CONFIDENCE BANDS FROM CENSORED SAMPLES

BY

SÁNDOR CSÖRGŐ AND LAJOS HORVÁTH

TECHNICAL REPORT NO. 105
AUGUST 1985

PREPARED UNDER THE AUSPICES
OF
PUBLIC HEALTH SERVICE GRANT 5 RO1 GM21215-11

DIVISION OF BIOSTATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
CONFIDENCE BANDS FROM CENSORED SAMPLES

By
Sándor Csörgő and Lajos Horváth

TECHNICAL REPORT NO. 105
August 1985

PREPARED UNDER THE AUSPICES
OF
PUBLIC HEALTH SERVICE GRANT 5 R01 GM21215-11

DIVISION OF BIOSTATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
Confidence bands from censored samples

SÁNDOR CSÖRGÖ and LAJOS HORVÁTH

Bolyai Institute, Szeged University, Szeged, Hungary

Key words and phrases. Censored data, confidence band, product-limit estimator, Brownian motion, narrowed Kolmogorov band.

AMS 1980 subject classifications. Primary 62G15; secondary 62E20, 62E25, 62N05

ABSTRACT

Based on isolated individual ideas, several types of asymptotic confidence bands have been proposed in the literature when the data are randomly censored on the right. Introducing new classes of bands we find the proper places of the old bands and their relationship to one another within a comprehensive theory of bands. A thorough analysis yields generally narrower bands and two kinds of modifications which are asymptotically distribution and censor-free. One of these kinds is useful when the interval, on which the bands are constructed, is predetermined and the width of the bands is random. The other one is advantageous when there is a predetermined bound on the width, and then the interval is random. We illustrate our bands on the Szeged Pacemaker Data. The considerations also provide a general modification of the Kolmogorov band in the uncensored case.

1. INTRODUCTION

Let $X_1^o, X_2^o, \ldots$ be a sequence of independent real random variables with common continuous distribution function $F^o$. Another sequence, independent of the $\{X_j^o\}$, of independent random variables $Y_1, Y_2, \ldots$ with common left-continuous distribution function $H$ censors on the right the preceding one, so that the observations available to us at the $n$-th stage consist of the pairs $(X_j, \delta_j)$, $1 \leq j \leq n$, where $X_j = \min(X_j^o, Y_j)$ and $\delta_j$ is the indicator of the event $\{X_j = X_j^o\}$. We define the Kaplan-Meier (1958) product-limit estimator $\hat{F}_n^o$ of $F^o$ by

$$1 - \hat{F}_n^o(t) = \begin{cases} \prod_{1 \leq j \leq n : X_j < t} \left( \frac{n - N_j^o}{n} \right)^{-1} \delta_j & (t \leq X_n:n), \\ 0 & (t > X_n:n), \end{cases}$$

where $X_n:n = \max(X_1, \ldots, X_n)$ and $N_j:n = \#\{k : 1 \leq k \leq n, X_k < X_j\}$. We note, however, that all the considerations in this paper remain valid if $\hat{F}_n^o$ is defined by any other quantity beyond $X_n:n$, different from 1. For a left continuous distribution function $G$ set $T_G = \inf\{t : G(t) = 1\}$. Consider the weighted product-limit processes $Z_n(t) = n \frac{\hat{F}_n^o(t) - F^o(t)}{1 - \hat{F}_n^o(t)} / \{1 - F^o(t)\}$ and

$$\hat{Z}_n(t) = n \frac{\hat{F}_n^o(t) - F^o(t)}{1 - \hat{F}_n^o(t)} / \{1 - F_n^o(t)\}$$

together with their limiting variance function

$$d(t) = \int_{-\infty}^{t} \{1 - F^o(s)\}^{-1} \{1 - H(s)\}^{-1} dF^o(s)$$

$$= \int_{-\infty}^{t} \{1 - F(s)\}^{-2} d\bar{F}(s),$$
where \( F \) is the left-continuous distribution function of the minimum
variables \( X_1, X_2, \ldots \), \( 1-F = (1-F^o)(1-H) \), and \( F \) is the sub-distribution
function of the uncensored observations: \( F(s) = \Pr(X_1 < s, \delta_1 = 1) \).
Integrals of the form \( \int_a^b \) are meant as \( \int_{[a,b]} \) throughout. The continuity
of \( F^o \) implies that \( d \) is a continuous nondecreasing function on
\((-\infty, T_F)\). Let \( a(t) = \inf\{u: d(u) \geq t\} \) be its inverse, so that \( d(a(t)) = t \).
Efron (1967) noted that the processes \( Z_n(a(t)) \) approach as \( n \to \infty \)
to the standard Wiener process \( W \). Indeed, if \( F^o(0) = H(0) = 0 \) and
\( H \) is also continuous Breslow & Crowley (1974) and Aalen (1976), the
latter in a special case, proved that if \( T < T_F \) then \( Z_n(\cdot) \) converges
weakly in Skorohod's space \( D[0,T] \) to \( W(d(\cdot)) \), and the same is true
for \( \hat{Z}_n(\cdot) \).

One of the basic problems of practical interest is the construction
of confidence bands for \( F^o \). Efron's above noted transformation into
the Brownian motion \( W \) easily leads to confidence bands, we call them
the E-bands, for \( F^o \) on \((-\infty, T)\) by considering the processes \( Z_n \) and
\( \hat{Z}_n \) between the constant boundaries \( \pm \lambda \). Gillespie & Fisher (1979)
generalised this approach by considering the more general boundaries
\( \pm\{\lambda_1 + \lambda_2 d_n(t)\} \), where \( d_n(t) \) is an appropriate empirical version of
d\( (t) \) defined in Section 2 below and \( \lambda_1 > 0 \) and \( \lambda_2 \geq 0 \). We shall
refer to the resulting band as the GF-band. However, they considered this
only for the process \( Z_n \) and not for \( \hat{Z}_n \). Hall & Wellner (1980) based
their considerations on the scale-changed Brownian bridge obtained from
\( W(d(t)) \) by Doob's transformation, and transformed the process \( \hat{Z}_n \)
accordingly. However, their band may be viewed as one obtained by
considering \( \hat{Z}_n \) between the boundaries \( \pm \lambda \{1 + d_n(t)\} \). They did not
consider the analogous one for the process \( Z_n \). The HW-band has the
property of reducing to the Kolmogorov band in the absence of censoring.
Although the HW-band may be used conservatively in a distribution-free
manner, the common property of all these bands is that their limiting
distribution depends on the usually unknown quantity \( d(T) \) determined by
a mutual relationship between \( F_0 \) and \( H \). To overcome this difficulty,
and following up an earlier proposition of Aalen (1976), Nair (1981, 1982)
has considered the process \( Z_n \) between the boundaries \( \pm \lambda d_n^{1/2}(T) \).
The resulting AN-band is asymptotically distribution and censor-free. The
Introduction of Hall & Wellner (1980) and Section 2 of Nair (1981)
suggest that the relationship of all these bands to each other remained
unclear and the same is true for the earlier papers of the present authors
in the reference list. In Section 2 we analyse all these bands. It
turns out advantageous to introduce all the above type bands based both
on \( Z_n \) and \( \hat{Z}_n \). Then the HW-band appears as a special case of the
GF-type bands based on \( \hat{Z}_n \), while the AN-bands, based on both \( Z_n \) and
\( \hat{Z}_n \), appear as distribution-free versions of the special E-bands.
Considering both classes of bands based on \( Z_n \) and \( \hat{Z}_n \) is also advan-
tageous in that the lower contours of the bands based on \( \hat{Z}_n \) are always
above those based on \( Z_n \), and the same is true vice versa of the upper
contours. This observation leads to the notion of (somewhat narrowed)
mixed bands, the lower contour chosen from one class while the upper one
from the other. This notion will be convenient in the discussion of the
properties of the bands. But then the relationship between the E and
AN-bands suggests finding asymptotically distribution and censor-free
modifications of the whole class of mixed bands such that the mixed
AN-band would be just one member of the new classes of bands. We present two kinds of such modifications. The one in Section 3 is advantageous if the length of the interval on which we construct the band is pre-determined. The other one in Section 4, based on a theoretical result in Appendix 1 on the empirically Efron-transformed processes $Z_n$ and $\hat{Z}_n$, can be useful if we wish to bound the width of the bands. The two kinds of modifications are compared in Section 5. In Section 6 we illustrate some of our bands on the censored survival times of 647 patients of the Department of Heart Surgery, University Medical School, Szeged, who have had heart pacemakers implanted.

