Ann. Statist.* **16** 1113–1132.
- MASON, D. M. (1985). The asymptotic distribution of generalized Rényi statistics. *Acta Sci. Math. (Szeged)* **48** 315–323.
- MASON, D. M. and VAN ZWET, W. R. (1987). A refinement of the KMT inequality for the uniform empirical process. *Ann. Probab.* **15** 871–884.
- NAIR, V. N. (1984). Confidence bands for survival functions with censored data: a comparative study. *Technometrics* **26** 265–275.
- NELSON, W. (1972). Theory and applications of hazard plotting for censored failure data. *Technometrics* **14** 945–966.
- POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York.
- SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.
- SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.
- STUTE, W. (1994a).  $U$ -statistic processes: a martingale approach. *Ann. Probab.* **22** 1725–1744.
- STUTE, W. (1994b). Strong and weak representations of cumulative hazard functions and Kaplan–Meier estimators on increasing sets. *J. Statist. Plann. Inference* **42** 315–329.
- 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.
- WELLNER, J. A. (1978). Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. Verw. Gebiete* **45** 73–88.
- YANG, S. (1992). Some inequalities about the Kaplan–Meier estimator. *Ann. Statist.* **20** 535–544.

- YING, Z. (1989). A note on the asymptotic properties of the product-limit estimator on the whole line. *Statist. Probab. Lett.* **7** 311–314.
- ZHOU, M. (1991). Some properties of the Kaplan–Meier estimator for independent nonidentically distributed random variables. *Ann. Statist.* **19** 2266–2274.

DEPARTMENT OF STATISTICS  
UNIVERSITY OF MICHIGAN  
1440 MASON HALL  
ANN ARBOR, MICHIGAN 48109-1027