Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

R. Helmers, P. Janssen, R. Serfling

Berry-Esseen rates and bootstrap results for generalized L-statistics

Department of Mathematical Statistics Report MS-R8808 June

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Berry-Esseen Rates and Bootstrap Results for Generalized L-statistics

R. Helmers
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

P. Janssen
Limburgs Universitair Centrum, Universitaire Campus, B-3610
Diepenbeek, Belgium
and

R. Serfling
Department of Mathematical Sciences
Johns Hopkins University, Baltimore, Maryland 21218, U.S.A.

The rate of convergence of the distribution of a generalized L-statistic to its normal limit is established. Based on this result the corresponding bootstrap approximation is shown to be asymptotically valid, thus providing an alternative to the use of the normal approximation. By the same method of proof the asymptoic accuracy of the bootstrap approximation of generalized L-statistics is also obtained.

AMS 1980 *subject classifications:* Primary 62E20; secondary 62G05.
Keywords and phrases. Bootstrap, generalized L-statistics, Berry-Esseen theorem, rate of convergence, empirical measures.
Note: Research supported by the U.S. Department of Navy under Office of Naval Research Contract No. N00014-79-C-0801 and by NATO under Research Grant No. 0034/87.

1. Introduction

Let $X_1,...,X_n$ be independent random variables having common probability distribution F, and let $h(x_1,...,x_m)$ be a *kernel* of *degree m* (i.e. a real-valued measurable function in m arguments). Further let H_F denote the distribution function of the random variable $h(X_1,...,X_m)$ and let $H_F(y)$ be estimated by

$$H_n(y) = n_{(m)}^{-1}\Sigma 1\{h(X_{i_1},...,X_{i_m})\leq y\},\ y\in\mathbb{R}$$

where the sum extends over the $n_{(m)}=n(n-1)...(n-m+1)$ m-tuples $(i_1,...,i_m)$ with $1\leq i_1,...,i_m\leq n$ and $i_1,...,i_m$ distinct. Denote by

$$W_{n,1}\leq W_{n,2}\leq \cdots \leq W_{n,n_{(m)}}$$

the ordered evaluations $h(X_{i_1},...,X_{i_m})$ taken over these $n_{(m)}$ m-tuples $(i_1,...,i_m)$. Serfling (1984), Silverman (1983), Janssen, Serfling and Veraverbeke (1984) and Helmers and Ruymgaart (1988) recently investigated the asymptotic normality of various classes of *generalized L-statistics*, including the class we consider here, defined as $T_n=T(H_n)$, where $T(\cdot)$ is an *L-functional*,

$$T(G) = \int_0^1 J(t)G^{-1}(t)dt,$$

Report MS-R8808
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

with J a given weight function on (0,1) and $G^{-1}(t)=\inf\{y:G(y)\geq t\}$ for any df G. In particular, for $J=1, T_n$ reduces to the usual *U-statistics* based on the kernel h. Note that, with

$$c_{ni} = \int_{(i-1)/n_{(m)}}^{i/n_{(m)}} J(t)dt,$$

the statistics T_n may be expressed as $T_n=\Sigma_{i=1}^{n_{(m)}} c_{ni} W_{n,i}$. The basic asymptotic normality result for T_n is (e.g., see (3.7) in HELMERS and RUYMGAART (1988)).

$$n^{1/2}(T_n - T(H_F)) \xrightarrow[d]{} N(0,\sigma^2(T,H_F)), \tag{1.1}$$

where

$$\sigma^2(T,G) = m^2 \int\int J(G(x))J(G(y))\{G(x)\wedge G(y) - G(x)G(y)\}dxdy.$$

The parameter $T(H_F)$ estimated by T_n is called a *generalized L-functional.*

In this paper we establish, under a Lipschitz condition on J, the *Berry-Esseen rate* associated with the convergence (1.1), and we apply this result to show the asymptotic validity of the *bootstrap approximation* for the exact df of the statistic T_n. By the same method of proof the asymptotic accuracy of the bootstrap approximation for generalized L-statistics is also obtained. Such a result provides an alternative for the normal approximation discussed above and extends the work of BICKEL and FREEDMAN (1981), on asymptotic validity of bootstrap approximations for the distribution functions of non-degenerate U- and L-statistics. Further relevant background include BRETAGNOLLE (1983), treating von Mises statistics, and BOOS, JANSSEN and VERAVERBEKE (1988), where resampling plans for two-sample U-statistics with estimated parameters are studied.