Although the reduction property of the HW-band is nice, it is well-known that the Kolmogorov band is not optimal in a number of uncensored situations. Hence this property, or symmetry about $\hat{F}_n^o$, cannot serve as a criterion to rule out other bands. Indeed, our considerations with censored bands lead us to an interesting and simple modification of the Kolmogorov band in the uncensored case. This is presented in Appendix 2.

We note that Gill (1983) has extended the HW-band to the whole interval $[0,T_F]$, for positive random variables, instead of $[0,T]$ with $T < T_F$, provided

$$\int_0^{T_F} (1-H)^{-1} d\Phi^o < \infty.$$  

We prefer to work on intervals $(-\infty,T]$ because we find it more convenient to find an appropriate value $T < T_F$ in practice than to check conditions like Gill's on the heaviness of censoring which, if not satisfied, may force us back to the choice of a $T < T_F$. 

Throughout we concentrate on confidence bands and will not comment on the new asymptotically distribution-free Kolmogorov-Smirnov and Cramér-von Mises type statistics for goodness of fit, arising from Sections 3 and 4. This has been done by Koziol (1980) for the Hall-Wellner case, and it would perhaps be of interest to conduct similar studies for the present statistics in a separate article. The present authors do not plan to do this. Likewise, to keep the exposition of the bands as simple as possible, we do not consider extra weighting functions as in Nair (1981, 1982), though it would be possible to do so.

2. THE E, GF, HW, AND AN-BANDS

As before, we assume throughout this and the following sections that \( T < T_f \). Let \( F_n(t) \) be the ordinary left-continuous sample distribution function of the minimum variables \( X_1, \ldots, X_n \) and let
\[
\tilde{F}_n(t) = n^{-1} \# \{ k : 1 \leq k \leq n, X_k < t, \delta_k = 1 \} = n^{-1} \# \{ k : 1 \leq k \leq \pi_n, U_k < t \}
\]
be the empirical sub-distribution function of the ordered uncensored observations \( U_{1:n} \leq \cdots \leq U_{\pi_n:n} \). Here \( \pi_n = \# \{ k : 1 \leq k \leq n, \delta_k = 1 \} \) is the number of the uncensored observations out of the first \( n \) observations. In order to remain in accord with the reduction principle of Hall & Wellner (1980), we define the estimate of the asymptotic variance function as
\[
d_n(t) = \int_{-\infty}^{t} (1-F_n^+(1-F_n))^{-1} d\tilde{F}_n
\]
\[
= n \sum_{\{ 1 \leq j \leq \pi_n: X_j < t \}} (n-j)^{-1} (n-j-1)^{-1} \delta_j
\]
\[ n^{-1} \sum_{1 \leq i < j \leq n: U_{ij} < t} (1 - F_n^+(U_{ij}: n))^{-1} (1 - F_n(U_{ij}: n))^{-1}, \]

\[ -\infty < t < \infty, \]

where \( F_n^+ \) is the right-continuous version of \( F_n \).

The respective \( E, HW, GF, \) and \( AN \)-bands \( E_n(t; \lambda_E), H_n(t; \lambda_H), \)
\( G_n(t; \lambda_1, \lambda_2), \) and \( A_n(t; \lambda_A), \) based on the process \( Z_n \), all have the form

\[
\left[ \frac{F_n^O(t) - r_n(t)}{1 - r_n(t)}, \frac{F_n^O(t) + r_n(t)}{1 + r_n(t)} \right], \quad -\infty < t < T,
\]

while the respective \( E, HW, GF, \) and \( AN \)-bands \( \hat{E}_n(t; \lambda_E), \hat{H}_n(t; \lambda_H), \)
\( \hat{G}_n(t; \lambda_1, \lambda_2), \) and \( \hat{A}_n(t; \lambda_A), \) based on the process \( \hat{Z}_n \), all have the form

\[
[\hat{F}_n^O(t) - \hat{r}_n(t), \hat{F}_n^O(t) + \hat{r}_n(t)], \quad -\infty < t < T,
\]

where

\[ r_n(t) \quad \hat{r}_n(t) \]

<table>
<thead>
<tr>
<th>in ( E_n ) resp. ( \hat{E}_n )</th>
<th>( \lambda_E n^{-\frac{1}{2}} )</th>
<th>( \lambda_E n^{-\frac{1}{2}} (1 - \hat{F}_n^O(t)) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>in ( H_n ) resp. ( \hat{H}_n )</td>
<td>( \lambda_H n^{-\frac{1}{2}} (1 + d_n(t)) )</td>
<td>( \lambda_H n^{-\frac{1}{2}} (1 - F_n^O(t)) { 1 + d_n(t) } )</td>
</tr>
<tr>
<td>in ( G_n ) resp. ( \hat{G}_n )</td>
<td>( n^{-\frac{1}{2}} { \lambda_1 + \lambda_2 d_n(t) } )</td>
<td>( n^{-\frac{1}{2}} (1 - F_n^O(t)) { \lambda_1 + \lambda_2 d_n(t) } )</td>
</tr>
<tr>
<td>in ( A_n ) resp. ( \hat{A}_n )</td>
<td>( \lambda_A n^{-\frac{1}{2}} d_n^2(T) )</td>
<td>( \lambda_A n^{-\frac{1}{2}} (1 - \hat{F}_n^O(t)) d_n^2(T) )</td>
</tr>
</tbody>
</table>

We note that \( H_n(t; \lambda_H) \) was not considered by Hall & Wellner (1980) and
\( \hat{G}_n(t; \lambda_1, \lambda_2) \) was not considered by Gillespie & Fisher (1979). The above
\( \lambda \)'s are chosen as follows. For \( c > 0 \) and \( \lambda > 0 \), set
\( K_c(\lambda) = \Pr\{ \sup_{0 \leq t \leq c} |W(t)| \leq \lambda \} = \Pr\{-\lambda \leq W(t) \leq \lambda, 0 \leq t \leq c\} \) (2.1)

and for \( \lambda_1 > 0 \) and \( \lambda_2 \geq 0 \), set

\[ Q_c(\lambda_1, \lambda_2) = \Pr\{-(\lambda_1 + \lambda_2 t) \leq W(t) \leq \lambda_1 + \lambda_2 t, 0 \leq t \leq c\}. \] (2.2)

Of course \( Q_c(\lambda, 0) = K_c(\lambda) \), and since \( B(s) = (1-s)W(s(1-s)^{-1}), 0 < s < 1 \),
is a Brownian bridge by Doob’s transformation, we may write

\[ Q_c(\lambda, \lambda) = \Pr\{ \sup_{0 \leq t \leq c/(1+c)} |B(t)| \leq \lambda \}. \]

Anderson’s (1960) formula for \( Q_c(\lambda_1, \lambda_2) \) is given in Gillespie & Fisher (1979), containing the formula for \( Q_c(\lambda, \lambda) \) given in Hall & Wellner (1980) who give various references for the tabulation of the latter.

