The Annals of Statistics
1987, Vol. 15, No. 1, 360–375

A LARGE SAMPLE STUDY OF THE BAYESIAN BOOTSTRAP[1]

BY ALBERT Y. LO

SUNY at Buffalo

An asymptotic justification of the Bayesian bootstrap is given. Large-sample Bayesian bootstrap probability intervals for the mean, the variance and bands for the distribution, the smoothed density and smoothed rate function are also provided.

1. Introduction. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables having an arbitrary unknown distribution function F. Denote $(X_1, X_2, \ldots, X_n)$ by $\mathbf{X}$. Given a specific functional $\theta(F, \mathbf{X})$, depending on both F and $\mathbf{X}$, the problem is to access the "posterior" opinion of $\theta(F, \mathbf{X})$ given $\mathbf{X} = \mathbf{x}$. A Bayes approach to this problem is to construct a prior distribution on the space of F's and then use the posterior distribution of $\theta(F, \mathbf{X})$ given $\mathbf{X} = \mathbf{x}$ to summarize the "posterior" opinion of $\theta(F, \mathbf{X})$ given $\mathbf{X} = \mathbf{x}$. Recently, Rubin (1981) introduced the Bayesian bootstrap method by constructing a random distribution D_n by replacing the jump-sizes of the empirical distribution function by the gaps of $n - 1$ i.i.d. $U(0,1)$ random variables and suggesting that the conditional distribution of $\theta(D_n, \mathbf{X})|\mathbf{X} = \mathbf{x}$ can be used as the posterior distribution of $\theta(F, \mathbf{X})|\mathbf{X} = \mathbf{x}$.

A great advantage of Rubin's constructive approach is that the Monte Carlo method can be applied to simulate the approximated posterior distribution $\angle\{\theta(D_n, \mathbf{X})|\mathbf{X} = \mathbf{x}\}$ for any given functional θ. Specifically, let $\mathbf{X} = \mathbf{x}$ be a fixed sample.

(1.1) *Step* 1: Simulate $n - 1$ i.i.d. $U(0,1)$ random variables (independent of the X's) and denote the ordered statistics of the U's by $0 = U_{0:\,n-1} < U_{1:\,n-1} < \cdots < U_{n-1:\,n-1} < U_{n:\,n-1} = 1$. Let $\Delta_{j:\,n} = U_{j:\,n-1} - U_{j-1:\,n-1}$, $j = 1, \ldots, n$, be the n gaps of the $U_{j:\,n-1}$'s. Construct a random discrete distribution function D_n with weights $\Delta_{j:\,n}$ at x_j for $j = 1, \ldots, n$.

(1.2) *Step* 2: Repeat Step 1 a large number of times, say B times, to obtain $D_{n1}, D_{n2}, \ldots, D_{nB}$ and compute $\theta(D_{n1}, \mathbf{x}), \ldots, \theta(D_{nB}, \mathbf{x})$, denoted by $\theta_1, \ldots, \theta_B$, respectively.

(1.3) *Step* 3: The empirical distribution function of $\theta_1, \ldots, \theta_B$, putting mass $1/B$ on θ_j for $j = 1, \ldots, B$, approximates $\angle\{\theta(D_n, \mathbf{X})|\mathbf{X} = \mathbf{x}\}$ for a large B.

The convergence of the above algorithm (i.e., Step 3 with B tends to infinity) follows from the law of large numbers since $\theta_1, \ldots, \theta_B$ is an i.i.d. sample from

Received June 1984; revised July 1986.

[1]This research was supported in part by NSF Grant MCS81-02523-01.

AMS 1980 *subject classification*. Primary 62099.

Key words and phrases. Bayesian bootstrap, strong approximation.

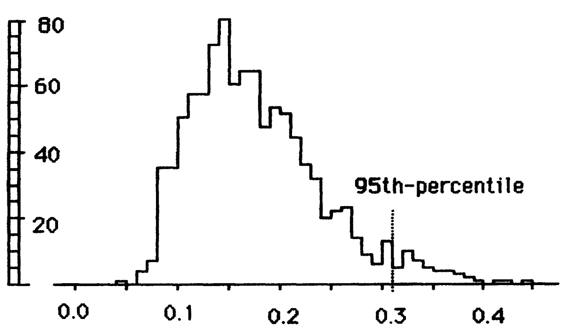

FIG. 1.

$\angle\{\theta(D_n, \mathbf{X})|\mathbf{X} = \mathbf{x}\}$. For the same reason, if the number of Bayesian bootstrap replications is large, $E[\theta(D_n, \mathbf{X})|\mathbf{X} = \mathbf{x}] < \infty$ implies that the sample average of $\theta_1, \ldots, \theta_B$ is near $E[\theta(D_n, \mathbf{X})|\mathbf{X} = \mathbf{x}]$ and $E[\theta^2(D_n, \mathbf{X})|\mathbf{X} = \mathbf{x}] < \infty$ implies that the sample variance of $\theta_1, \ldots, \theta_B$ is close to the conditional variance of $\theta(D_n, X)$ given $\mathbf{X} = \mathbf{x}$, etc.

The following example illustrates the Monte Carlo steps (1.1)–(1.3).

EXAMPLE 1.1. To obtain a 95% probability band for the distribution F, consider the absolute deviation of D_n centered at the empirical distribution function on the reals given by

$$(1.4) \quad \theta_0(D_n, \mathbf{X}) = \sup_x |D_n(x) - F_n(x)| = \sup_{1 \le k \le e} |U_{n(k):\, n-1} - n(k)/n|,$$

where e is the number of distinct observations, $n(k) = \sum_{j:\, 1 \le j \le k} d_j$ for $k = 1, \ldots, e$ and $\{d_1, \ldots, d_e\}$ is the configuration of the ties of the X's (i.e., the numbers of observations tied at the smallest value, the next smallest value, etc.). The result of 1000 BB replications of the above functional for the fifteen GPA's of the law school data in Efron (1982) is displayed in Figure 1. The 95th percentile point of $\theta_0(D_n, \mathbf{X})$ is 0.31 and hence a 95% Bayes probability band of D_n is given by $F_n \pm 0.31$. Note that if $\angle\{\theta_0(D_n, \mathbf{X})|\mathbf{X} = \mathbf{x}\}$ approximates the posterior distribution $\angle\{\theta_0(F, \mathbf{X})|\mathbf{X} = \mathbf{x}\}$, then the above band is approximately a 95% Bayes probability band for the unknown F.