Our method of proof resembles that of SINGH (1981), who employed the classical Berry-Esseen theorem as his main tool to establish the asymptotic validity of the bootstrap approximation for the distribution function of the ordinary sample mean. Similarly, our proof will rely on a Berry-Esseen result for generalized L-statistics. Also instrumental in our proof will be certain Glivenko-Cantelli results for the empirical df H_n recently established by HELMERS, JANSSEN and SERFLING (1988). Our Berry-Esseen result is developed in Section 2 and our bootstrap results in Sections 3 and 4.

Finally note that the discussion in the next sections restrict attention to the case of *symmetric h.* At the cost of some straightforward technicalities our results can be shown to hold for arbitrary h. We omit this general discussion in the interest of brevity.

2. BERRY-ESSEEN BOUNDS FOR GENERALIZED L-STATISTICS.

We assume the weight function J to be Lipschitz of order 1, i.e.,

$$|J(s)-J(t)| \leq K|s-t|, \quad 0<s,\ t<1, \quad \text{for some } K>0,$$

in which case also J is bounded: $|J|\leq M<\infty$. As further notation, put $\theta = Eh(X_1,...,X_m)$,

$$g(x) = \int \cdots \int h(x,x_2...,x_m)dF(x_2)\cdots dF(x_m) - \theta$$

and define for $n\geq m$

$$\tilde{F}_{T_n}(x) = P[n^{1/2}(T_n - T(H_F))/\sigma(T,H_F) \leq x], \quad x\in\mathbb{R}.$$

Finally let $\Phi(x)$ denote the standard normal distribution function.

THEOREM 1. *Suppose that*

(i) *J is Lipschitz of order 1;*

(ii) *$E|g(X_1)|^3<\infty$ and $Eh^2(X_1,...,X_m)<\infty$;*

(iii) $\sigma^2 = \sigma^2(T,H_F)>0.$

Then there exists a universal constant $C>0$ such that for all $n\geq 2m$

$$\sup_x |\tilde{F}_{T_n}(x) - \Phi(x)| \tag{2.1}$$
$$\leq Cn^{-1/2}\{\frac{M^3}{\sigma^3}[E|g|^3 + (E|h|)^3] + \frac{Eh^2}{\sigma^2}(2m-1)^2(M^2 + \gamma_m K^2 n^{-1/2}) + 1 + \frac{KE|h|}{\sigma}\}$$

where γ_m is a constant depending only on m.

PROOF. Let G_1, G_2 be df's satisfying $\int |J(t)G_i^{-1}(t)|dt < \infty, i=1,2$. As in SERFLING (1980), p. 265 we have

$$T(G_2) - T(G_1) = -\int [\psi(G_2(y)) - \psi(G_1(y))]dy \tag{2.2}$$

where $\psi(u) = \int_0^u J(t)dt$. The Lipschitz condition on J implies that

$$|\psi(G_2(y)) - \psi(G_1(y)) - (G_2(y) - G_1(y))J(G_1(y))| \leq K(G_2(y) - G_1(y))^2. \tag{2.3}$$

From (2.2) and (2.3) follows

$$|T_n - T(H_F) + \int (H_n(y) - H_F(y))J(H_F(y))dy| \leq K\int (H_n(y) - H_F(y))^2 dy. \tag{2.4}$$

Now note that

$$-\int (H_n(y) - H_F(y))J(H_F(y))dy = U_n \tag{2.5}$$

where

$$U_n = \binom{n}{m}^{-1} \sum_{C_{n,m}} h_1(X_{i_1}, \ldots, X_{i_m})$$

with $C_{n,m} = \{(i_1, \ldots, i_m): 1 \leq i_1 < \cdots < i_m \leq n\}$ and

$$h_1(x_1, \ldots, x_m) = -\int [\mathbb{1}\{h(x_1, \ldots, x_m) \leq y\} - H_F(y)]J(H_F(y))dy.$$