The well-known formula for \( K_1(\lambda) = Q_1(\lambda, 0) \) is extensively tabulated in Csörgő & Horváth (1985). Now \( \lambda_A > 0 \) of the AN-bands above is chosen as \( K_1(\lambda_A) = 1-\alpha \), for a given level \( \alpha \) \((0 < \alpha < 1)\), and the \( \lambda_1 > 0 \)
and \( \lambda_2 \geq 0 \) in the GF-bands are chosen to satisfy

\[ Q_d(T)(\lambda_1, \lambda_2) = 1-\alpha. \] (2.3)

Of course, we may fix one of \( \lambda_1 \) and \( \lambda_2 \) and vary the other one to achieve this equality. Hence, fixing \( \lambda_2 = 0 \), we choose \( \lambda_E \) such that

\[ K_d(T)(\lambda_E) = Q_d(T)(\lambda_E, 0) = 1-\alpha. \] Indeed, we see that \( E_n(t; \lambda_E) \equiv G_n(t; \lambda_E, 0) \) and \( \hat{E}_n(t; \lambda_E) \equiv \hat{G}_n(t; \lambda_E, 0) \) for each \( t \) and \( n \). Similarly, \( H_n(t; \lambda_H) \equiv G_n(t; \lambda_H, \lambda_H) \) and \( \hat{H}_n(t; \lambda_H) \equiv \hat{G}_n(t; \lambda_H, \lambda_H) \) for each \( t \) and \( n \), where \( \lambda_H \) is then chosen such that \( Q_d(T)(\lambda_H, \lambda_H) = 1-\alpha. \)

This means that the E-bands and the HW-bands are special cases of the corresponding general GF-bands. Therefore, it is sufficient to talk about the latter, in general, and about the AN-bands. In this notation,
it is the special GF-band $\hat{G}_n(t; \lambda_1, \lambda_2)$ of Hall & Wellner (1980) which reduces to the Kolmogorov band in the absence of censoring. As a consequence, basically, of the Breslow-Crowley theorem, we have

$$\lim_{n \to \infty} \Pr\left\{ F^O(t) \in G_n(t; \lambda_1, \lambda_2), \quad -\infty < t < T \right\} = \lim_{n \to \infty} \Pr\left\{ F^O(t) \in \hat{G}_n(t; \lambda_1, \lambda_2), \quad -\infty < t < T \right\} = 1 - \alpha$$

and

$$\lim_{n \to \infty} \Pr\left\{ F^O(t) \in A_n(t; \lambda_A), \quad -\infty < t < T \right\} = \lim_{n \to \infty} \Pr\left\{ F^O(t) \in \hat{A}_n(t; \lambda_A), \quad -\infty < t < T \right\} = 1 - \alpha.$$  

We note that the previously mentioned authors assumed that $F^O(0) = 0$, and hence considered bands on $[0, T]$ only, and they also assumed that $H$ is also continuous. Formally, the above four limit relations follow from Corollary 6.1 in Burke et al. (1981) and Theorem 3.2 in Horváth (1980). In fact the latter results and the ones for the AN-bands in Csergő & Horváth (1982a) imply that the rate of convergence in the four limit theorems above is at least $O(n^{-1/3} (\log n)^{3/2})$. Including the negative half line is of course not so important, but it is useful to allow discontinuous censoring distributions. For example, $Y_1, Y_2, \ldots$ may now degenerate at a truncation constant $Y$, a situation often arising in practice.

Now we turn to the discussion of the bands. Although we only have to deal with GF- and AN-bands, it is still inconvenient that we have two sets of asymptotically equivalent such bands based on the asymptotically equivalent two processes $Z_n$ and $\hat{Z}_n$. However, introducing the mixed GF-band

$$G_n^*(t; \lambda_1, \lambda_2) = \left[ \frac{1}{n^{1/2}} \left\{ \frac{F_n^O(t) \{\lambda_1 + \lambda_2 d_n(t)\}}{n^{1/2}} - \frac{\dot{F}_n^O(t) + \frac{n^{-1/2} \{\lambda_1 + \lambda_2 d_n(t)\}}{1 + n^{-1/2} \{\lambda_1 + \lambda_2 d_n(t)\}}}{1 + n^{-1/2} \{\lambda_1 + \lambda_2 d_n(t)\}} \right\} \right]$$
and the mixed AN-band

\[ A^*_n(t; \lambda_A) = \left[ \frac{\hat{G}_n(t) - \lambda_A \left( 1 - \frac{\hat{G}_n(t)}{n} \right) d_n^1(T)}{n^{1/2}} , \frac{\hat{F}_n(t) + \lambda_A n^{-1/2} d_n^1(T)}{1 + \lambda_A n^{-1/2} d_n^1(T)} \right] \]

we notice that if \( n \) is so large that \( n^{-1/2} \{ \lambda_1 + \lambda_2 d_n(T) \} < 1 \) and \( n^{-1/2} \lambda d_n^1(T) < 1 \), respectively, then

\[ G^*_n(t; \lambda_1, \lambda_2) \subset G_n(t; \lambda_1, \lambda_2) \cap \hat{G}_n(t; \lambda_1, \lambda_2) \text{ and } A^*_n(t; \lambda_A) \subset A_n(t; \lambda_A) \cap \hat{A}_n(t; \lambda_A), \quad (2.4) \]

and these inclusions are strict for any \(-\infty < t \leq T\). The mentioned extensions of the Breslow-Crowley theorem easily imply

\[ \lim_{n \to \infty} \Pr \{ \tilde{F}_n(t) \in G^*_n(t; \lambda_1, \lambda_2), \; -\infty < t \leq T \} = 1 - \alpha \quad (2.5) \]

and

\[ \lim_{n \to \infty} \Pr \{ \tilde{F}_n(t) \in A^*_n(t; \lambda_A), \; -\infty < t \leq T \} = 1 - \alpha ; \quad (2.6) \]

replacing the above four limit relations. The mixed bands are always narrower, though they differ from the originals only in terms of order \( n^{-1} \). As the Associate Editor has pointed out, the asymptotic probabilities are unlikely to be realized in a situation where the two bands being mixed differ noticeably. Reversing this argument, a situation when the two bands based on \( Z_n \) and \( \hat{Z}_n \) do not differ noticeably may be indicative that the finite sample coverage probabilities are close to the asymptotic ones, and then we may perhaps use the slightly narrower mixed bands.

