The Annals of Statistics
1992, Vol. 20, No. 3, 1611–1624

A RANK STATISTICS APPROACH TO THE CONSISTENCY OF A GENERAL BOOTSTRAP

BY DAVID M. MASON[1] AND MICHAEL A. NEWTON[2]

University of Delaware and University of Washington

A general notion of a bootstrapped mean constructed by exchangeably weighting sample points is introduced. Consistency of this generalized bootstrapped mean, which includes proposals of Efron and Rubin among others, is proved by classical linear rank statistics theory. The consistency of generalized bootstrapped empirical and quantile processes is also established.

1. Introduction. We introduce a generalized bootstrap procedure and study some of its properties by means of rank statistics methods. To motivate our generalized bootstrap, let $X_1, X_2, \dots, X_n$ be i.i.d. F. Efron's (1979) bootstrapped mean can be written as

$$\overline{X}_{\mathscr{M},n} = \sum_{i=1}^{n} M_{n,i} X_i, \tag{1.1}$$

where $M_n := (M_{n,1}, M_{n,2}, \dots, M_{n,n})$ is $1/n$ times a multinomial random vector formed from n draws on n equally likely cells, and independent of $X_1, \dots, X_n$. Notationally,

$$nM_n \sim \mathrm{Mult}(n; 1/n, 1/n, \dots, 1/n).$$

Analogously, Rubin's (1981) Bayesian bootstrapped mean can be written as

$$\overline{X}_{\mathscr{D},n} = \sum_{i=1}^{n} D_{n,i} X_i, \tag{1.2}$$

where $D_n := (D_{n,1}, D_{n,2}, \dots, D_{n,n})$ is equal in distribution to the vector of n 1 spacings of $n-1$ ordered uniform $(0,1)$ random variables independent of $X_1, \dots, X_n$. That is,

$$D_n \sim \mathrm{Dirichlet}(n; 1, 1, \dots, 1).$$

[If instead, D_n is chosen to be Dirichlet $(n; 4, 4, \dots, 4)$, then $\overline{X}_{\mathscr{D},n}$ corresponds to a suggestion of Weng (1989), Remark 2.3, and Zheng and Tu (1988), Remark 5.]

A natural generalization of these two forms of the bootstrapped mean is the following. Consider a vector of random weights $W_n = (W_{n,1}, W_{n,2}, \dots, W_{n,n})$

Received May 1990; revised April 1991.

[1]Research partially supported by NSF Grant DMS-88-03209.

[2]Research partially supported by ONR Contract N-00014-88-K-0265, and the Applied Physics Laboratory at the University of Washington.

AMS 1980 *subject classifications*. Primary 62G99; secondary 62G30.

Key words and phrases. Bootstrap, central limit theorem, exchangeable, rank statistics.

independent of the data $X_1, \dots, X_n$. Assume that for each integer $n \ge 1$, the components of W_n are exchangeable. Now form the generalized bootstrapped mean corresponding to the weight vector W_n:

$$\bar{X}_{\mathscr{W},n} = \sum_{i=1}^{n} W_{n,i} X_i. \tag{1.3}$$

Note that $\bar{X}_{\mathscr{M},n}$ and $\bar{X}_{\mathscr{D},n}$ are special cases of $\bar{X}_{\mathscr{W},n}$. Typically, the weights W_n will also satisfy

$$(\mathscr{W}_{\mathrm{I}}) \qquad W_{n,i} \ge 0, \qquad i = 1, 2, \dots, n, \, n \ge 1,$$

$$(\mathscr{W}_{\mathrm{II}}) \qquad \sum_{i=1}^{n} W_{n,i} = 1,$$

and for some $c > 0$,

$$(\mathscr{W}_{\mathrm{III}}) \qquad n \sum_{i=1}^{n} (W_{n,i} - 1/n)^2 \to_P c \quad \text{as } n \to \infty.$$

Let $\bar{X}_n$ denote the sample mean of $X_1, \dots, X_n$ and S_n^2, the sample variance, $\sum_{i=1}^{n}(X_i - \bar{X}_n)^2/n$. Assuming that $0 < \operatorname{Var} X := \sigma^2 < \infty$, it is known in the two special cases (1.1) and (1.2), that uniformly in t,

$$P\Big(\sqrt{n}\left(\bar{X}_{\mathscr{W},n} - \bar{X}_n\right)/S_n \le t \Big| X_1, \dots, X_n\Big) \to \Phi(t) \quad \text{a.s.,} \tag{1.4}$$

as $n \to \infty$, where Φ is the cumulative distribution function of the standard normal random variable. A bootstrap procedure satisfying (1.4) is said to be consistent. See Bickel and Freedman (1981) for the case $\mathscr{W} = \mathscr{M}$ and Lo (1987) for $\mathscr{W} = \mathscr{D}$. For Edgeworth expansions of the conditional distribution in (1.4), refer to Singh (1981) when $\mathscr{W} = \mathscr{M}$ and Weng (1989) when $\mathscr{W} = \mathscr{D}$.