This article is the result of an investigation into the first-order approximation of the conditional distribution $\angle\{\theta(D_n, \mathbf{X})|\mathbf{X}\}$ to the posterior distribution $\angle\{\theta(F, \mathbf{X})|\mathbf{X}\}$ for a variety of functionals. The approximation is to be interpreted in the sense that the distance (say, the distance induced by weak convergence) between $\angle\{\theta(D_n, \mathbf{X})|\mathbf{X}\}$ and $\angle\{\theta(F, \mathbf{X})|\mathbf{X}\}$ tends to zero for almost all $\mathbf{X}$. (A sufficient condition is that they both have the same weak limit for almost all $\mathbf{X}$.) Our study indicates that, if the prior probability is a Dirichlet process [Ferguson (1973)], the approximation is valid. This result, however, is not surprising since Rubin's choice of D_n is related to the work of Ferguson (1973) in the sense that $\angle\{D_n|\mathbf{X}\}$ is a Dirichlet process with shape measure $\sum_i \delta_{X_i}$. Thus Theorem 1 of

Ferguson (1973, Section 3) implies that one may view $\angle\{D_n|\mathbf{X}\}$ as $\angle\{F|\mathbf{X}\}$ where F has a "flat" Dirichlet process prior.

Our study also indicates that asymptotically the Bayesian bootstrap and the bootstrap of Efron (1979) are essentially equivalent in the sense that for almost all sample sequences they achieve the same limiting conditional distribution at the same $n^{1/2}$-rate for a variety of θ's [the finite sample similarities between the BB and the B have already been pointed out by Rubin (1981) and Efron (1982)]. This result raises the question that if a Bayesian can use either the B or the BB as his/her bootstrap, which one is better and in what sense? (Apparently the last question is equally applicable to a frequentist.) We do not know a clear-cut answer to this question.

The results obtained in the following sections relate to the article of Rubin (1981) in the same sense as Bickel and Freedman (1981) and Singh (1981) relate to the paper of Efron (1979). [Note, however, that Singh (1981) also provides some second-order properties of the bootstrap.] In Section 2, we show that, for almost all $\mathbf{X}$, D_n, a posterior Dirichlet process and a bootstrap empirical process F^* (i.e., the chemical distribution of the bootstrap sample), centered and rescaled, can be approximated by a Kiefer process in absolute deviation with rate $O(n^{-1/4}(\log\log n)^{1/4}(\log n)^{1/2})$. This approximation is applied to derive a large-sample BB probability band for F in Section 3. We then discuss the large-sample theory for the Bayesian bootstrapping the mean and the variance in Section 4. In Section 5, the approximation is applied to show that the smoothed BB density approximates a smoothed posterior Dirichlet process density and to construct large-sample probability bands for the smoothed density and rate function. Section 6 collects proofs.

REMARK 1.1. During the first revision, the author was informed by Kjell Doksum that the bootstrap part of Theorem 5.1 in Section 5 was obtained independently by Jhun (1985). During the second revision, the author was again informed by Kjell Doksum that the questions concerning the Bayesian bootstrap in Sections 3 and 4 were discussed independently by Hjort (1985).

2. Strong approximations and the bootstraps. Since the bulk of the functionals $\theta(F, \mathbf{X})$ of interest can be written, at least approximately, as $g(D_n) - g(F_n)$, where g is some appropriate functional, a unified treatment of the asymptotic theory can be achieved by studying the behavior of $d_n = n^{1/2}\{D_n - F_n\}$ for large n's. Our goal is to approximate the d_n-process by Brownian bridges for almost all $\mathbf{X}$. In fact, we will prove the following: One can construct a rescaled Kiefer process which approximates d_n, d^* ($d^* = n^{1/2}\{F^* - F_n\}$) and $d_{\alpha n}$ ($d_{\alpha n} = n^{1/2}\{D_{\alpha n} - F_n\}$) ($F^*$ and $D_{\alpha n}$ are the empirical distribution function of the bootstrap sample and a posterior Dirichlet process with shape measure α, respectively; the precise definitions are given in the next two paragraphs) in supremum distance with rate $O(n^{-1/4}(\log\log n)^{1/4}(\log n)^{1/2})$ for almost all $\mathbf{X}$. An interpretation of this result is that the supremum distance between any two of them is bounded essentially by $n^{-1/4}(\log\log n)^{1/4}(\log n)^{1/2}$ for almost all $\mathbf{X}$; hence their weak limits, conditioned on $\mathbf{X}$, should be identical for almost all $\mathbf{X}$.

The following notation will be used. Let $U_1, U_2, \dots, U_n$ be i.i.d. $U(0,1)$ random variables independent of the X's. Using the first $n-1$ U's, we define D_n described in (1.1) explicitly as follows:

$$(2.1) \qquad D_n(x) = \sum_{j:\, 1 \le j \le n} \Delta_{j:\, n} \delta_{X_{j:\, n}}(x), \qquad -\infty < x < \infty,$$

where δ_x is a point mass at x and the $X_{j:\, n}$'s are the ordered statistics of the sample. Next is some notation for a bootstrap empirical process. The left-continuous inverse of F_n, denoted by F_n^{-1}, is defined by $F_n^{-1}(s) = \inf\{x\colon s \le F_n(x)\}$, $0 \le s \le 1$. For each $n \ge 1$, let $X_{jn}^* = F_n^{-1}(U_j)$ for $j = 1, \dots, n$. It follows that $X_{1n}^*, \dots, X_{nn}^*$ is an i.i.d. sample from F_n and hence is a bootstrap sample. Let F^* be the empirical distribution function of $X_{1n}^*, \dots, X_{nn}^*$. Note that $E[F^* | \mathbf{X}] = E[D_n | \mathbf{X}] = F_n$.

Next, we define a "posterior Dirichlet process" $D_{\alpha n}$. Let $Y_1, \dots, Y_n$ be independent standard exponential random variables and $S_j = Y_1 + \cdots + Y_j$, $j = 1, \dots, n$. Let $\{\mu(x);\ -\infty < x < \infty\}$ be a gamma process with finite shape measure α. That is, $\mu(x)$ is an independent increment process and, for each x, $\mu(x)$ is a gamma $(\alpha(x); 1)$ random variable. We assume that $\{X_i\}$, $\{U_i\}$, $\{Y_i\}$ and $\{\mu(x)\}$ are independent and are defined on a rich enough probability space $(\Omega, \mathscr{F}, P)$ and the elements of Ω are denoted by ω. Define $D_{\alpha n}$ by

$$(2.2) \qquad D_{\alpha n}(x) = [\mu(x) + S_n D_n(x)] / [\mu(\infty) + S_n], \qquad -\infty < x < \infty.$$

Given $X_1, \dots, X_n$, $\mu(x) + S_n D_n(x)$ is a gamma process with shape measure $\alpha + \sum_i \delta_{X_i}$ and hence $D_{\alpha n} | X_1, \dots, X_n$ is a Dirichlet process with shape measure $\alpha + \sum_i \delta_{X_i}$. Therefore, $\angle\{\theta(D_{\alpha n}, \mathbf{X}) | \mathbf{X} = \mathbf{x}\} = \angle\{\theta(F, \mathbf{X}) | \mathbf{X} = \mathbf{x}\}$, where F has a Dirichlet process prior with shape measure α. Denote the centered and rescaled posterior Dirichlet process $n^{1/2}\{D_{\alpha n} - F_n\}$ by $d_{\alpha n}$.