Be it as it may, the notion of a mixed band is at least convenient in that it saves a two-fold asymptotic analysis.

It is interesting to note that, according to the reduction property of Hall & Wellner (1980), the lower contour of the mixed HW-band
\( G_n^*(t; \lambda_H', \lambda_H) \) reduces to the Kolmogorov contour \( F_n - \lambda_H n^{-\frac{1}{2}} \) in the absence of censoring, while the upper contour

\[
G_n^{(2)}(t; \lambda_H', \lambda_H) = \frac{F_n(t)\{1-F_n(t)\} + \lambda_H n^{-\frac{1}{2}}}{1-F_n(t) + \lambda_H n^{-\frac{1}{2}}}
\]

\[
= F_n(t) + \lambda_H n^{-\frac{1}{2}} - \lambda_H^2 n\{1-F_n(t)\} + \lambda_H n^{-\frac{1}{2}}, \quad -1
\]

is strictly below the Kolmogorov upper contour \( F_n + \lambda_H n^{-\frac{1}{2}} \). This observation leads us to a simple improvement of the Kolmogorov band in Appendix 2. Incidentally, the lower and upper contours of the mixed E-band \( G_n^*(t; \lambda_E', 0) \) reduce in the uncensored case to the respective contours of Rényi's (1953) bands.

The width of the mixed AN-band \( A_n^*(t; \lambda_A) \) is

\[
\{1-F_n(t)\}\{2\lambda_A^{-\frac{1}{2}}d_n^*(T) + \lambda_A^{-1}d_n(T)\}/\{1 + \lambda_A^{-\frac{1}{2}}d_n(T)\}.
\]

Hence this band has the property that it narrows as \( t \) grows and is narrowest for large times \( t \) where information is less due to censoring.

The width of the mixed GF-band \( G_n^*(t; \lambda_1', \lambda_2) \) is

\[
\left\{1-F_n(t)\right\} \frac{2n^{-\frac{1}{2}}(\lambda_1 + \lambda_2 d_n(t)) + n^{-1}(\lambda_1 + \lambda_2 d_n(t))^2}{1 + n^{-\frac{1}{2}}(\lambda_1 + \lambda_2 d_n(t))}
\]

which is asymptotically \( 2n^{-\frac{1}{2}}\{1-F_n(t)\}\{\lambda_1 + \lambda_2 d(t)\} = 2n^{-\frac{1}{2}}w(t) \). Now if the density function \( f^0(t) = dF^0(t)/dt \) exists, then

\[
\frac{dw(t)}{dt} = f^0(t)\{\lambda_2 g(t) + \lambda_2 - \lambda_1\}
\]

where
\[ g(t) = \left[1 - F(t)\right]^{-1} - 1 - d(t) \] 

(2.7)

\[ = \int_{-\infty}^{t} \left[1 - F(s)\right]^{-2} d\Pr\{X_1 < s, \delta_1 = 0\} \geq 0. \]

Thus, if \( \lambda_2 \geq \lambda_1 \), including the mixed HW-case, then \( G_n^*(t; \lambda_1, \lambda_2) \) usually widens, for large \( n \), as \( t \) grows. If \( \lambda_2 < \lambda_1 \) then the situation is generally undecided. However, if \( g(T) < (\lambda_1 - \lambda_2)/\lambda_2 \), then \( G_n^*(t; \lambda_1, \lambda_2) \) will narrow as \( t \) grows. In particular, the latter is true for the mixed E-band \( G_n^*(t; \lambda_1, 0) \) for each \( n \), and even without the existence of a density \( f^o \), just as for the mixed AN-band. The \( t \) value near to which \( A_n^*(t; \lambda_A) \) and \( G_n^*(t; \lambda_1, \lambda_2) \) are asymptotically of the same width is the one for which \( \lambda_1 + \lambda_2 d(t) = \lambda_A d^1(t) \). This means that in general we cannot predict the subintervals on which one band would be better than the other, because this depends on a mutual relationship of \( F^o \) and \( H \). Therefore, we may come to an idea of considering some "narrowest possible" bands available, obtained by intersecting two bands, one narrow on the left, the other one narrow on the right, at least when the sample size is reasonably large. We shall mention this possibility at the end of the discussion in Section 5.

We have already pointed out that the qualitative behaviour of \( A_n^*(t; \lambda_A) \) and of the mixed E-band \( E_n^*(t; \lambda_E) = G_n^*(t; \lambda_E, 0) \) is the same. Indeed, \( A_n^*(t; \lambda_A) \) should be viewed as an asymptotically distribution-free version of \( E_n^*(t; \lambda_E) \) which is just a single member of the general family of narrowed mixed GF-bands \( G_n^*(t; \lambda_1, \lambda_2) \). This fact poses the problem if there exists an asymptotically distribution-free version \( M_n^*(t; \lambda_1, \lambda_2) \) of the whole family \( G_n^*(t; \lambda_1, \lambda_2) \), including the E-
HW-special cases, such that $\Lambda_n^*(t; \lambda_A)$ would just be a single member of the family of these versions. In the next section we give an affirmative solution to this problem by a generalised form of the trick of Aalen and Nair, i.e., by building the estimated standard deviation at the endpoint $T$ into the bands.

As noted in the Introduction, the just sketched transformation of $G_n^*$ into $M_n^*$ is one possible way. The other one, based on the empirical Efron transform of $Z_n$ and $\hat{Z}_n$, is described in Section 4.

3. ASYMPTOTICALLY DISTRIBUTION AND CENSOR-FREE GF-BANDS BASED ON THE ESTIMATED VARIANCE AT THE ENDPOINT

In order to motivate the following modified distribution-free bands, note that the relationship between the $\lambda_E$ in the E-bands and the $\lambda_A$ in their distribution-free versions, the AN-bands, is $d^T(T) \lambda_A = \lambda_E$. This can be seen by applying a scale transformation to the Brownian motion in the limiting distribution of the E-bands. We see, therefore, that the distribution-free versions are obtained as if $d(T)$ were estimated by $d_n(T)$ in the limiting distribution of the E-bands. One must, of course, check that this procedure is legal. This is done by considering the processes $Z_n$ and $\hat{Z}_n$ between the boundaries $\pm \lambda_A d_n^T(T)$. The underlying idea in the construction of the following bands is the same: they are obtained as if $d(T)$ in the limiting distribution of the GF-bands were estimated by $d_n(T)$.