In Section 2, we prove a theorem which provides conditions on the weights W_n under which (1.4) is true. In particular, this yields (1.4) for $\mathscr{W} = \mathscr{M}$ and $\mathscr{W} = \mathscr{D}$ as special cases. Our approach relies heavily on a classical result due to Hájek (1961) on the asymptotic normality of linear rank statistics. We remark in passing that a natural extension of the Singh (1981) and Weng (1989) results is feasible using Edgeworth expansions for linear rank statistics as in Does (1983) and Schneller (1989). In Section 3, we consider generalized versions of bootstrapped empirical and quantile processes. Our treatment of these processes is indirectly based on rank statistics methods through our use of Theorem 24.2 of Billingsley (1968).

Zheng and Tu in a series of papers [see Zheng and Tu (1988) and the references therein] have studied and established the consistency of a method of random weighting which is essentially a special case of (1.3). The novelty of our point of view is that classical linear rank statistics theory provides a ready method for verifying the consistency of a variety of bootstrap procedures.

2. Consistency of the generalized bootstrap mean. We begin with a theorem on the asymptotic normality of linear combinations of exchangeable arrays. This theorem is an extension of Theorem 1 of Chernoff and Teicher

(1958) and relies on Theorem 4.1 of Hájek (1961). It is our basic tool for establishing the consistency of the generalized bootstrapped mean.

Let $\{X_{n,k}\colon k = 1, 2, \dots, k_n,\ n \geq 1\}$ be an array of real-valued random variables sitting on a probability space $(\Omega_1, \mathscr{A}_1, P_1)$. Let $\{Y_{n,k}\colon k = 1, 2, \dots, k_n,\ n \geq 1\}$ be another array of real-valued random variables independent of the $X_{n,k}$ array, and sitting on a probability space $(\Omega_2, \mathscr{A}_2, P_2)$. Suppose that $k_n \to \infty$ as $n \to \infty$ and for each n, the variables $Y_{n,1}, Y_{n,2}, \dots, Y_{n,k}$ are exchangeable. Let $\overline{X}_n$ and $\overline{Y}_n$ be the row averages of these arrays. Define the following random variables:

$$Z_n = \frac{\sqrt{k_n}\,\Sigma_k X_{n,k}\big(Y_{n,k} - \overline{Y}_n\big)}{\Big(\Sigma_i\big(X_{n,i} - \overline{X}_n\big)^2 \Sigma_j\big(Y_{n,j} - \overline{Y}_n\big)^2\Big)^{1/2}}, \tag{2.1}$$

$$U_{n,k} = \frac{\big(X_{n,k} - \overline{X}_n\big)}{\Big(\Sigma_i\big(X_{n,i} - \overline{X}_n\big)^2\Big)^{1/2}}, \tag{2.2}$$

$$V_{n,k} = \frac{\big(Y_{n,k} - \overline{Y}_n\big)}{\Big(\Sigma_i\big(Y_{n,i} - \overline{Y}_n\big)^2\Big)^{1/2}}. \tag{2.3}$$

(We define $0/0 = 1$.) Note that Z_n are governed by the product measure $P_1 \times P_2$ while $U_{n,k}$ and $V_{n,k}$ are governed by the respective marginal measures P_1 and P_2. Each $\omega_1 \in \Omega_1$ defines two infinite arrays of constants $x_{n,k} = X_{n,k}(\omega_1)$, $u_{n,k} = U_{n,k}(\omega_1)$, and similarly each $\omega_2 \in \Omega_2$ defines arrays $y_{n,k} = Y_{n,k}(\omega_2)$, $v_{n,k} = V_{n,k}(\omega_2)$.

We have the following main theorem.

THEOREM 2.1. *Suppose that as $n \to \infty$,*

$$\max_{1 \leq k \leq k_n} U_{n,k}^2 \to 0 \quad a.s.\ P_1, \tag{2.4}$$

$$\max_{1 \leq k \leq k_n} V_{n,k}^2 \to_{P_2} 0, \tag{2.5}$$

and for almost every $\omega_1 \in \Omega_1$ and all $\tau > 0$,

$$D_n(\tau) = \sum_i \sum_j u_{n,i}^2 V_{n,j}^2 1\big[k_n u_{n,i}^2 V_{n,j}^2 > \tau\big] \to_{P_2} 0. \tag{2.6}$$

Then for almost every $\omega_1 \in \Omega_1$ and uniformly in t,

$$P\big(Z_n \leq t | X_{n,k}(\omega_1) = x_{n,k},\ k = 1, 2, \dots, k_n\big) \to \Phi(t), \tag{2.7}$$

where $\Phi(t)$ is the standard normal cumulative distribution function.

PROOF. Choose a point $\omega_1 \in \Omega_1$ for which $\max_{1 \leq k \leq k_n} U_{n,k}^2(\omega_1) \to 0$ and $D_n(\tau) \to 0$ in P_2 probability. The proof relies on the introduction of ranks as follows. Let $(R_{n,1}, R_{n,2}, \dots, R_{n,k_n})$ be a random vector taking each permutation of $(1, 2, \dots, k_n)$ with equal probability, and independent of the $X_{n,k}$ and $Y_{n,k}$. Construct a permuted version of Z_n as

$$Z_n^* = \frac{\sqrt{k_n}\,\Sigma_k X_{n,k}\left(Y_{n,R_{n,k}} - \bar{Y}_n\right)}{\left(\Sigma_i\left(X_{n,i} - \bar{X}_n\right)^2 \Sigma_j\left(Y_{n,j} - \bar{Y}_n\right)^2\right)^{1/2}}.$$

It is a simple consequence of the rowwise exchangeability of $Y_{n,k}$ that Z_n^* has the same distribution as Z_n. Thus it suffices to prove (2.7) with Z_n replaced by Z_n^*.