We now present some standard results on a Kiefer process. A Brownian bridge $\{B(s);\ 0 \le s \le 1\}$ is a Gaussian process with a zero mean function and a covariance function defined by $E[B(s_1)B(s_2)] = \min\{s_1, s_2\} - s_1 s_2$. A Kiefer process $\{K(s,t);\ 0 \le s \le 1,\ 0 \le t\}$ is a two-parameter Gaussian process with a zero mean function and a covariance function defined by $E[K(s_1, t_1)K(s_2, t_2)] = (\min\{s_1, s_2\} - s_1 s_2)\min\{t_1, t_2\}$. Note that for each $t > 0$, $\{t^{-1/2}K(s,t); 0 \le s \le 1\}$ is a Brownian bridge and that $\{K(s, n+1) - K(s, n);\ n \ge 1\}$ is a sequence of independent Brownian bridges. For more properties of $B(s)$ and $K(s,t)$, see Csörgő and Révész (1981).

We also denote $\sup_{-\infty < x < \infty}$ by $\sup_x$, $n^{-1/4}(\log\log n)^{1/4}(\log n)^{1/2}$ by $l(n)$, and the sample sequence $X_1, \dots, X_n, \dots$ by $\mathbf{X}$. The symbol $O(\cdot)$ is used in the usual Landau sense. When the corresponding relation holds with probability one, the constant of $O(\cdot)$ could be a finite random variable.

Denote the conditional distribution of $\mathbf{X}$, $\{U_i\}$, $\{Y_i\}$, $\{\mu(x)\}$ and $\{K(s,t)\}$ given $\mathbf{X}$ by $P(\cdot | \mathbf{X})$ and the "true" distribution function by F_0. The main result of this section is the following approximation theorem.

THEOREM 2.1. *There exists a Kiefer process $K(s,t)$ that is independent of* $\mathbf{X}$ *and a F_0-null set N of sample sequences such that* $\mathbf{X} \notin N$ *implies*

$$\sup_x |d_n(x) - n^{-1/2}K(F_0(x), n)| = O(l(n)) \quad a.s.\ [P(\cdot|\mathbf{X})], \tag{2.3}$$

$$\sup_x |d^*(x) - n^{-1/2}K(F_0(x), n)| = O(l(n)) \quad a.s.\ [P(\cdot|\mathbf{X})] \tag{2.4}$$

and

$$\sup_x |d_{\alpha n}(x) - n^{-1/2}K(F_0(x), n)| = O(l(n)) \quad a.s.\ [P(\cdot|\mathbf{X})]. \tag{2.5}$$

The proof of this theorem is given in Section 6. Note that Theorem 2.1 provides the Bayesian and the bootstrap analogies of Finkelstein's LIL law for an empirical distribution function, which is perhaps of interest from a probabilist's viewpoint. Specifically, let

$$\beta_n = d_n/(2\log\log n)^{1/2}, \qquad \beta_{\alpha n} = d_{\alpha n}/(2\log\log n)^{1/2}$$

and

$$\beta^* = d^*/(2\log\log n)^{1/2}.$$

COROLLARY 2.1. *For* $\mathbf{X} \notin N$, *on a set with* $P(\cdot|\mathbf{X})$*-probability one,* β_n, $\beta_{\alpha n}$ *and* β^* *are relatively compact with respect to the supremum distance on* $(-\infty, \infty)$ *with limit points* $\{h(F_0(\cdot));\ h \in K\}$ *where* K *is the Finkelstein set* [*page* 608 *in Finkelstein* (1971)] *containing absolutely continuous* (*with respect to Lebesgue measure*) *functions* h *defined on* $[0,1]$ *such that* $\int [h'(s)]^2\,ds \le 1$ *and* $h(0) = h(1) = 0$.

The proof of this corollary is given in Section 6.

REMARK 2.1. A more careful analysis reveals that the conclusions of Theorem 2.1 and Corollary 2.1 hold on the set of sample sequences $\mathbf{X}$ such that

$$\limsup \sup_x n^{1/2}|F_n(x) - F_0(x)|/(2\log\log n)^{1/2} \le \tfrac{1}{2}. \tag{2.6}$$

3. The Bayesian bootstrap of a distribution function. In this section, we discuss the large-sample effect for the Bayesian bootstrap of the unknown distribution function F and construct a large-sample BB confidence band of F given a sample $X_1, \dots, X_n$ from a continuous F_0. This band is then compared with the exact (up to Monte Carlo accuracy) BB band obtained in Example 1.1. First, we state a direct consequence of Theorem 2.1.

THEOREM 3.1. *For* $X \notin N$,

$$\left|\sup_x |d_n(x)| - n^{-1/2}|K(F_0(x), n)|\right| = O(l(n)) \quad a.s.\ P(\cdot|\mathbf{X}) \tag{3.1}$$

and

(3.2) *statement* (3.1) *remains valid if* d_n *is replaced by* d^* *or* $d_{\alpha n}$.

If F_0 is continuous, $\mathcal{L}\{n^{-1/2}\sup_x|K(F_0(x), n)|\} = \mathcal{L}\{\sup_{0 \le s \le 1}|B(s)|\}$ where B is a standard Brownian bridge. According to (3.1), $F_n \pm n^{-1/2}\lambda$ gives a $(1 - \alpha)$ asymptotic BB probability band for F where λ is defined by

$$\alpha = \sum_{j:\, 1 \le j < \infty} (-1)^{j+1}\exp(-2j^2\lambda^2).$$

For example, $\alpha = 5\%$ corresponds to $\lambda = 1.359$ and if the sample size is $n = 15$, a 95% large-sample BB probability band for F is $F_n \pm 0.35$. This result compares favorably with the exact BB band given in Example 1.1.