Considering, therefore, our two basic processes $Z_n(t)$ and $\hat{Z}_n(t)$ between the boundaries $\pm (\mu_1 d_n^T(T) + \mu_2 (d_n(t)/d_n^T(T)))$ where $\mu_1 > 0$ and $\mu_2 > 0$ are determined in (3.1) below, we are at once led to the
respective general modified GF-bands

\[ M_n(t; \mu_1, \mu_2) = [N_n^{(1)}(t; \mu_1, \mu_2), M_n^{(2)}(t; \mu_1, \mu_2)] \]

\[
= \left[ \frac{F_n^{(1)}(t) - n^{-\frac{1}{2}}(\mu_1 d_n^{1}(T) + \mu_2 \{d_n(t)/d_n^{1}(T)\})}{1 - n^{-\frac{1}{2}}(\mu_1 d_n^{1}(T) + \mu_2 \{d_n(t)/d_n^{1}(T)\})}, \frac{F_n^{(1)}(t) + n^{-\frac{1}{2}}(\mu_1 d_n^{1}(T) + \mu_2 \{d_n(t)/d_n^{1}(T)\})}{1 + n^{-\frac{1}{2}}(\mu_1 d_n^{1}(T) + \mu_2 \{d_n(t)/d_n^{1}(T)\})} \right]
\]

and

\[
\hat{M}_n(t; \mu_1, \mu_2) = [\hat{N}_n^{(1)}(t; \mu_1, \mu_2), \hat{N}_n^{(2)}(t; \mu_1, \mu_2)]
\]

\[
= \left[ \frac{\hat{F}_n^{(1)}(t) - n^{-\frac{1}{2}}(1 - \hat{F}_n^{(1)}(t)) \{\mu_1 d_n^{1}(T) + \mu_2 \frac{d_n(t)}{d_n^{1}(T)}\}, \hat{F}_n^{(1)}(t) + n^{-\frac{1}{2}}(1 - \hat{F}_n^{(1)}(t)) \{\mu_1 d_n^{1}(T) + \mu_2 \frac{d_n(t)}{d_n^{1}(T)}\}} \right]
\]

Since (2.4) again holds with \( G \) replaced by \( M \), whenever the denominator in \( M_n^{(1)}(t; \mu_1, \mu_2) \) is positive, we again consider the narrowed mixed band

\[ M_n^{*}(t; \mu_1, \mu_2) = [\hat{N}_n^{(1)}(t; \mu_1, \mu_2), M_n^{(2)}(t; \mu_1, \mu_2)]. \]

Choosing now \( \mu_1 > 0 \) and \( \mu_2 \geq 0 \) such that

\[ Q_1(\mu_1, \mu_2) = \text{Pr}\{-(\mu_1+\mu_2) t \leq W(t) \leq \mu_1+\mu_2 t, \ 0 \leq t \leq 1\} = 1-\alpha, \quad (3.1) \]

Corollary 6.1 and Lemma 6.2 in Burke et al. (1981), line-10 of page 27 in Csörgő & Horváth (1982a), the continuity of \( d(t) \), and the fact that \( \{W(d(t))/d^1(T), -\infty < t < T\} \) equals in distribution to a Wiener process on [0,1] at once give the following result.

**THEOREM 3.1.** If \( F(T) < 1 \), then as \( n \to \infty \),

\[ \lim_{n \to \infty} \text{Pr}\{F_n^{(1)}(t) \in M_n^{*}(t; \mu_1, \mu_2), \ -\infty < t < T\} = 1-\alpha. \]
The pleasant feature of these bands is that $\mu_1$ and $\mu_2$ do not depend on any quantity other than the level $1-\alpha$. As we have foreseen, the original mixed AN-band is a member, $A_n^*(t; \lambda_A) \equiv M_n^*(t; \lambda_A, 0)$. The distribution-free version of the mixed HW-band is $M_n^*(t; \mu_H^*, \mu_H^*)$, where $\mu_H^*$ is then defined as

$$Q_1(\mu_H^*, \mu_H^*) = \Pr\{ \sup_{0 \leq t \leq 1/2} |B(t)| \leq \mu_H^* \} = 1-\alpha,$$  \hspace{1cm} (3.2)

with a Brownian bridge $B$, again by Doob's transformation.

The asymptotic width of $M_n^*(t; \mu_1, \mu_2)$ is

$$2n^{-1/2} \{1 - F(t)\} \{ \mu_1 d(T) + \mu_2 d(t)/d(T) \}.$$  \hspace{1cm} (3.3)

Hence, just as we saw in the preceding section, if the density function $f^0$ of $F^0$ exists then $M_n^*(t; \mu_1, \mu_2)$ will widen, for large enough $n$, as $t$ grows when $(\mu_2/\mu_1) > d(T)$ and will narrow when $g(T) < (\mu_1 d(T) - \mu_2)/\mu_2$. Here $g$ is as in (2.7). We have already pointed out that $A_n^*(t; \lambda_A) = M_n^*(t; \lambda_A, 0)$ always narrows.

4. ASYMPTOTICALLY DISTRIBUTION AND CENSOR-FREE MIXED GF-TYPE BANDS

BASED ON EMPirical EFRON TRANSFORMS

Working with the notations introduced at the beginning of Section 2, we now introduce the inverse function of the estimated variance function $d_n(t)$ as

$$a_n(t) = \inf \{ u : d_n(u) \geq t \}$$

$$= u_{j:n} \text{ if } d_n(u_{j:n}) < t \leq d_n(u_{j+1:n})$$,
(j=0,...,π_n), where u_0:n = -∞ and d_n(u_{π_n+1:n}) = ∞. This random step function, a generalized quantile function, will be of technical value only, and what we really need is a continuous version of it defined as

\[ \tilde{a}_n(t) = \begin{cases} \frac{1}{t} + \frac{1}{d_n(u_{1:n})} + u_{1:n}, & 0 \leq t \leq d_n(u_{2:n}), \\ \frac{u_{j+1:n} - u_{j:n}}{d_n(u_{j+1:n}) - d_n(u_{j:n})} \{t = d_n(u_{j+1:n})\}, & d_n(u_{j+1:n}) \leq t < d_n(u_{j+2:n}); \\ \frac{t - d_n(u_{π_n:n}) + u_{π_n-1:n}}{π_n - 1}, & t \geq d_n(u_{π_n:n}). \end{cases} \]

This means that we interpolated linearly, and the special definitions on the first and the last intervals serve only to ensure that the range of \( \tilde{a}_n(t) \) be the whole line. If we happen to know that our underlying random variables are nonnegative, as in most applications, then \( \tilde{a}_n(t) \) on \([0,d_n(u_{2:n})]\) is defined by linearly interpolating \( 0 \) and \( u_{1:n} \).

We call \( Z_n^E(t) = Z_n(\tilde{a}_n(t)), t \geq 0 \), and \( \hat{Z}_n^E(t) = \hat{Z}_n(\tilde{a}_n(t)), t \geq 0 \), the empirical Efron transforms of the product limit process, based on \( Z_n \) and \( \hat{Z}_n \) of the Introduction, respectively. In Appendix 1, we shall approximate these transformed processes by a sequence \( \{W_n\} \) of Wiener processes on an interval \([0,L]\), where

\[ L < d(T_{F-}) \leq ∞. \]  

(4.1)

Note that \( d(T_{F-}) \) depends on the heaviness of censoring. Condition (4.1) is parallel to the condition \( T < T_F \), used in all our statements so far.
Let \( \mu_1 > 0 \) and \( \mu_2 \geq 0 \) be defined by (3.1). Noting that
\[
Q_1(\mu_1, \mu_2) = Q_L(\mu_1 L^{-\frac{1}{4}}, \mu_2 L^{-\frac{1}{4}}) = 1 - \alpha
\]
and considering the narrowed mixed GF-band \( G_n^*(t; \mu_1 L^{-\frac{1}{4}}, \mu_2 L^{-\frac{1}{4}}) \) of Section 2, but on the interval \( (-\infty, \tilde{a}_n(L)) \), the theorem in Appendix 1 easily implies the following result.