Let $x_{n,k}$ and $y_{n,k}$ be two arrays of real numbers (dimensioned as $X_{n,k}$) with row averages $\bar{x}_n$ and $\bar{y}_n$, and define $u_{n,k}$ and $v_{n,k}$ as in (2.2) and (2.3)

$$u_{n,k} = \frac{(x_{n,k} - \bar{x}_n)}{\left(\Sigma_j\left(x_{n,j} - \bar{x}_n\right)^2\right)^{1/2}}, \qquad v_{n,k} = \frac{(y_{n,k} - \bar{y}_n)}{\left(\Sigma_j\left(y_{n,j} - \bar{y}_n\right)^2\right)^{1/2}}$$

and set for $\tau > 0$,

$$d_n(\tau) = \sum_{i,j} u_{n,i}^2 v_{n,j}^2 1\left[k_n u_{n,i}^2 v_{n,j}^2 > \tau\right].$$

Introduce the linear rank statistic

$$T_n = \frac{\sqrt{k_n}\,\Sigma_k x_{n,k}\left(y_{n,R_{n,k}} - \bar{y}_n\right)}{\left(\Sigma_i\left(x_{n,i} - \bar{x}_n\right)^2 \Sigma_j\left(y_{n,j} - \bar{y}_n\right)^2\right)^{1/2}}.$$

Note that the randomness in T_n comes only through the ranks $(R_{n,1}, \dots, R_{n,k_n})$.

Hájek's Theorem 4.1 [actually its proof combined with that of the Lindeberg–Feller theorem as given in Billingsley (1968)] says that for every $\varepsilon > 0$ there exist $\delta > 0$ and $\tau > 0$ such that whenever

$$\max_{1 \le k \le k_n} u_{n,k}^2 < \delta, \qquad \max_{1 \le k \le k_n} v_{n,k}^2 < \delta, \qquad d_n(\tau) < \delta, \tag{2.8}$$

then uniformly in t,

$$\left|P(T_n \le t) - \Phi(t)\right| < \varepsilon/2.$$

Treating $x_{n,k}$ and $y_{n,k}$ as realizations $X_{n,k}(\omega_1)$ and $Y_{n,k}(\omega_1)$ and noting that given these realizations $T_n = Z_n^*$, we get immediately when (2.8) holds that uniformly in t,

$$\left|P(Z_n^* \le t | X_{n,k} = x_{n,k}, Y_{n,k} = y_{n,k}, k = 1, \dots, k_n) - \Phi(t)\right| < \varepsilon/2. \tag{2.9}$$

In what follows, we use an obvious abbreviated notation for the variables being conditioned on.

Almost sure convergence of $\max_{1 \le k \le k_n} U_{n,k}^2$ to 0 means that for every $\delta > 0$ and almost every ω_1 there exists $N_1 = N_1(\omega_1, \delta)$ such that for all $n > N_1$, $\max_{1 \le k \le k_n} u_{n,k}^2 < \delta$. Let

$$B_n = \left\{\omega_2 \in \Omega_2 \colon \max_{1 \le k \le k_n} v_{n,k}^2 \ge \delta\right\}$$

and for any $\tau > 0$ set

$$C_n(\tau) = \{\omega_2 \in \Omega_2 \colon d_n(\tau) \ge \delta\}.$$

By assumption, there exist $N_2 = N_2(\delta, \varepsilon)$ and $N_3 = N_3(\omega_1, \delta, \varepsilon, \tau)$ such that for all $n > N_2$, $P_2(B_n) < \varepsilon/4$ and all $n > N_3$, $P_2(C_n(\tau)) < \varepsilon/4$. Combining these facts with (2.8), we have for $n > \max\{N_1, N_2, N_3\}$, uniformly in t,

$$\begin{aligned}
&|P(Z_n^* \le t | X_{n,k}(\omega_1)) - \Phi(t)| \\
&\quad = \left| \int [P(Z_n^* \le t | X_{n,k}(\omega_1), Y_{n,k}(\omega_2)) - \Phi(t)]\, dP_2(\omega_2) \right| \\
&\quad \le \int_{B_n^c \cap C_n^c(\tau)} |P(Z_n^* \le t | X_{n,k}(\omega_1), Y_{n,k}(\omega_2)) - \Phi(t)|\, dP_2(\omega_2) \\
&\qquad + \int_{B_n \cup C_n(\tau)} dP_2(\omega_2) \\
&\quad \le (\varepsilon/2) P_2(B_n^c \cap C_n^c(\tau)) + P_2(B_n) + P_2(C_n(\tau)) \\
&\quad < \varepsilon/2 + \varepsilon/4 + \varepsilon/4 = \varepsilon.
\end{aligned}$$

By the arbitrary choice of $\varepsilon > 0$, the proof is complete. □

We remark that the proof of our Theorem 2.1 was based on ideas in the proof of Theorem 1 of Chernoff and Teicher (1958).

The following two corollaries provide useful sets of sufficient conditions for the assumptions of Theorem 2.1 to hold.