Suppose F_0 is arbitrary yet nondegenerate. Given the sample $X_1, \dots, X_n$, there will be distinct X's for n sufficiently large. We have

$$\text{(3.3)} \qquad \begin{aligned} \sup_x |d_n(x)| &= \sup_x n^{1/2}|D_n(x) - F_n(x)| \\ &= n^{1/2} \sup_{1 \le k \le e} \left|U_{n(k):\, n-1} - n(k)/n\right|. \end{aligned}$$

Hence, given $X_1, \dots, X_n$, $\sup_x|d_n(x)|$ is a continuous random variable, implying that a Bayesian bootstrapper can choose a unique c_n (depending on the sample) such that

$$\text{(3.4)} \qquad P\left\{\sup_x |d_n(x)| \le c_n | X_1, \dots, X_n\right\} = 1 - \alpha$$

and a $(1 - \alpha)$ BB probability band for F is given by $F_n \pm n^{-1/2}c_n$. Example 1.1 exemplifies the case of $\alpha = 5\%$ and $n = 15$, obtaining $F_n \pm 0.31$ as the 95% BB band. Note that c_n converges to c, which is the $(1 - \alpha)$-percentile point of the continuous random variable $\sup_x|B(F_0(x))|$ [Bickel and Freedman (1981)] and can be approximated by the $(1 - \alpha)$-percentile point of $\sup_x|B(F_n(x))|$ in view of Lemma 6.2 in Section 6.

On the other hand, given $X_1, \dots, X_n$, $\sup_x|d^*(x)|$ is a discrete random variable supported by at most n points and, according to Bickel and Freedman (1981, Corollary 4.1), a frequentist bootstrapper chooses the sequence $\{c_n\}$ such that

$$\text{(3.5)} \qquad P\left\{\sup_x |d^*(x)| \le c_n | X_1, \dots, X_n\right\} \to 1 - \alpha.$$

Here the choice of c_n is not unique. However, c_n also converges to c.

REMARK 3.1. If F_0 is continuous and supported by a finite interval, the weak convergence arguments in Lo (1983, 1986) can also be modified to provide the asymptotic results in this section.

4. The Bayesian bootstrap of the mean and the variance. In this section we discuss the large-sample effect of the Bayesian bootstrap of some simple functionals of the form $\theta(F, \mathbf{X}) = g(F) - g(F_n)$. We only consider the following simplest linear and nonlinear g's:

$$(4.1)\qquad \begin{aligned} g_1(H) &= \int xH(dx), \\ g_2(H) &= \int\int (x-y)^2 H(dx)H(dy) = 2\operatorname{var}(H). \end{aligned}$$

For these g's, provided that F_0 is supported by a bounded interval, the limit of $\angle\{g(D_n) - g(F_n)|\mathbf{X}\}$ can also be obtained as direct consequences of Theorem 2.1 and an integration-by-parts argument. The situation is more delicate if F_0 has unbounded support, and we give the results in the following theorem. Note that according to Bickel and Freedman (1981) and Singh (1981), a frequentist bootstrapper would obtain the same limiting distributions under the same conditions.

THEOREM 4.1. *$\int x^2 F_0(dx) < \infty$ implies*

$$(4.2)\qquad \angle\{n^{1/2}[g_1(D_n) - g_1(F_n)][\sigma_1(F_0)]^{-1}|\mathbf{X}\} \to N(0,1) \quad a.s.\,[F_0],$$

where $[\sigma_1(F_0)]^2$ is the variance of F_0, and if, in addition, $\int x^2\alpha(dx) < \infty$, (4.2) *remains valid with D_n replaced by $D_{\alpha n}$.*

$\int y^4 F_0(dy) < \infty$ implies

$$(4.3)\qquad \angle\{n^{1/2}[g_2(D_n) - g_2(F_n)][\sigma_2(F_0)]^{-1}|\mathbf{X}\} \to N(0,1) \quad a.s.\,[F_0],$$

where

$$[\sigma_2(F_0)]^2 = 4\left\{\int\left[\int (x-y)^2 F_0(dx)\right]^2 F_0(dy) - [g_2(F_0)]^2\right\},$$

and if, in addition, $\int x^4\alpha(dx) < \infty$, (4.3) *remains valid with D_n replaced by $D_{\alpha n}$.*

PROOF. We only prove (4.3) since the proof of (4.2) is similar and in fact simpler. Note that

$$(4.4)\qquad n^{1/2}\{g_2(D_n) - g_2(F_n)\} = n^{1/2} \sum_{j:\, 1\le j\le n} \Delta_{j:\,n} A_{jn} + n^{1/2}\Delta(D_n, F_n),$$

where $A_{jn} = 2\{\int (X_j - y)^2 F_n(dy) - g_2(F_n)\}$ and

$$(4.5)\qquad \Delta(D_n, F_n) = \int\int (x-y)^2 (D_n - F_n)(dx)(D_n - F_n)(dy).$$

Expand $\Delta^2(D_n, F_n)$ and compute $E[\Delta^2(D_n, F_n)|\mathbf{X}]$ with the joint moments of $\{\Delta_{j:\,n};\ j = 1, \dots, n\}$, which is a Dirichlet random vector. Using the strong consistency of U-statistics, we conclude

$$(4.6)\qquad nE[\Delta^2(D_n, F_n)|\mathbf{X}] = O(n^{-1}) \quad \text{a.s.}\,[F_0].$$

Since $\mathscr{L}\{\Delta_{j:\,n};\ j = 1,\ldots,n\} = \mathscr{L}\{Z_j/(Z_1 + \cdots + Z_n);\ j = 1,\ldots,n\}$, where the Z's are i.i.d. random variables (independent of the X's) having a standard exponential distribution and $(Z_1 + \cdots + Z_n)/n \to 1$ with probability one, it suffices to show

$$\mathscr{L}\Big\{n^{-1/2} \sum_{j:\,1\le j\le n} Z_j A_{jn}|\mathbf{X}\Big\} \to N\big(0, [\sigma_2(F_0)]^2\big) \quad \text{a.s. } [F_0]. \tag{4.7}$$