THEOREM 4.1. If \( L \) is chosen such that condition (4.1) is satisfied, then, as \( n \to \infty \)
\[
\lim_{n \to \infty} \Pr \{ F^O(t) \in G_n^*(t; \mu_1 L^{-\frac{1}{4}}, \mu_2 L^{-\frac{1}{4}}), -\infty < t < \tilde{a}_n(L) \} = 1 - \alpha.
\]
In particular, \( G_n^*(t; \mu_1 L^{-\frac{1}{4}}, 0) \) is, with \( \mu_1 = \lambda_A \) satisfying \( Q_1(\mu_1, 0) = K_1(\mu_1) = 1 - \alpha \), an Aalen-Nair type narrowing band, but with an extra weighting by \( L^{-\frac{1}{4}} \) instead of the estimated standard deviation at the end-point. It is, in fact, the distribution-free version, on the random interval \( (-\infty, \tilde{a}_n(L)) \), of the original mixed E-band \( E_n^*(t; \lambda_E) \)
\[
\equiv G_n^*(t; \lambda_E, 0)
\]
of Section 2. Also, \( G_n^*(t; \mu L^{-\frac{1}{4}}, \mu L^{-\frac{1}{4}}) \), where \( \mu = \mu_H \) satisfies (3.2), is the distribution-free original mixed HW-band, with the same reduction property, on the random interval \( (-\infty, \tilde{a}_n(L)) \). All the qualitative properties such as narrowing, widening, asymptotic bandwidth, discussed earlier remain in force for the present general GF-bands on the data-dependent interval \( (-\infty, \tilde{a}_n(L)) \).

5. DISCUSSION

When applying the bands of Section 2, we must know beforehand that the right end-point \( T \) of the interval \( (-\infty, T) \) on which we are constructing the bands is strictly less than \( T_F \). The necessity of such
a preliminary knowledge seems to be unavoidable with randomly censored
data. We proposed two general classes of confidence bands. The one in
Section 3 builds the estimate \( d_n(T) \) of \( d(T) \) into the bands themselves
and is an extension of the approach of Aalen and Nair. The other one in
Section 4 replaces \( d_n(T) \) by a predetermined quantity \( L \), chosen to
satisfy (4.1), and then the band will be valid on the random interval
\((-\infty, \tilde{a}_n(L))\). Of course, if \( T \) is such that \( d(T) = L \), then the two
approaches will coincide in practice. The first approach in Section 3
is useful when we wish to determine the interval \((-\infty, T]\) in advance,
on which the bands are constructed, but we do not have a preliminary
requirement on the width of the resulting bands. The second approach
in Section 4 is useful when we wish to bound uniformly the width of the
bands in advance. Indeed, let \( 0 < w_0 < 1 \). Then a simple computation
shows that if

\[
L \leq n\left(2+w_0-2(1+w_0)^{1/2}\right)/(\mu_1+\mu_2)^2
\]

then the width of \( G_n^*(t; \mu_1L^{1/2}, \mu_2L^{-1/2}) \) will nowhere be greater than \( w_0 \).
The price we pay is that this will hold on the data-dependent interval
\((-\infty, \tilde{a}_n(L))\).

One last basic question is the required sample size \( n \) at which
the asymptotic results can be applied. The simulation study of Gillespie
& Fisher (1979) cast doubt on the general \textit{ad hoc} applicability for sample
size \( n = 200 \). It follows from our earlier results that the rate of
convergence in (2.5), (2.6) and Theorem 3.1 is at least \( O(n^{-1/3}(\log n)^{3/2}) \),
while in Theorem 4.1 it is at least \( O(n^{-1/4}(\log n)^{3/4}) \), following from
the Theorem in Appendix 1. These are the presently available best
estimates. These rates are of little practical value without knowledge
of the constants in the indicated O's. As a consequence of our earlier papers cited, we in fact have these constants, at least in the first three statements. They are expressed as polynomials of the quantity $b = \{1 - F(T)\}^{-1}$, where $T = a(L)$ in the case of the bands in Section 4. Therefore, it is apparent that the rate of convergence will always depend on the heaviness of censoring and the length of our interval as measured by $b$. Theoretical considerations in Csörgő & Horváth (1982b) suggest that limit theory is applicable at about $n = 10000$ and $b \approx 2$. In the simulation examples of Gillespie & Fisher the values of $b$ are 8.15, 5.44, 3.62 and 2.73. This roughly means that they perhaps work on too long intervals under the given kinds of censorship, and therefore their results are not surprising. We also conducted Monte Carlo experiments, on a smaller scale, reported partly in Csörgő et al. (1983) and in an unpublished Hungarian diploma-work by a student of ours, Pál Pusztai. We considered the cases when exponential distribution is censored by uniform, the case considered by Gillespie & Fisher, when uniform is censored by uniform and exponential, and exponential is censored by exponential. Our general finding in these examples were more optimistic than those, referred to above, based on theoretical considerations. For HW and E-type bands of level $1 - \alpha = 0.9$ we found that when $b \approx 2$, the sample size $n = 500$ is enough for the applicability of the limit theory in these examples. However, much further empirical work would still be needed to acquire a firmer base for more decisive conclusions. For more positive simulation results for the HW and his "equal precision" bands see Nair (1984). We would like to emphasize, at the same time, that in the presence of censoring no theoretical or simulation study
can provide a statement that a certain type of a band "is reliable whenever \( n \geq n_0 \)." These threshold numbers differ from situation to situation and are functions of an intricate relationship of the censored and censoring distributions and of the length of the interval on which the band is constructed.

As an empirical test of the sufficiency of the sample size for the reliable applicability of the asymptotic bands we suggest to visually compare the pair of asymptotically equivalent bands \( M_n(t; \mu_1, \mu_2) \) and \( \hat{M}_n(t; \mu_1, \mu_2) \) for the three choices \( (\lambda_A, 0), (\mu_H, \mu_H) \) and \( (\mu_1, \mu_2), \mu_1 \neq \mu_2, \mu_2 \neq 0 \), of the parameters. The lack of noticeable difference between the members of all three pairs may be a good indication of a reliable situation. We note that if \( \alpha = 0.1 \), then \( \lambda_A = 1.96, \mu_H = 1.13 \), and \( \mu_2 \) corresponding to \( \mu_1 = 1 \) is \( \mu_2 = 0.28 \) (cf. the next section).

We end this discussion with two further remarks. One is that the recent "equal precision" band of Nair (1984) results from considering the process \( \hat{Z}_n \) between the boundaries \( \pm \lambda_n^1(t) \). Another "equal precision" band results from considering the other process \( Z_n \) between these boundaries. Then the corresponding version of (2.4) holds again and one may consider mixing the two bands. The resulting mixed band, however, does not fit into the family \( G_n^* \) of Section 2. It is because this band, having a limiting distribution expressed in terms of the absolute supremum of the standardized Brownian bridge, is valid only on random intervals of the form \( \{ t : A \leq d_n(t) \leq B \} \cap (-\infty, T) \), where \( 0 < A < B \) and \( P(T) < 1 \).