COROLLARY 2.2. *Suppose* $X_{n,k} = X_k$, $k = 1, 2, \dots, k_n$, *where* $X_1, X_2, \dots$ *are i.i.d.* F *with* $0 < \operatorname{Var} X = \sigma^2 < \infty$, *and also assume that* (2.5) *holds. Then* (2.7) *follows.*

PROOF. It is well known that $EX^2 < \infty$ if and only if

$$\max_{1 \le j \le m} |X_j| / \sqrt{m} \to 0 \quad \text{a.s. } P_1 \text{ as } m \to \infty. \tag{2.10}$$

Also we have

$$S_m^2 \to \sigma^2 \quad \text{a.s. } P_1 \text{ as } m \to \infty. \tag{2.11}$$

From these two facts, we easily infer (2.4). By (2.5), we get for any $\tau > 0$ and $0 < \varepsilon < 1$ that for all n sufficiently large with P_2 probability greater than $1 - \varepsilon$,

$$D_n(\tau) \le \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} u_{n,i}^2 V_{n,j}^2 1\left[k_n u_{n,j}^2 > \tau/\varepsilon\right]$$

$$= \sum_{j=1}^{k_n} u_{n,j}^2 1\left[k_n u_{n,j}^2 > \tau/\varepsilon\right] := \Delta_n(\tau/\varepsilon).$$

Now $\Delta_n(\tau/\varepsilon)$, which does not depend on the weights $V_{n,j}$, converges almost surely P_1 by the strong law to

$$\Delta(\tau/\varepsilon) := \frac{1}{\sigma^2} E\left\{(X_1 - \mu)^2 1\left[(X_1 - \mu)^2 > \frac{\tau\sigma^2}{\varepsilon}\right]\right\},$$

where $\mu = EX_1$. Since $\Delta(\tau/\varepsilon)$, in turn, converges to 0 as ε goes to 0 (by finite variance of X_1), we easily see that (2.7) holds. □

COROLLARY 2.3. *Assume, in addition to* (2.4) *and* (2.5), *that for some* $c > 0$,

$$k_n \sum_{i=1}^{k_n} \left(Y_{n,i} - \bar{Y}_n\right)^2 \to_{P_2} c \quad \text{as } n \to \infty \tag{2.12}$$

and

$$\lim_{\tau \to \infty} \limsup_{n \to \infty} E\left\{\tilde{Y}_{n,1}^2 1\left[\tilde{Y}_{n,1}^2 > \tau\right]\right\} = 0, \tag{2.13}$$

where $\tilde{Y}_{n,1} = k_n(Y_{n,1} - \bar{Y}_n)$. *Then* (2.7) *holds.*

PROOF. For every $\tau > 0$ and $0 < \varepsilon < 1$, we get after applying some elementary bounds along with (2.4) and (2.5) that for all large enough n with P_2 probability greater than $1 - \varepsilon$,

$$D_n(\tau) \le \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} u_{n,i}^2 V_{n,j}^2 1\left(V_{n,j}^2 > \tau/\varepsilon\right)$$

$$\le \frac{2}{c} \sum_{j=1}^{k_n} k_n\left(Y_{n,j} - \bar{Y}_n\right)^2 1\left[\left(k_n\left(Y_{n,j} - \bar{Y}_n\right)\right)^2 > \tau/\varepsilon\right] := \Gamma_n(\tau/\varepsilon).$$

From exchangeability we obtain

$$E\Gamma_n(\tau/\varepsilon) = \frac{2}{c} E\left\{\tilde{Y}_{n,1}^2 1\left[\tilde{Y}_{n,1}^2 > \tau/\varepsilon\right]\right\}.$$

Next, assumption (2.13) says that

$$\lim_{\tau\to\infty} \limsup_{n\to\infty} E\Gamma_n(\tau/\varepsilon) = 0,$$

from which we readily conclude (2.6) and hence (2.7). □

In the two examples to follow, we show how Theorem 2.1 is used to establish the consistency of various bootstrapped means.

EXAMPLE 2.1 (A generalized randomly weighted mean). Let $X_1, X_2, \dots$ be i.i.d. with $0 < \operatorname{Var} X_1 = \sigma^2 < \infty$ and set $X_{n,k} = X_k$ for $k = 1, 2, \dots, n$ and $n \geq 1$. Independent of the X_k's, let $\xi_1, \xi_2, \dots$ be a sequence of i.i.d. strictly positive random variables with $E\xi^2 < \infty$ and set

$$Y_{n,k} = \xi_k \Big/ \sum_{j=1}^{n} \xi_j, \qquad k = 1, 2, \dots, n,\ n \geq 1.$$

It is clear that the assumptions of Corollary 2.2 are satisfied with $k_n = n$. Also

$$\sum_{j=1}^{n} n\left(Y_{j,n} - \overline{Y}_n\right)^2 = n \sum_{j=1}^{n} \xi_j^2 \Big/ \left(\sum_{i=1}^{n} \xi_i\right)^2 - 1, \tag{2.14}$$

which converges almost surely as $n \to \infty$ to

$$\rho_\xi^2 := E\xi_1^2/(E\xi_1)^2 - 1.$$

From (2.7) we get uniformly in t,

$$P\left(\sqrt{n}\left(\overline{X}_{\mathscr{W},n} - \overline{X}_n\right)/(S_n \rho_\xi) \leq t \middle| X_1, \dots, X_n\right) \to \Phi(t) \quad \text{a.s.}, \tag{2.15}$$

as $n \to \infty$. Here $W_n = (Y_{n,1}, Y_{n,2}, \dots, Y_{n,n})$.