For the r.h.s. of (2.4) we have

$$K\int (H_n(y) - H_F(y))^2 dy = W_n + R_n \tag{2.6}$$

with

$$W_n = \binom{n}{m}^{-2} \Sigma^* h_2(X_{i_1}, \ldots, X_{i_m}, X_{j_1}, \ldots, X_{j_m})$$

and

$$R_n = \binom{n}{m}^{-2} \Sigma^{**} h_2(X_{i_1}, \ldots, X_{i_m}, X_{j_1}, \ldots, X_{j_m})$$

where Σ^*, resp. Σ^{**}, denotes the sum over all pairs of m-tuples $(i_1, \ldots, i_m), (j_1, \ldots, j_m) \in C_{n,m}$ having *all indices different, resp. at least two indices equal* and with

$$h_2(x_1, \ldots, x_{2m})$$
$$= K\int [\mathbb{1}\{h(x_1, \ldots, x_m) \leq y\} - H_F(y)][\mathbb{1}\{h(x_{m+1}, \ldots, x_{2m}) \leq y\} - H_F(y)]dy.$$

Note that W_n is a U-statistic with kernel h_{2n} depending on n. Indeed

$$W_n = \binom{n}{2m}^{-1} \sum_{C_{n,2m}} h_{2n}(X_{k_1}, \ldots, X_{k_{2m}})$$

with

$$h_{2n}(x_{k_1}, \ldots, x_{k_{2m}}) = \binom{n}{2m}\binom{n}{m}^{-2} \Sigma^+ h_2(x_{i_1}, \ldots, x_{i_m}, x_{j_1}, \ldots, x_{j_m}) \tag{2.7}$$

where Σ^+ denotes the sum over all $\binom{2m}{m}$ summands $h_2(x_{i_1}, \ldots, x_{i_m}, x_{j_1}, \ldots, x_{j_m})$ for which

$\{i_1,...,i_m,j_1,...,j_m\}=\{k_1,k_2,...,k_{2m}\}$. For the second term R_n in the r.h.s of (2.6) we will exploit that its magnitude is of lower order.

As in HELMERS (1981), see also HELMERS (1982), we now use (2.4)-(2.6) to approximate $T_n - T(H_F)$ from above and below by $U_n + W_n + R_n$ and $U_n - W_n - R_n$, i.e., for all $n \geq 2m$

$$U_n - W_n - R_n \leq T_n - T(H_F) \leq U_n + W_n + R_n.$$

From this approximation we directly see that

$$P[n^{1/2}\sigma^{-1}(U_n + W_n + ER_n) \leq x - n^{-1/2}] + P[|R_n - ER_n| \geq \sigma n^{-1}] \leq \tag{2.8}$$
$$\tilde{F}_{T_n}(x) \leq P[n^{1/2}\sigma^{-1}(U_n - W_n - ER_n) \leq x + n^{-1/2}] +$$
$$+ P[|R_n - ER_n| \geq \sigma n^{-1}].$$

Since both $U_n - W_n$ and $U_n + W_n$ are U-statistics of degree $2m$ with varying kernels $h_n^{\mp} = h_{1n} \pm h_{2n}$ where h_{2n} is defined by (2.7) and

$$h_{1n}(x_{k_1},...,x_{k_{2m}})$$
$$= \binom{n}{2m}\binom{n}{m}^{-1}\binom{n-m}{m}^{-1}\Sigma^+ \tfrac{1}{2}[h_1(x_{i_1},...,x_{i_m}) + h_1(x_{j_1},...,x_{j_m})]$$

with Σ^+ as in (2.7), the Berry-Esseen result for U-statistics due to VAN ZWET (1984) applies. So, using also the inequality,

$$\sup_x |\Phi(x+q) - \Phi(x)| \leq |q| \quad \text{we get for } U_n - W_n - ER_n$$
$$\sup_x |P[n^{1/2}\sigma^{-1}(U_n - W_n - ER_n) \leq x + n^{-1/2}] - \Phi(x)| \tag{2.9}$$
$$\leq \sup_x |P[n^{1/2}\sigma^{-1}(U_n - W_n) \leq x] - \Phi(x)|$$
$$+ \sup_x |\Phi(x + n^{-1/2} + n^{1/2}\sigma^{-1}ER_n) - \Phi(x)|$$
$$\leq Cn^{-1/2}\left[\frac{E|g_n(X_1)|^3}{(Eg_n^2(X_1))^{3/2}} + \frac{(2m-1)^2 E(h_n^-(X_1,...,X_{2m}))^2}{Eg_n^2(X_1)}\right]$$
$$+ n^{-1/2}(1 + \sigma^{-1}|ER_n|)$$

where

$$g_n(x) = \int \cdots \int h_n^-(x, x_2,...,x_{2m}) dF(x_2) \cdots dF(x_{2m}). \tag{2.10}$$