The rate of convergence in this limit theorem is again at least \( O(n^{-1/3} (\log n)^{3/2}) \).
The other remark is that it is possible to intersect two members of the family of bands $M^*_n$ of Section 3, (or the family in Section 4), one narrow on the left, the other narrow on the right, to obtain narrowest possible bands. The arising limiting distribution is quite complicated but still manageable in some cases. Details on such inner envelope bands can be obtained from the authors on request.

6. ILLUSTRATION: THE SZEGED PACEMAKER DATA

Professor Béla Felkai of the Department of Heart Surgery, University Medical School, Szeged, has given us the survival data of 647 patients who have had heart pacemakers implanted in the period January 1969 - July 1982. The data are in months and range from 0 to 163. On the closing date of the data collection 475 patients have been alive and their survival times are considered as censored observations. We have, therefore, 172 uncensored observations. The data are available from the authors on request. Figure 1 depicts bands of Section 3, where we have chosen $T = 150$ on the basis that 151 is the largest uncensored observation. The asymptotic coverage probability is $1 - \alpha = 0.9$. The value $\lambda_A = 1.96$ in (a) is taken from Table 1 of Csörgö and Horváth (1985), $\mu_H = 1.13$ in (b) is taken from Table 1 of Hall & Wellner (1980), while in (c) we have initially fixed $\mu_1 = 1$ and solved the equation $Q_1(1, \mu_2) = 0.9$ on the computer, obtaining $\mu_2 = 0.28$, by using Anderson's (1960) formula for $Q_1$, given also in Gillespie & Fisher (1979).

Figure 1 about here
The relatively large sample size makes us believe that the true coverage probabilities should be close to 0.9 for all the three bands. The modified mixed GF-band used here is uniformly the best for the data at hand. The upper contour of the HW-band in (b) is useless in the far tail. However, since the two bands in (b) and (c) are extremely close to each other up to the 120-th month, any cardiologist would vote that the survival tendency prevails after this date. Therefore we do not see any reason not to trust the band in (c) even up to the 150-th month.

APPENDIX I

EMPIRICAL EFRON TRANSFORMS OF THE PRODUCT-LIMIT PROCESS

For a sequence of random variables $R_n$ and a sequence of constants $r_n$ we write $R_n \overset{a.s.}{=} O(r_n)$ if $\limsup R_n / r_n \leq A$ almost surely with a non-random constant $A$. The following theorem is a complete generalisation of the main result in Csörgő & Horváth (1981).

**Theorem.** If $L$ is as in (4.1), then on some probability space there exists a two-parameter Wiener process $\{W(t,s); t,s \geq 0\}$ such that for the standard Wiener processes $W_n(t) = n^{-\frac{1}{4}} W(t,n)$ we have

$$\sup_{0 \leq t \leq L} |Z_n^E(t) - W_n(t)| \overset{a.s.}{=} O(n^{-1/4} (\log n)^{3/4})$$

and

$$\sup_{0 \leq t \leq L} |Z_n^F(t) - W_n(t)| \overset{a.s.}{=} O(n^{-1/4} (\log n)^{3/4}).$$

**Proof.** The basic probability space will be that of Burke et al.
As a preliminary first step we show that if $T < T_F$, then

$$
\Delta_n^{(1)}(T) = \sup_{0 \leq t < \delta_n(T)} |d(\hat{a}_n(t)) - t| \overset{a.s.}{=} n^{-1/2}(\log n)^{1/2}.
$$

(A.1)

Indeed,

$$
\Delta_n^{(1)}(T) \leq \sup_{0 \leq t < \delta_n(T)} |d(\hat{a}_n(t)) - d(a_n(t))| + \sup_{0 \leq t < \delta_n(T)} |d(a_n(t)) - t|.
$$

(A.2)

By reasons of symmetry, the second term is not greater than

$$
\sup_{0 \leq t < \delta_n(T)} |d_n(a(t)) - t| = \sup_{-\infty < t < T} |d_n(t) - d(t)| \overset{a.s.}{=} O(n^{-1/2}(\log n)^{1/2}).
$$

using Lemma 6.2 in Burke et al. (1981). Since $T < T_F$, for almost all elementary event $\omega$ there exists a threshold $n_0 = n_0(\omega)$ such that $U_{\pi_n} > T$ for $n \geq n_0$. So if $n \geq n_0$, then the first term in (A.2) is not greater than

$$
\max_{1 \leq j \leq \rho_n} \int_{U_{j-1:n}}^{U_{j:n}} (1-F^0)^{-1}(1-H)^{-2}dF^0
$$

$$
\leq (1-F^0(T))^{-1}(1-H(T))^{-2} \max_{1 \leq j \leq \rho_n} \{F^0(U_{j:n}) - F^0(U_{j-1:n})\},
$$

where $\rho_n = \max\{j : 1 \leq j \leq n, U_{j:n} \leq T\}$. Since $\{F^0(U_{j:n})\}$ equally in distribution to the sequence of uncensored ordered elements of a sample drawn from the uniform $(0,1)$ distribution and censored by random variables with distribution function $H(\mathcal{Q}(\cdot))$, where $\mathcal{Q}$ is the quantile
function belonging to \( P^0 \), Theorem 3.1 of Aly, M. Csörgö and Horváth (1985) implies that the latter bound is \( O(n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}) \), almost surely. Hence we have (A.1).

As a second preliminary, we show that

\[
\Delta_n^{(2)}(L) = \sup_{0 < t < L} \left| W_n(d(\tilde{a}_n(t))) - W_n(t) \right| a_n^{1/4} \cdot O(n^{-1/4}(\log n)^{3/4}). \tag{A.3}
\]

Indeed, by condition (4.1) there exists a \( T < T_F \) such that

\[
L < d(T) < \infty. \tag{A.4}
\]

Then, by Lemma 6.2 in Burke et al. (1981), for almost all elementary event \( \omega \) there exists a threshold \( n_1 = n_1(\omega) \) such that \( d_n(T) \geq L \) if \( n \geq n_1 \). Therefore, if \( n \geq n_1 \),

\[
\Delta_n^{(2)}(L) \leq \sup_{0 < t < L} \sup_{0 < h < \Delta_n^{(1)}(T)} \left| W_n(t + h) - W_n(t) \right| a_n^{1/4} \cdot O(n^{-1/4}(\log n)^{3/4})
\]

by (A.1) and an application of Chan's theorem (Theorem 1.14.2 in M. Csörgö & Révész, 1981).