EXAMPLE 2.2 (Efron's and Rubin's bootstrapped means). Let $X_1, X_2, \dots$ be i.i.d. F, where F is in the domain of attraction of a normal law. This can be shown to be equivalent to

$$\max_{1 \leq i \leq n} X_i^2/\left(nS_n^2\right) \to_{P_1} 0 \quad \text{as } n \to \infty. \tag{2.16}$$

[See Section 5 of S. Csörgő and Mason (1989) and the references therein.] Now by the subsequence principle, (2.16) holds if and only if for every subsequence $\{l_n\}$ of $\{n\}$ with $l_n \to \infty$ as $n \to \infty$ there is a further subsequence $\{k_n\} \subset \{l_n\}$ with $k_n \to \infty$ as $n \to \infty$ such that

$$\max_{1 \leq i \leq k_n} X_i^2/\left(k_n S_{k_n}^2\right) \to 0 \quad \text{a.s. } P_1 \text{ as } n \to \infty. \tag{2.17}$$

Now let $X_{n,i} = X_i$ for $i = 1, 2, \dots, k_n$ and $n \geq 1$. Obviously (2.4) follows from (2.17). Next set $Y_{n,i} = W_{k_n,i}$ again with $i = 1, \dots, k_n$ and $n \geq 1$, where $W_{k_n,i}$

equals either the multinomial weights $M_{k_n, i}$ in (1.1) or the Dirichlet weights $D_{k_n, i}$ in (1.2). For either choice, routine calculations show that

$$\lim_{n \to \infty} E\Big(k_n\big(Y_{n,1} - \bar{Y}_n\big)\Big)^4 < \infty,$$

from which (2.13) readily follows by Chebyshev's inequality. That assumptions (2.5) and (2.12) hold for the Dirichlet weights is a consequence of the discussion in Example 2.1, since in this case, the vector of $D_{k_n, i}$ is equal in distribution to the vector of $\xi_i / \sum_{j=1}^{k_n} \xi_j$, where $\xi_1, \xi_2, \dots$ are i.i.d. exponential random variables with mean 1. That these two assumptions also hold for the multinomial weights follows from our Lemma 4.1 in the Appendix and Theorem 1 of Hoeffding (1951).

From all of this we can conclude, using Corollary 2.3 and the fact that $c = 1$ in (2.12) for both multinomial and Dirichlet weights, that uniformly in t, as $n \to \infty$,

$$(2.18)\quad P\Big(\sqrt{k_n}\,\big(\bar{X}_{\mathscr{W}, k_n} - \bar{X}_{k_n}\big)/S_{k_n} \le t \Big| X_1, \dots, X_{k_n}\Big) \to \Phi(t) \quad \text{a.s. } P_1,$$

where $\mathscr{W}$ equals $\mathscr{M}$ or $\mathscr{D}$. Consequently, by the subsequence principle we obtain that uniformly in t,

$$(2.19)\qquad P\Big(\sqrt{n}\,\big(\bar{X}_{\mathscr{W}, n} - \bar{X}_n\big)/S_n \le t \Big| X_1, \dots, X_n\Big) \to_{P_1} \Phi(t)$$

as $n \to \infty$. The case $\mathscr{W} = \mathscr{M}$ has been previously proved in Section 5 of S. Csörgő and Mason (1989), where it is shown that for (2.19) to hold, it is necessary that F be in the domain of attraction of a normal law. If $0 < \operatorname{Var} X_1 < \infty$, then from (2.4) we get (2.17) along the full sequence $\{n\}$, which implies that (2.18) also occurs along $\{n\}$.

Note that Theorem 2.1 does not allow the weights $W_{n,i}$ of (1.3) to depend upon the data themselves. Theorem 2.1 can be reformulated such that the weights are conditionally exchangeable given the data $X_1, \dots, X_n$.

3. Consistency of the generalized bootstrapped empirical and quantile process. Let $X_1, X_2, \dots, X_n$ be a sequence of i.i.d. F random variables and for each integer $n \ge 1$ let

$$(3.1)\qquad F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1[X_i \le x], \qquad -\infty < x < \infty,$$

denote the empirical distribution function based on the sample with corresponding empirical quantile function or inverse of F_n:

$$(3.2)\qquad Q_n(s) = \inf\{x \colon F_n(x) \ge s\}, \qquad 0 < s < 1.$$

The quantile function Q of F is defined as in (3.2) with F replacing F_n.

Throughout this section, $\{W_{n,i} \colon i = 1, 2, \dots, n,\ n \ge 1\}$ will denote an array of random weights independent of $X_1, X_2, \dots,$ which are exchangeable on each row and satisfy $(\mathscr{W}_{\mathrm{I}})$, $(\mathscr{W}_{\mathrm{II}})$ and $(\mathscr{W}_{\mathrm{III}})$. The generalized bootstrapped empirical distribution becomes

$$(3.3)\qquad F_{\mathscr{W},n}(x) = \sum_{i=1}^{n} W_{n,i}1[X_i \le x], \qquad -\infty < x < \infty.$$

Notice that exchangeability of $W_{n,1}, \dots, W_{n,n}$ and $(\mathscr{W}_{\mathrm{II}})$ imply that

$$E(F_{\mathscr{W},n}(x)|X_1, \dots, X_n) = F_n(x), \qquad -\infty < x < \infty.$$