According to Lindeberg's theorem, (4.7) is valid if for any $\varepsilon > 0$,

$$n^{-1} \sum_{j:\,1\le j\le n} E\Big[(Z_1 A_{jn})^2 I_{jn}(\varepsilon)|\mathbf{X}\Big] \to 0 \quad \text{a.s. } [F_0], \tag{4.8}$$

where $I_{jn}(\varepsilon)$ is the indicator function of the event $\{(Z_1 A_{jn})^2 > n\varepsilon\}$. Since Z_1 is a standard exponential random variable which is independent of $\mathbf{X}$, the left-hand side of (4.8) becomes

$$n^{-1} \sum_{j:\,1\le j\le n} (A_{jn})^2 h(X_j, F_n), \tag{4.9}$$

where

$$h(X_j, F_n) = \Big\{\Big[(\varepsilon n)^{1/2}|A_{jn}|^{-1} + 1\Big]^2 + 1\Big\}\exp\Big\{-(\varepsilon n)^{1/2}|A_{jn}|^{-1}\Big\}.$$

Denote $(X_1 + \cdots + X_n)/n$ by μ_n, $(\mu_n)^2 - g_2(F_n)$ by a_n and expand $(A_{jn})^2$. Expression (4.9) becomes

$$\begin{aligned} &4\Big\{E\big[(X_1)^4 h(X_1, F_n)|\mathscr{S}_n\big] - 4\mu_n E\big[(X_1)^3 h(X_1, F_n)|\mathscr{S}_n\big] \\ &\quad + \big(4(\mu_n)^2 + 2a_n\big)E\big[(X_1)^2 h(X_1, F_n)|\mathscr{S}_n\big] \\ &\quad - 2a_n\mu_n E\big[X_1 h(X_1, F_n)|\mathscr{S}_n\big] + (a_n)^2 E\big[h(X_1, F_n)|\mathscr{S}_n\big]\Big\}, \end{aligned} \tag{4.10}$$

where $\mathscr{S}_n$ is the σ-field generated by the Borel functions of $X_1,\ldots,X_n$, depending symmetrically upon $X_1,\ldots,X_n$. Denote the intersection of the decreasing $\mathscr{S}_n$ by $\mathscr{S}_\infty$.

It suffices to show that the five terms appearing in (4.10) converge to zero a.s. $[F_0]$. Take the first expectation, denoted concisely by $E[V_n|\mathscr{S}_n]$. Note that $V_n \to 0$ a.s. $[F_0]$. Next, $|h(X_1, F_n)| \le 2$ implies that $|V_n|$ is dominated by $2|X_1|^4$ which is integrable. By the generalized martingale convergence theorem of Blackwell and Dubins (1962, Theorem 2),

$$E[V_n|\mathscr{S}_n] \to E\Big[\lim_{n\to\infty} V_n|\mathscr{S}_\infty\Big] = 0 \quad \text{a.s. } [F_0]. \tag{4.11}$$

Similar arguments applied to the other terms complete the proof.

The second part of (4.3) follows from the first part and the fact that $\mathscr{L}\{n^{1/2}[g_2(D_n) - g_2(D_{\alpha n})]|\mathbf{X}\} \to$ a point mass at zero a.s. $[F_0]$. □

Theorem 4.1 is not directly applicable in the construction of large-sample BB probability intervals since the $\sigma(F_0)$'s are not known. The following results bridge this gap.

COROLLARY 4.1. *Theorem* 4.1 *remains valid if* $[\sigma_1(F_0)]^2$ *and* $[\sigma_2(F_0)]^2$ *are replaced by* $[\sigma_1(F_n)]^2$ *and* $[\sigma_2(F_n)]^2$, *respectively.*

PROOF. This result follows from the fact that $[\sigma_1(F_n)]^2 \to [\sigma_1(F_0)]^2$ and $[\sigma_2(F_n)]^2 \to [\sigma_2(F_0)]^2$ a.s. $[F_0]$ and Slutsky's theorem [Bickel and Doksum (1977)]. □

The construction of large-sample BB probability intervals for $g_i(D_n)$, and hence for $g_i(F)$, $i = 1, 2$, is routine and is omitted.

5. The smoothed Bayesian bootstrap. In this section we assume that the "true" distribution F_0 has a density (with respect to Lebesgue measure) given by $f_0(s) = dF_0(s)/ds$. Then the lumpiness of D_n and $D_{\alpha n}$ may not be desirable. One way out is to attribute some smoothness to D_n and $D_{\alpha n}$, and to substitute the smoothed D_n for D_n in Step 1 of the algorithm. We define the smoothed D_n as follows: Let w be a prescribed kernel (density) on $(-\infty, \infty)$ and $b(n) = n^{-\delta}$ for some $\delta \in (0, \frac{1}{2})$. For any distribution function H of the y's, let

$$f_n(s|H) = b(n)^{-1} \int w((s-y)/b(n)) H(dy). \tag{5.1}$$

We call $f_n(s|F)$, $f_n(s|D_n)$, $f_n(s|F^*)$ and $f_n(s|D_{\alpha n})$ the smoothed density, the smoothed BB density, the smoothed B density and the smoothed posterior Dirichlet process density. Note that $\angle\{f_n(s|D_{\alpha n})|\mathbf{X}\} = \angle|\{f_n(s|F)|\,|\mathbf{X}\}$ for a Dirichlet process prior on F, and that $E[f_n(s|D_n)|\mathbf{X}] = E[f_n(s|F^*)|\mathbf{X}] = f_n(s|F_n)$ which is the classical kernel estimate.

In the rest of the section, we discuss the large-sample behavior of $\angle\{f_n(s|D_n)|\mathbf{X}\}$. Specifically, we find the limit of $\angle\{\sup_s|f_n(s|D_n) - f_n(s|F_n)|\,|\mathbf{X}\}$ for almost all $\mathbf{X}$ where the supremum is over a bounded interval. It will be shown that the same limit is also achieved by $\angle\{\sup_s|f_n(s|F^*) - f_n(s|F_n)|\,|\mathbf{X}\}$ and, $\angle\{\sup_s|f_n(s|D_{\alpha n}) - f_n(s|F_n)|\mathbf{X}\}$, implying that for large sample sizes the smoothed BB differs little from the smoothed B and the smoothed posterior Dirichlet process. The basic idea of the proof is due to Bickel and Rosenblatt (1973). Assume that the densities w and f_0 satisfy a set of regularity conditions, say A1–A4 in Bickel and Rosenblatt (1973). Define $Y_n(s)$ by

$$\begin{aligned} Y_n(s) &= [nb(n)/f_0(s)]^{1/2}\{f_n(s|D_n) - f_n(s|F_n)\} \\ &= [b(n)/f_0(s)]^{-1/2} \int w((s-x)/b(n)) d_n(dx), \qquad 0 \le s \le 1. \end{aligned} \tag{5.2}$$

Using Theorem 2.1 instead of Brillinger's theorem [Brillinger (1969)] and arguing as in the proof of Theorem 3.1 of Bickel and Rosenblatt (1973), we obtain the following proposition:

PROPOSITION 5.1. *Let* $g(w) = \int w^2(x)\,dx$ *and* $b(n) = n^{-\delta}$, $0 < \delta < \frac{1}{2}$. *Then, for almost all sample sequences* $\mathbf{X}$,

$$(5.3)\quad P\left\{(2\delta \log n)^{1/2}\left[g(w)^{-1/2} \sup_{0 \le s \le 1} |Y_n(s)| - a_n\right] \le x \middle| \mathbf{X}\right\} \to \exp(-2e^{-x}),$$

where

$$a_n = (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2}\{\log[K_1(w)/\pi^{1/2}] - (\log \delta + \log\log n)/2\},$$

if (a) *of* A1 *holds and* $K_1(w) = (w^2(A) + w^2(-A))/(2g(w)) > 0$, *and otherwise*

$$a_n = (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2}\log[K_2(w)/(2\pi)],$$

where $K_2(w) = \int [w'(x)]^2\,dx/(2g(w))$;

(5.4) *statement* (5.3) *remains valid if* d_n *in* Y_n *is replaced by* d^* *or* $d_{\alpha n}$.

Evidently, a large-sample probability band for $f_n(s|F)$, $0 \le s \le 1$, given $\mathbf{X}$, can be obtained if we replace $f_0(s)$ by $f_n(s|F_n)$ in the denominator of Y_n. The resulting process, say Z_n, obeys

$$(5.5)\quad \begin{aligned} &\sup_{0 \le s \le 1} |Z_n(s) - Y_n(s)| \\ &\quad \le \sup_{0 \le s \le 1} \left|Y_n(s)\{f_n(s|F_n) - f_0(s)\}[f_n(s|F_n)]^{-1}\right| \\ &\quad \le \varepsilon_n(\log n)^{-1/2} \sup_{0 \le s \le 1} |Y_n(s)| \sup_{0 \le s \le 1} \left|[f_n(s|F_n)]^{-1}\right|, \end{aligned}$$

where $\varepsilon_n = (\log n)^{1/2}\sup_{0 \le s \le 1}|f_n(s|F_n) - f_0(s)|$. Our goal is to show that $\sup_{0 \le s \le 1}|Z_n(s) - Y_n(s)| \to 0$ in $P(\cdot|\mathbf{X})$-probability a.s. $[F_0]$. Since $(\log n)^{-1/2}\sup_{0 \le s \le 1}|Y_n(s)|$ is bounded in $P(\cdot|\mathbf{X})$-probability a.s. $[F_0]$ by Proposition 5.1, it suffices to show that ε_n tends to zero a.s. $[F_0]$.

LEMMA 5.1. *In the present context,* ε_n *tends to zero a.s.* $[F_0]$.

The proof of this lemma is given in Section 6.

Similar arguments applied in the bootstrap and the posterior Dirichlet cases with Y_n and Z_n defined analogously give the following theorem.

THEOREM 5.1. *Proposition* 5.1 *remains valid if* Y_n *is replaced by* Z_n.

The above Theorem 5.1 justifies the use of the smoothed BB to construct a probability band for the smoothed density. In addition, according to Theorem 5.1, given $X_1, \dots, X_n$, a $(1 - \alpha)$ large-sample probability band for the smoothed density $f_n(s|F)$ is given by

$$(5.6)\quad f_n(\cdot|F_n) \pm [f_n(\cdot|F_n)g(w)]^{1/2}[nb(n)]^{-1/2}\{(2\delta\log n)^{-1/2}\lambda + a_n\},$$

where λ is defined by $\exp(-2e^{-\lambda}) = 1 - \alpha$.

The construction of large-sample probability band for the corresponding smoothed hazard function is simple. The hazard function r of F is defined by $r(x) = f(x)[1 - F(x)]^{-1}$ and a smoothed BB hazard function is defined by

$$(5.7)\qquad r_n(s|D_n) = f_n(s|D_n)[1 - F_n(s)]^{-1}, \qquad 0 \le s \le 1.$$

Note that r_n is well defined for sufficiently large n. Denote

$$E[r_n(s|D_n)|X_1,\dots,X_n]$$

by $r_n(s|F_n)$, i.e.,

$$(5.8)\qquad r_n(s|F_n) = f_n(s|F_n)[1 - F_n(s)]^{-1}, \qquad 0 \le s \le 1.$$

Since

$$(5.9)\quad [nb(n)(1 - F_n(s))/r_n(s|F_n)]^{1/2}\{r_n(s|D_n) - r_n(s|F_n)\} = Z_n(s),$$

where Z_n is given in Theorem 5.1, we can apply the last theorem to obtain a $(1 - \alpha)$ large-sample probability band for $r_n(s|F)$ given by

$$(5.10)\qquad \begin{aligned} &r_n(\cdot|F_n) \pm [r_n(\cdot|F_n)g(w)]^{1/2}[nb(n)(1 - F_n(\cdot))]^{-1/2} \\ &\qquad \times\{(2\delta\log n)^{-1/2}\lambda + a_n\}. \end{aligned}$$

Apparently, using the bootstrap part of Theorem 5.1, a frequentist bootstrapper obtains the same bands.

REMARK 5.1. The results in this section hold for the bands restricted to a finite interval. They can be extended to cover the case of a sequence of finite intervals with widths tending to infinity at an appropriate rate; see Rice and Rosenblatt (1976) for details.

6. Proofs. We proceed to prove Theorem 2.1. The proof given here is based on the strong approximation method developed by Komlós, Major and Tusnády (1975). Following Brillinger (1969), we first establish a series of lemmas which will imply the main results. The first lemma extends a part of the Chung–Smirnov LIL law [Csörgő and Révész (1981)] from a continuous F_0 to an arbitrary one and is well known. The proof is inserted for completeness. Denote the empirical distribution function of $U_1, U_2, \dots, U_n$ by G_n.