Now we can prove the theorem. We have

\[
\sup_{0 < t < L} \left| Z_n^E(t) - W_n(t) \right| \leq \sup_{0 < t < L} \left| Z_n^E(t) - W_n(d(\tilde{a}_n(t))) \right| + \Delta_n^{(2)}(L),
\]

where the second term is estimated in (A.3). As to the first one, we note first that \( \tilde{a}_n(L) \leq a_n(L) \). Now \( a_n(L) \leq T \) if and only if \( d_n(T) \geq L \), and by (A.4) the latter holds if \( n \geq n_1 \). So if \( n \geq n_1 \) then \( \tilde{a}_n(L) \leq T \) with some \( T \) for which \( F(T) < 1 \). Consequently,
\[ \sup_{0 < t < L} |Z_n^E(t) - W_n(d(\tilde{a}_n(t)))| = \sup_{-\infty < t < \tilde{a}_n(L)} |Z_n(t) - W_n(d(t))| \]
\[ \leq \sup_{-\infty < t < T} |Z_n(t) - W_n(d(t))| \]
\[ a.s. \quad O(n^{-1/3} (\log n)^{5/2}) , \]

where the latter rate is provided by Corollary 6.1 and Section 9 in Burke et al. (1981) and Remark 3.3 in Horváth (1980). The first statement of the theorem is proved, while the second statement follows from the first with little extra work, using, for example, the log log law from Csörgő & Horváth (1983).

APPENDIX 2

NARROWING THE KOLMOGOROV BAND IN THE UNCENSORED CASE

Let \( X_1, \ldots, X_n \) be independent and identically distributed random variables with a continuous distribution function \( F \), and let \( F_n \) denote the left-continuous empirical distribution function of this sample. Determine \( \lambda > 0 \) from the Kolmogorov distribution function:

\[ \Pr\{ \sup_{0 < t < 1} |B(t)| \leq \lambda \} = 1 - \alpha. \]

The usual Kolmogorov band \( [F_n - n^{-1}, F_n + n^{-1}] \) is narrowed by the following result.

THEOREM.

\[ \lim_{n \to \infty} \Pr\{ F_n(t) - \frac{\lambda}{\sqrt{n}} + \frac{\lambda^2}{nF_n(t) + \lambda n^{1/2}} \leq F(t) \leq F_n(t) + \frac{\lambda}{\sqrt{n}} - \frac{\lambda^2}{n(1 - F_n(t)) + \lambda n^{1/2}} , \]
\[ -\infty < t < \infty \} = 1 - \alpha . \]
Proof. Introducing the ordinary empirical process \( B_n(t) = n^{-1} \{ F_n(t) - F(t) \} \), on an appropriate probability space we have the Skorohod construction

\[
\sup_{-\infty < t < \infty} |B_n(t) - B(F(t))| \to 0, \quad \text{almost surely},
\]

by the Doob-Donsker invariance principle. The probability in question is

\[
\Pr\left\{ \left\{ -\lambda \leq B_n(t) \frac{F_n(t)}{F(t)}, \quad -\infty < t < \infty \right\} \cap \left\{ B_n(t) \frac{1-F_n(t)}{1-F(t)} \leq \lambda, \quad -\infty < t < \infty \right\} \right\}
\]

Therefore, the theorem will follow if we show that

\[
\sup_{-\infty < t < \infty} |B_n(t) \frac{F_n(t)}{F(t)} - B(F(t))| + \sup_{-\infty < t < \infty} |B_n(t) \frac{1-F_n(t)}{1-F(t)} - B(F(t))| \to 0
\]

in probability as \( n \to \infty \), on the probability space of the above construction.

Breaking up the two suprema into three parts by appropriately chosen large fixed constants, using the above construction in the middle, and elementary properties of \( B_n \) and the Brownian bridge on the remote half lines in conjunction with the first relation of Remark 1 of Wellner (1978), the latter convergence can be proved in a standard fashion.

We have conducted a small scale Monte Carlo simulation in trying to check the applicability of the Theorem at \( n = 50 \). We generated 40 samples of size 50 from the uniform \((0,1)\) distribution and constructed the bands of the Theorem with \( \lambda = 1.23 \) corresponding to \( 1-\alpha = 0.9 \). Maybe we were too lucky, but \( F(t) = t \) went out of the band on \((0,1)\) only once. It would be desirable to make more extensive empirical studies. Note that even in the middle where \( F_n(t) \approx 1/2 \), the width of our band is less than the width of the Kolmogorov band by the quantity 0.09 when \( 1-\alpha = 0.9 \) and \( n = 50 \). The essential gain is, however, in the tails.
Letting \([S_n^{(1)}(t; \lambda), S_n^{(2)}(t; \lambda)]\) denote the band in the Theorem, note that
\[ S_n^{(1)}(t; \lambda) = 0 \text{ if } t \leq X_{1:n} \text{ and } S_n^{(1)}(t; \lambda) > 0 \text{ if } t > X_{1:n}, \]
while \( S_n^{(2)}(t; \lambda) < 1 \text{ if } t \leq X_{n:n} \text{ and } S_n^{(2)}(t; \lambda) = 1 \text{ if } t > X_{n:n}, \)
where \(X_{1:n}\) and \(X_{n:n}\) are the smallest and the largest observations, respectively. Figure 2 shows typical pictures of the wider Kolmogorov band and our band on simulated uniform \((0,1)\) data at \(n = 50\) and \(n = 100\) with \(\lambda = 1.23\) corresponding to \(1 - \alpha = 0.9\).

---

**Figure 2** about here

---

Finally, we note that all the lower contours of the bands considered in this paper for censored data may now be lifted. Namely, if we add the random functions
\[
\frac{(\lambda_1 + \lambda_2 d_n(t))^2(1 - F_n(t))^2}{nF_n(t) + n \{\lambda_1 + \lambda_2 d_n(t)\}(1 - F_n(t))},
\]
and
\[
\frac{[\mu_1 d_n^+_n(T) + \mu_2 (d_n(t)/d_n^+(T))]^2(1 - F_n(t))^2}{nF_n(t) + n \{\mu_1 d_n^+_n(T) + \mu_2 (d_n(t)/d_n^+(T))\}(1 - F_n(t))},
\]
and
\[
\frac{(\mu_1 L^{-1} + \mu_2 L^{-1} d_n(t))^2(1 - F_n(t))^2}{nF_n(t) + n \{\mu_1 L^{-1} + \mu_2 L^{-1} d_n(t)\}(1 - F_n(t))},
\]
to the lower contours \(\hat{G}_n^{(1)}(t; \lambda_1, \lambda_2), \hat{M}_n^{(1)}(t; \mu_1, \mu_2)\) and \(\hat{G}_n^{(1)}(t; \mu_1 L^{-1}, \mu_2 L^{-1})\)
of the bands \(G_n^*, M_n^*\) and \(G_n^*\) of Sections 2, 3 and 4, respectively, than
the asymptotic theory remains the same. The thus obtained bands reduce in the case \( \lambda_1 = \lambda_2 \) and \( \mu_1 = \mu_2 \) to the band of the Theorem in the absence of censoring.

We wish to thank Professor Béla Fekai for providing the data and Mr. Pál Pusztai for the computer work. We also thank Professor J.N.K. Rao for drawing our attention to Nair (1984) and the referees and the Associate Editor for their many perceptive and helpful remarks.

REFERENCES


Figure 1. (a) The AN-band $A^*_{647}(t; 1.96) = M^*_{647}(t; 1.96, 0)$
(b) The modified HW-band $M^*_{647}(t; 1.13, 1.13)$.
(c) The modified GF-band $M^*_{647}(t; 1, 0.28)$.
The middle curve is the survival estimator $1-\hat{F}_0^{647}(t)$.

Figure 2. The Kolmogorov and narrowed Kolmogorov bands based on simulated uniform (0,1) variables. (a) $n=50$ (b) $n=100$. 