The bootstrapped empirical quantile function $Q_{\mathscr{W},n}$ is the left-continuous inverse of $F_{\mathscr{W},n}$:

$$(3.4)\qquad Q_{\mathscr{W},n}(s) = \inf\{x\colon F_{\mathscr{W},n}(x) \ge s\}, \qquad 0 < s < 1.$$

We now define the generalized bootstrapped empirical process to be

$$(3.5)\qquad \alpha_{\mathscr{W},n}(x) = n^{1/2}(F_{\mathscr{W},n}(x) - F_n(x)), \qquad -\infty < x < \infty,$$

and the generalized bootstrapped quantile process to be

$$(3.6)\qquad \beta_{\mathscr{W},n}(s) = n^{1/2}(Q_n(s) - Q_{\mathscr{W},n}(s)), \qquad 0 < s < 1.$$

The following theorem establishes the consistency of these two bootstrapped processes and is the analog of Theorems 4.1 and 5.1 of Bickel and Freedman (1981).

THEOREM 3.1. *Suppose that in addition to* $(\mathscr{W}_{\mathrm{I}})$, $(\mathscr{W}_{\mathrm{II}})$ *and* $(\mathscr{W}_{\mathrm{III}})$, *and exchangeability, the weights satisfy*

$$(3.7)\qquad \max_{1 \le i \le n} nW_{n,i}^2 \to_{P_2} 0 \qquad \text{as } n \to \infty.$$

Then along almost all sample sequences, given $X_1, \dots, X_n$,

$$(3.8)\qquad \alpha_{\mathscr{W},n} \text{ converges weakly to } c^{1/2}B(F),$$

where B is a Brownian bridge. Furthermore, if F has a positive continuous density quantile function $f(Q)$, then along almost all sample sequences, given $X_1, \dots, X_n$, *for all* $0 < a < b < 1$,

$$(3.9)\qquad \beta_{\mathscr{W},n} \text{ converges weakly to } c^{1/2}hB$$

on $D_L[a, b]$ (the left-continuous version of $D[a, b]$), where $h = 1/f(Q)$.

PROOF. For integers $n \ge 1$, set

$$(3.10)\qquad W_n(t) = \sum_{i \le nt} W_{n,i}, \qquad 0 \le t \le 1,$$

where the empty sum is defined to be 0 and let

$$(3.11)\qquad V_n(t) = \inf\{u\colon W_n(u) \ge t,\ u \in [0,1]\}, \qquad 0 \le t \le 1,$$

be the left-continuous inverse of W_n.

The assumptions on the weights allow us to apply Billingsley's (1968) Theorem 24.2 to infer that the process

$$\overline{w}_n(t) = n^{1/2} \sum_{i \le nt} (W_{n,i} - 1/n), \qquad 0 \le t \le 1,$$

converges weakly to $c^{1/2}B$. Moreover, since it is readily checked that

$$\sup_{0 \le t \le 1} \left| \overline{w}_n(t) - n^{1/2}\{W_n(t) - t\} \right| \to 0 \quad \text{as } n \to \infty,$$

we then conclude, with I denoting the identity function, that

$$n^{1/2}(W_n(I) - I) \text{ converges weakly to } c^{1/2}B. \tag{3.12}$$

Furthermore, from (3.12) and Theorem 2 of Vervaat (1972), we also get more generally that

$$\begin{aligned} &\left(n^{1/2}(W_n(I) - I), n^{1/2}(I - V_n(I))\right) \\ &\qquad \text{converges weakly to } (c^{1/2}B, c^{1/2}B). \end{aligned} \tag{3.13}$$

Next, the Skorokod representation theorem permits us to construct a probability space on which there sit sequences of probabilistically equivalent versions $\tilde{W}_n$ and $\tilde{V}_n$ of W_n and V_n, respectively, for $n \ge 1$, and a fixed Brownian bridge B such that almost surely both

$$\sup_{0 \le t \le 1} \left| n^{1/2}\left(\tilde{W}_n(t) - t\right) - c^{1/2}B(t) \right| \to 0 \quad \text{as } n \to \infty \tag{3.14}$$

and

$$\sup_{0 \le t \le 1} \left| n^{1/2}\left(t - \tilde{V}_n(t)\right) - c^{1/2}B(t) \right| \to 0 \quad \text{as } n \to \infty. \tag{3.15}$$

Now extend this probability space to include a sequence $X_1, X_2, \dots$ of i.i.d. F random variables independent of $\{(\tilde{W}_n, \tilde{V}_n)\colon n \ge 1\}$, and B. Moreover, since $\{X_n\}_{n \ge 1} =_{\mathscr{D}} \{Q(U_n)\}_{n \ge 1}$, where $U_1, U_2, \dots$ is a sequence of independent uniform $(0, 1)$ random variables, we can assume that the probability space is formed so that

$$X_n = Q(U_n), \qquad n \ge 1. \tag{3.16}$$

By exchangeability of $W_{n,1}, \dots, W_{n,n}$, $\alpha_{\mathscr{W},n}$ is equal in distribution (conditioned on $X_1, \dots, X_n$) to

$$n^{1/2}\left(\sum_{i=1}^{n} W_{n,i} 1\left[X_{(n,i)} \le x\right] - F_n(x) \right), \tag{3.17}$$

where $X_{(n,1)} \le X_{(n,2)} \le \cdots \le X_{(n,n)}$ are the order statistics of $X_1, \dots, X_n$. Therefore, conditioned on $X_1, \dots, X_n$,

$$\begin{aligned} \alpha_{\mathscr{W},n}(x) &=_{\mathscr{D}} n^{1/2}\left(\sum_{i=1}^{nF_n(x)} W_{n,i} - F_n(x) \right) \\ &= n^{1/2}(W_n(F_n) - F_n)(x) \quad \text{by (3.10)} \\ &=_{\mathscr{D}} n^{1/2}\left(\tilde{W}_n(F_n) - F_n\right)(x). \end{aligned}$$