LEMMA 6.1.

$$(6.1)\quad \limsup \sup_x n^{1/2}|F_n(x) - F_0(x)|/(2\log\log n)^{1/2} \le \tfrac{1}{2} \quad a.s.\ [F_0].$$

PROOF. Let F_0^{-1} be the left-continuous inverse of F_0 defined by $F_0^{-1}(s) = \inf\{x;\ s \le F_0(x)\}$, $0 \le s \le 1$. Denote the empirical distribution function of $F_0^{-1}(U_1), \dots, F_0^{-1}(U_n)$, putting mass n^{-1} on each $F_0^{-1}(U_i)$, by H_n. Evidently, $\angle\{H_n;\ n \ge 1\} = \angle\{F_n;\ n \ge 1\}$. Next, $H_n(x) = G_n(F_0(x))$ for each ω and each x implies

$$(6.2)\quad \sup_x |H_n(x) - F_0(x)| = \sup_x |G_n(F_0(x)) - F_0(x)| \le \sup_{0 \le s \le 1} |G_n(s) - s|.$$

An application of the Chung–Smirnov LIL law implies that (6.1) holds if F_n is replaced by H_n and hence the result. □

LEMMA 6.2. *Let $K(s, t)$ be a Kiefer process. Then*

$$(6.3)\qquad \sup_x |K(F_n(x), n) - K(F_0(x), n)| = O(n^{1/2} l(n)) \quad a.s.\ [P].$$

PROOF. This result follows from Lemma 6.1 and Theorem 1.15.2 in Csörgő and Révész (1981), which provides a Hölder condition for a Kiefer process. The arguments are standard [see page 146 in Csörgő and Révész (1981)] and are omitted. □

LEMMA 6.3. *There exists a Kiefer process $K(s, t)$ that is independent of* $\mathbf{X}$ *such that*

$$(6.4)\qquad \sup_x |d_n(x) - n^{-1/2} K(F_n(x), n)| = O(l(n)) \quad a.s.\ [P],$$

$$(6.5)\qquad \sup_x |d^*(x) - n^{-1/2} K(F_n(x), n)| = O(n^{-1/2} \log^2 n) \quad a.s.\ [P]$$

and

$$(6.6)\qquad \sup_x |d_{\alpha n}(x) - n^{-1/2} K(F_n(x), n)| = O(l(n)) \quad a.s.\ [P].$$

PROOF. Define $\alpha_n(s) = n^{1/2}\{G_n(s) - s\}$, $0 \le s \le 1$. The KMT inequality [Theorem 4 in Komlós, Major and Tusnády (1975)] implies the existence of a Kiefer process $K(s, t)$ such that

$$(6.7)\qquad \sup_{0 \le s \le 1} |\alpha_n(s) - n^{-1/2} K(s, n)| = O(n^{-1/2} \log^2 n) \quad \text{a.s. } [P].$$

We prove (6.5) first. Note that for this $K(s, t)$,

$$(6.8)\qquad \begin{aligned} &\sup_x |d^*(x) - n^{-1/2} K(F_n(x), n)| \\ &\quad = \sup_x |n^{1/2}\{G_n(F_n(x)) - F_n(x)\} - n^{-1/2} K(F_n(x), n)| \\ &\quad \le \sup_{0 \le s \le 1} |n^{1/2}\{G_n(s) - s\} - n^{-1/2} K(s, n)| \\ &\quad = O(n^{-1/2} \log^2 n) \quad \text{a.s. } [P], \end{aligned}$$

where the first equality follows from $F^*(x) = G_n(F_n(x))$ for each ω and each x

and the last equality is by (6.7). Next, note that

$$\begin{aligned}
&\sup_x \left|d_n(x) - n^{-1/2}K(F_n(x), n)\right| \\
&\quad \le \sup_{1\le j\le n} \left|n^{1/2}\{U_{j:\,n-1} - j/n\} - n^{1/2}K(j/n, n)\right| \\
&\quad \le \sup_{1\le j\le n} \left|n^{1/2}\{U_{j:\,n-1} - j/(n-1)\} - n^{-1/2}K(j/n, n)\right| \\
&\qquad + (n-1)^{-1}n^{1/2} \\
&\quad \le \sup_{1\le j\le n-1} n^{1/2}\Big|\{U_{j:\,n-1} - j/(n-1)\} \\
&\qquad\qquad - (n-1)^{-1}K(j/(n-1), n-1)\Big| \\
&\qquad + \sup_{1\le j\le n-1} \left|(n-1)^{-1}n^{1/2}K(j/(n-1), n-1) - n^{-1/2}K(j/n, n)\right| \\
&\qquad + (n-1)^{-1}n^{1/2} \\
&\quad = A_n + B_n + (n-1)^{-1}n^{1/2}.
\end{aligned} \tag{6.9}$$

It remains to bound A_n and B_n appropriately.

$$\begin{aligned}
A_n &\le 2 \sup_{1\le j\le n-1} \left|\alpha_{n-1}(U_{j:\,n-1}) - (n-1)^{-1/2}K(U_{j:\,n-1}, n-1)\right| \\
&\quad + 2(n-1)^{-1/2} \sup_{1\le j\le n-1} \left|K(U_{j:\,n-1}, n-1) - K\left(\frac{j}{n-1}, n-1\right)\right| \\
&\le 2 \sup_{0\le s\le 1} \left|\alpha_{n-1}(s) - (n-1)^{-1/2}K(s, n-1)\right| \\
&\quad + 2(n-1)^{-1/2} \sup_{0\le s\le 1} \left|K(s, n-1) - K(G_{n-1}(s), n-1)\right| \\
&= O(l(n)) \quad \text{a.s. } [P],
\end{aligned} \tag{6.10}$$

where the last equality follows from (6.7) and Lemma 6.2.

Next we bound B_n.

$$\begin{aligned}
B_n &\le 2(n-1)^{-1/2} \sup_{1\le j\le n-1} \left|K(j/(n-1), n-1) - K(j/n, n-1)\right| \\
&\quad + 2(n-1)^{-1/2} \sup_{1\le j\le n-1} \left|K(j/n, n) - K(j/n, n-1)\right| \\
&\quad + (n-1)^{-1}n^{-1/2} \sup_{1\le j\le n-1} \left|K(j/n, n)\right| \\
&= B_{1n} + B_{2n} + B_{3n}.
\end{aligned} \tag{6.11}$$

First, $B_{3n} \le (n-1)^{-1}n^{-1/2}\sup_{0\le s\le 1}|K(s, n)| = O(l(n))$ a.s. $[P]$ by the LIL law for a Kiefer process. Next,

$$B_{2n} \le 2(n-1)^{-1/2} \sup_{0\le s\le 1} |K(s,n) - K(s,n-1)| = O(l(n)) \quad \text{a.s. } [P].$$

Finally, $B_{1n} = O(l(n))$ a.s. $[P]$ by the Hölder condition [Theorem 1.15.2 in Csörgő and Révész (1981)] for a Kiefer process.