The discussion just given deals with the first term of the upper bound for $\tilde{F}_{T_n}(x)$ given in (2.8). To handle the first term of the lower bound a similar argument holds. We therefore restrict the discussion for the lower bound to the remark that, since h_{2n} is a degenerate kernel, an alternative way to define $g_n(x)$ is

$$g_n(x) = \int \cdots \int h_n^+(x, x_2,...,x_{2m}) dF(x_2) \cdots dF(x_{2m}).$$

Hence h_n^+ and h_n^- have the same projection. Now it can be shown, by some elementary calculations, that

$$|ER_n| \leq 2K \frac{m}{n-m+1} E|h(X_1,...,X_m)| \tag{2.11}$$
$$E|g_n(X_1)|^3 \leq C_1 M^3 [E|g(X_1)|^3 + (E|h(X_1,...,X_m)|)^3] \tag{2.12}$$
$$E(h_n^{\pm}(X_1,...,X_{2m}))^2 \leq C_2(M^2 + K^2) Eh^2(X_1,...,X_m). \tag{2.13}$$

Finally note that

$$Eg_n^2(X_1) = \sigma^2(T,H_F). \tag{2.14}$$

The latter equality follows since h_{1n} is obtained by rewriting the kernel h_1 of U_n is such a way that U_n, a U-statistic with kernel of degree m, transforms into a U-statistic of degree $2m$ and since h_{2n}, being a degenerate kernel, has no contribution to g_n. From (2.9)-(2.14) the appropriate bound for the first term in the r.h.s. of (2.8) is obtained. It remains to show that the second term in the r.h.s. of (2.8) is of the right order. To verify this note that

$$P[|R_n - ER_n| \geqslant \sigma n^{-1}] \leqslant n^2 \sigma^{-2} VarR_n. \tag{2.15}$$

Since R_n is a linear combination of U-statistics of degree at most $2m-1$ with coefficients which are at most of the order n^{-1}, a simple calculation (using Lemma A(i) in SERFLING (1980), p. 183) yields

$$VarR_n \leqslant K^2 \frac{\gamma_m}{n^3} Eh^2(X_1,...,X_m), \tag{2.16}$$

where γ_m is a constant depending only on m. From (2.15) and (2.16) we get that

$$P[|R_n - ER_n| \geqslant \sigma n^{-1}] \leqslant n^{-1} \sigma^{-2} \gamma_m K^2 Eh^2(X_1,...,X_m). \qquad \square$$

The following extension of Theorem 1 provides the appropriate order bound for application in Section 3.

COROLLARY 1. *Assume the conditions of Theorem 1 with* $E|g(X_1)|^3 < \infty$ *replaced by* $E|g(X_1)|^{2+\delta} < \infty$, *for some* $0 < \delta \leqslant 1$. *Then there exists a positive constant* C_δ, *depending only on* δ, *such that for all* $n \geqslant 2m$

$$\sup_x |\tilde{F}_{T_n}(x) - \Phi(x)| \tag{2.17}$$

$$\leqslant C_\delta n^{-\delta/2} \{ \frac{M^{2+\delta}}{\sigma^{2+\delta}} [E|g|^{2+\delta} + (E|h|)^{2+\delta}] + \frac{Eh^2}{\sigma^2} (2m-1)^2 (M^2 + \gamma_m K^2 n^{(\delta-2)/2}) + 1 + \frac{KE|h|}{\sigma} \},$$

with M, K *and* γ_m *as in Theorem 1.*

PROOF. First note that only minor changes are needed to obtain the appropriate modification of the Berry-Esseen theorem for U-statistics as proved in VAN ZWET (1984). The classical argument leading to his (3.7) must be replaced by a similar computation with $n^{-1/2}$ replaced by $n^{-\delta/2}$ and $E|g(X_1)|^3 < \infty$ by $E|g(X_1)|^{2+\delta} < \infty$ (see e.g. PETROV (1975), p. 115). This provides an appropriate modification of (2.9). Finally note that the inequality (2.12) is now replaced by

$$E|g_n(X_1)|^{2+\delta} \leqslant C_3 M^{2+\delta} [E|g(X_1)|^{2+\delta} + (E|h(X_1,...,X_m)|)^{2+\delta}]$$

with C_3 a positive constant depending only on δ. $\square$

3. BOOTSTRAPPING GENERALIZED L-STATISTICS

We shall need some additional notation. For $X_1,...,X_n$ a sequence of independent random variables with common distribution F, let F_n denote the corresponding empirical distribution function, i.e. $F_n(x) = n^{-1}\Sigma_{i=1}^n \mathbb{1}\{X_i \leqslant x\}$. The bootstrap method uses *resampling* with replacement from the observations $\{X_i\}$. Thus, given $X_1,...,X_n$, we obtain in each resample a collection of random variables $X_1^*,...,X_n^*$ which are conditionally independent with common distribution F_n.