Now the Glivenko–Cantelli theorem combined with the almost sure uniform continuity of B and (3.14) implies that almost surely along $X_1, X_2, \ldots$ as $n \to \infty$,

$$\sup_{-\infty < x < \infty} \left| n^{1/2}\{\tilde{W}_n(F_n(x)) - F_n(x)\} - c^{1/2}B(F(x)) \right| \to 0. \tag{3.18}$$

This establishess the first part of Theorem 3.1.

To prove the second part of Theorem 3.1, we require some facts. Define the uniform empirical quantile function based on $U_1, \ldots, U_n$ for $n \geq 1$ to be

$$\mathbf{U}_n(s) = \inf\{u: G_n(u) \geq s,\, u \in [0,1]\}, \qquad 0 \leq s \leq 1,$$

where G_n is the empirical distribution based on $U_1, \ldots, U_n$. Let

$$u_n(s) = n^{1/2}\{s - \mathbf{U}_n(s)\}, \qquad 0 \leq s \leq 1,$$

denote the uniform quantile process. By the Glivenko–Cantelli theorem

$$\sup_{0 \leq s \leq 1} |\mathbf{U}_n(s) - s| \to 0 \quad \text{a.s. as } n \to \infty. \tag{3.19}$$

We shall also require the following fact that can be derived from results in Mason (1984) [refer also to pages 581 and 582 of Shorack and Wellner (1986)], that for all $M > 0$, as $n \to \infty$,

$$\sup_{0 \leq h \leq M/n^{1/2}} \sup_{0 \leq s \leq 1-h} |u_n(s+h) - u_n(s)| \to 0 \quad \text{a.s.} \tag{3.20}$$

This fact will be very useful when it is combined with

$$\sup_{0 \leq t \leq 1} n^{1/2}\left|t - \tilde{V}_n(t)\right| \to c^{1/2} \sup_{0 \leq t \leq 1} |B(t)| \quad \text{a.s. as } n \to \infty, \tag{3.21}$$

which is a direct consequence of (3.15).

Finally, we observe that conditioned on $X_1 = Q(U_1), \ldots, X_n = Q(U_n)$,

$$\beta_{\mathscr{W},n} =_{\mathscr{D}} n^{1/2}\left\{Q(\mathbf{U}_n) - Q\left(\mathbf{U}_n\left(\tilde{V}_n\right)\right)\right\}.$$

Now for each $0 < a \leq s \leq b < 1$, we can write

$$n^{1/2}\left\{Q(\mathbf{U}_n(s)) - Q\left(\mathbf{U}_n\left(\tilde{V}_n(s)\right)\right)\right\} = h(s_1)n^{1/2}\left\{\mathbf{U}_n(s) - \mathbf{U}_n\left(\tilde{V}_n(s)\right)\right\}$$

for some s_1 between $\mathbf{U}_n(\tilde{V}_n(s))$ and $\tilde{V}_n(s)$, which in turn equals

$$\begin{aligned} c^{1/2}h(s)B(s) &+ \left\{h(s_1)n^{1/2}\left(s - \tilde{V}_n(s)\right) - c^{1/2}h(s)B(s)\right\} \\ &+ h(s_1)\left\{u_n(s) - u_n\left(\tilde{V}_n(s)\right)\right\} \\ &=: c^{1/2}h(s)B(s) + D_{n,1}(s) + D_{n,2}(s). \end{aligned}$$

By continuity of h, (3.15), (3.19) and (3.21), we get

$$\sup_{a \leq s \leq b} |D_{n,1}(s)| \to 0 \quad \text{a.s. as } n \to \infty,$$

and from (3.19), (3.20) and (3.21) along with boundedness of h on $[c,d]$ for any $0 < c < d < 1$, we have likewise

$$\sup_{a \le s \le b} |D_{n,2}(s)| \to 0 \quad \text{a.s. as } n \to \infty.$$

Thus along almost all sample sequences $X_1, X_2, \dots$ as $n \to \infty$,

$$\sup_{a \le s \le b} |\beta_{\mathscr{Y},n}(s) - c^{1/2}h(s)B(s)| \to 0 \quad \text{a.s.},$$

which finishes the proof of the second part of Theorem 3.1. □

The proof of the second part of our Theorem 3.1 follows roughly the lines of the proof that Swanepoel (1986) provided for the original Bickel and Freedman (1981) theorem on the bootstrapped quantile process.

APPENDIX

LEMMA 4.1. *Consider the scaled multinomial weights M_n defined in* (1.1). *For any integer $r \ge 2$, as $n \to \infty$,*

$$n^{r-1} \sum_{j=1}^{n} |M_{n,j} - 1/n|^r \to_P E|N(1) - 1|^r, \tag{4.1}$$

where $N(1)$ is a Poisson random variable with mean 1.