Statement (6.6) follows from (6.4) since

$$\sup_x |d_{\alpha n}(x) - d_n(x)| \le [\mu(\infty)n^{1/2}]/[\mu(\infty) + S_n]. \tag{6.12}$$

It remains to note that the $K(s,t)$ process is defined on the sequence $\{U_i\}$ and hence can be chosen to be independent of $\mathbf{X}$ if $(\Omega, \mathscr{F})$ is rich enough. □

LEMMA 6.4. *There exists a Kiefer process $K(s,t)$ that is independent of* $\mathbf{X}$ *such that*

$$\sup_x |d_n(x) - n^{-1/2}K(F_0(x), n)| = O(l(n)) \quad a.s. [P] \tag{6.13}$$

and

(6.14) (6.13) *remains valid if d_n is replaced by d^* or $d_{\alpha n}$.*

PROOF. Note that

$$\begin{aligned}\sup_x &|d_n(x) - n^{-1/2}K(F_0(x), n)| \\ &\le \sup_x |d_n(x) - n^{-1/2}K(F_n(x), n)| \\ &\quad + n^{-1/2}\sup_x |K(F_n(x), n) - K(F_0(x), n)|.\end{aligned}$$

Lemma 6.2 and (6.4) of Lemma 6.3 imply (6.13). Similarly, Lemma 6.2 and (6.5) and (6.6) of Lemma 6.3 give (6.14) □

PROOF OF THEOREM 2.1. Apply Fubini's theorem to Lemma 6.4. □

PROOF OF COROLLARY 2.1. Let $\gamma_n = n^{1/2}\{F_n - F_0\}$. Finkelstein's theorem [as extended by Richter (1973) for an arbitrary F_0] states that with F_0-probability one $\{\gamma_n;\ n \ge 1\}$ is relatively compact with respect to the supremum distance on $(-\infty, \infty)$ with limit points $\{h(F_0(\cdot));\ h \in K\}$. By the extended KMT theorem [Remark 4.4.3 in Csörgő and Révész (1981)],

$$\sup_x |\gamma_n(x) - n^{-1/2}K(F_0(x), n)| = O(n^{-1/2}\log^2 n) \quad \text{a.s. } [F_0], \tag{6.15}$$

for some Kiefer process $K(s,t)$, implying that $\{K(F_0, n)/(2n\log\log n)^{1/2};\ n \ge 1\}$ is almost surely relatively compact with respect to the supremum distance with limit points $\{h(F_0(\cdot));\ h \in K\}$. The conclusion then follows from Theorem 2.1. □

PROOF OF LEMMA 5.1. Since $(\log n)^{1/2}\sup_{0 \le s \le 1}|E[f_n(s|F_n)] - f_0(s)| \to 0$ (this result follows immediately from elementary analysis), it suffices to show that $(\log n)^{1/2}\sup_{0 \le s \le 1}|f_n(s|F_n) - E[f_n(s|F_n)]| \to 0$ a.s. $[F_0]$.

Denote $n^{-1/2}\log^2 n$ by $r(n)$. Let $K(s, t)$ be the Kiefer process given by Theorem 4 in Komlós, Major and Tusnády (1975). The integration by parts and the assumption A1 [in Bickel and Rosenblatt (1973)] can be applied to give

$$\sup_{0 \le s \le 1}\Big| f_n(s|F_n) - E[f_n(s|F_n)] - [b(n)n]^{-1}\int w((s-x)/b(n))\,d\{K(F_0(x), n)\}\Big| = O\big(r(n)b(n)^{-1}n^{-1/2}\big) \quad \text{a.s. } [F_0]. \tag{6.16}$$

It suffices to show that a.s. $[F_0]$

$$L_n = (\log n)^{1/2}[b(n)n]^{-1}\sup_{0 \le s \le 1}\left|\int w\left(\frac{s-x}{b(n)}\right)d\{K(F_0(x), n)\}\right| \to 0. \tag{6.17}$$

However, the last statement follows from

$$L_n \le (\log n)^{1/2}[b(n)n]^{-1}\sup_{0 \le s \le 1}|K(s, n)|\int |w'(y)|\,dy \to 0 \quad \text{a.s. } [F_0], \tag{6.18}$$

by the law of the iterated logarithm for a Kiefer process if $0 < \delta < \frac{1}{2}$. □

Acknowledgments. The author wishes to thank Michael Woodroofe for numerous stimulating conversations and Miklós Csörgő for a private communication. The revisions benefited from helpful suggestions from Kjell Doksum, Donald Rubin, Wing Hung Wong and two referees.

REFERENCES

BICKEL, P. J. and DOKSUM, K. A. (1977). *Mathematical Statistics: Basic Ideas and Selected Topics.* Holden Day, San Francisco.

BICKEL, P. J. and FREEDMAN, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217.

BICKEL, P. J. and ROSENBLATT, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095.

BLACKWELL, D. and DUBINS, L. (1962). Merging of opinion with increasing information. *Ann. Math. Statist.* **33** 882–886.

BRILLINGER, D. R. (1969). The asymptotic representation of the sample distribution function. *Bull. Amer. Math. Soc.* **75** 545–547.

CSÖRGŐ, M. and RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics.* Academic, New York.

EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26.

EFRON, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans.* SIAM, Philadelphia.

FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.

FINKELSTEIN, H. (1971). The law of the iterated logarithm for empirical distributions. *Ann. Math. Statist.* **42** 607–615.

HJORT, N. L. (1985). Bayesian nonparametric bootstrap confidence intervals. Technical Report 240, Dept. of Statistics, Stanford Univ.

JHUN, M. S. (1985). Ph.D. thesis, Dept. of Statistics, Univ. of California, Berkeley.

KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent rv's and the sample df, I. *Z. Wahrsch. verw. Gebiete* **32** 111–131.

LO, A. Y. (1983). Weak convergence for Dirichlet processes. *Sankhyā Ser. A* **45** 105–111.

LO, A. Y. (1986). A remark on the limiting posterior distribution of the multiparameter Dirichlet process. *Sankhyā Ser. A* **48** 247–249.

RICE, J. and ROSENBLATT, M. (1976). Estimation of the log survivor function and hazard function. *Sankhyā Ser. A* **38** 60–78.

RICHTER, H. (1973). Zum Gesetz vom iterierten Logarithmus fur empirische Verteilungsfunktionen und empirisches Chi-quadrat. *Manuscripta Math.* **9** 187–199.

RUBIN, D. B. (1981). The Bayesian bootstrap. *Ann. Statist.* **9** 130–134.

SINGH, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195.

DEPARTMENT OF STATISTICS
STATE UNIVERSITY OF NEW YORK
BUFFALO, NEW YORK 14261