PROOF. Introduce a sequence of i.i.d. uniform $(0,1)$ random variables $U_1, U_2, \dots$ and define the counting process $G_n(t) = \sum_{i=1}^{n} 1[U_i \le t]$. One way to construct the scaled multinomial n vector M_n is by putting

$$M_{n,j} = \frac{1}{n}\{G_n(j/n) - G_n((j-1)/n)\} \tag{4.2}$$

for $1 \le j \le n$.

For real t, let $N(t)$ be a Poisson process independent of $U_1, U_2, \dots$ and scaled so that $N(t) \sim \text{Poisson}(t)$. Put $N_n = N(n)$. It is easy to verify that $G_n^*(t) = \sum_{j=1}^{N_n} 1[U_i \le t]$ is another Poisson process with $G_n^*(t) \sim \text{Poisson}(nt)$. Mimicking (4.2), define the n vector M_n^* to have elements

$$M_{n,j}^* = \frac{1}{n}\{G_n^*(j/n) - G_n^*((j-1)/n)\} \tag{4.3}$$

for $1 \le j \le n$. The elements of M_n^* are $1/n$ times i.i.d. mean 1 Poisson variables. Hence, by the weak law of large numbers,

$$n^{r-1} \sum_{j=1}^{n} |M_{n,j}^* - \overline{M}_n^*|^r \to_P E|N(1) - 1|^r, \tag{4.4}$$

where $\overline{M}_n^*$ is the average of $M_{n,j}^*$. On account of (4.4), using the triangle and

Markov inequalities, we see that to prove (4.1) it suffices to show that for each $r \geq 2$, $ES_n \to 0$, where

$$S_n := n^{r-1} \sum_{j=1}^{n} \left| M_{n,j}^* - \overline{M}_n^* - M_{n,j} + 1/n \right|^r. \tag{4.5}$$

Upon conditioning and using exchangeability, we have $ES_n = ET_n$, where

$$T_n := n^r E\left(\left| M_{n,1}^* - M_{n,1} - \overline{M}_n^* + 1/n \right|^r \middle| N_n \right) \tag{4.6}$$

$$= E\left(\left| \sum_{i \in (n, N_n)} (1[U_i \leq 1/n] - 1/n) \right|^r \middle| N_n \right). \tag{4.7}$$

The notation $i \in (n, N_n)$ above means that the summation is from $n+1$ to N_n if $N_n > n$, from $N_n + 1$ to n if $N_n < n$, and 0 if $N_n = n$. Applying Lemma 1.1 of van Zuijlen (1978), there exists a constant A_r such that for all $n \geq 2$,

$$T_n \leq A_r \left(\frac{|N_n - n|}{n} + \frac{|N_n - n|^r}{n^r} \right). \tag{4.8}$$

A standard inequality [Loève (1977), page 276] and Liapunov's inequality (alternatively, the law of large numbers plus uniform integrability) now readily yield $ET_n \to 0$ as $n \to \infty$. □

Acknowledgments. We thank Jon Wellner for his interest in the problem and Adrian Raftery for research advice. We are also grateful to Martien van Zuijlen and the referee for a careful reading of the manuscript.

REFERENCES

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

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

CHERNOFF, H. and TEICHER, H. (1958). A central limit theorem for sums of interchangeable random variables. *Ann. Math. Statist.* **29** 118–130.

CSÖRGŐ, S. and MASON, D. M. (1989). Bootstrapping empirical functions. *Ann. Statist.* **17** 1447–1471.

DOES, R. J. M. M. (1983). An Edgeworth expansion for simple linear rank statistics under the null-hypothesis. *Ann. Statist.* **11** 607–624.

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

HÁJEK, J. (1961). Some extensions of the Wald–Wolfowitz–Noether theorem. *Ann. Math. Statist.* **32** 506–523.

HOEFFDING, W. (1951). Combinatorial central limit theorem. *Ann. Math. Statist.* **22** 558–566.

LO, A. Y. (1987). A large sample study of the Bayesian bootstrap. *Ann. Statist.* **15** 360–375.

LOÈVE, M. (1977). *Probability Theory* **1**, 4th ed. Springer, New York.

MASON, D. M. (1984). A strong limit theorem for the oscillation modulus of the uniform empirical process. *Stochastic Process. Appl.* **17** 127–136.

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

SCHNELLER, W. (1989). Edgeworth expansions for linear rank statistics. *Ann. Statist.* **17** 1103–1123.

SHORACK, G. and WELLNER, J. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

SINGH, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1186–1195.

SWANEPOEL, J. W. H. (1986). A note on proving that the (modified) bootstrap works. *Comm. Statist. Theory Methods* **15** 3193–3203.

VAN ZUIJLEN, M. C. A. (1978). Properties of the empirical distribution function for independent nonidentically distributed random variables. *Ann. Probab.* **6** 250–266.

VERVAAT, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrsch. Verw. Gebiete* **23** 245–253.

WENG, C. S. (1989). On a second-order asymptotic property of the Bayesian bootstrap mean. *Ann. Statist.* **17** 705–710.

ZHENG, Z. and TU, D. (1988). Random weighting method in regression models. *Sci. Sinica Ser. A* **31** 1442–1459.

DEPARTMENT OF MATHEMATICAL SCIENCES
501 EWING HALL
UNIVERSITY OF DELAWARE
NEWARK, DELAWARE 19716

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706