For any $n \geqslant m$ we define (h is assumed to be symmetric)

$$H_n^*(y) = \binom{n}{m}^{-1} C_{n,m}^{\Sigma} \mathbb{1}\{h(X_{i_1}^*,...,X_{i_m}^*) \leqslant y\},\ y \in \mathbb{R},$$

the empirical distribution function of U-statistic structure based on the bootstrap sample $X_1^*,...,X_n^*$. Let $T_n, T(H_F)$ and $\sigma^2(T,G)$ be as in Section 1 and $\tilde{F}_{T_n}$ as in Section 2. Also, define for $n \geqslant m$

$$F_{T_n}(x) = P[n^{1/2}(T_n - T(H_F)) \leqslant x],\ x \in \mathbb{R}.$$

The objects of our interest are bootstrap approximations for the df's F_{T_n} and $\tilde{F}_{T_n}$. These approximations are given by

$$F^*_{T_n}(x) = P^*[n^{1/2}(T(H^*_n)-T(H_n))\leqslant x],\ x\in\mathbb{R}$$

and

$$\tilde{F}^*_{T_n}(x) = P^*[n^{1/2}(T(H^*_n)-T(H_n))/\sigma(T,H_n)\leqslant x],\ x\in\mathbb{R}$$

where P^* denotes probability under F_n. Our main result is the following.

THEOREM 2. *Suppose that*
(i) *J is Lipschitz of order 1;*
(ii) $\max_{1\leqslant i_1\leqslant\cdots\leqslant i_m\leqslant m} E|h(X_{i_1},...,X_{i_m})|^{2+\delta}<\infty$, *for some* $0<\delta\leqslant 1$;
(iii) $\sigma^2=\sigma^2(T,H_F)>0$.
Then, with probability 1,

$$\lim_{n\to\infty}\sup_x|F_{T_n}(x)-F^*_{T_n}(x)| = 0 \tag{3.1}$$

and

$$\lim_{n\to\infty}\sup_x|\tilde{F}_{T_n}(x)-\tilde{F}^*_{T_n}(x)| = 0. \tag{3.2}$$

PROOF. To prove (3.1), note that

$$\sup_x|F_{T_n}(x)-F^*_{T_n}(x)|\leqslant A_n+B_n+C_n \tag{3.3}$$

where $A_n=\sup_x|F_{T_n}(x)-\Phi(x\sigma^{-1})|, B_n=\sup_x|F^*_{T_n}(x)-\Phi(x\sigma_n^{-1}|$, with

$\sigma_n^2=\sigma^2(T,H_n)$, and $C_n=\sup_x|\Phi(x\sigma^{-1})-\Phi(x\sigma_n^{-1})|$. The term A_n is not random, whereas B_n and C_n are random.

To see that $A_n\to 0, n\to\infty$, apply Theorem 3.2 in SERFLING (1984).

To show that $B_n\to 0, n\to\infty$ a.s. a more refined argument is needed. Let E^* denote expectation under F_n and define

$$\theta_n = E^*h(X^*_1,...,X^*_m) = n^{-m}\sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n}h(X_{i_1},...,X_{i_m}),$$

the natural estimator of $\theta=Eh(X_1,...,X_m)$, and

$$g^*_n(x) = \int\cdots\int h(x,x_2,...,x_m)dF_n(x_2)...dF_n(x_m)-\theta_n.$$

An application of Corollary 1 yields

$$B_n\leqslant Cn^{-\delta/2}\{\frac{E^*|g^*_n(X^*_1)|^{2+\delta}+(E^*|h(X^*_1,...,X^*_m)|)^{2+\delta}}{\sigma_n^{2+\delta}} \tag{3.4}$$

$$+\frac{E^*h^2(X^*_1,...,X^*_m)}{\sigma_n^2}+\frac{E^*|h(X^*_1,...,X^*_m)|}{\sigma_n}\},$$

where C is a constant depending on δ,M and K only. To proceed we evaluate the (conditional) moments appearing in the r.h.s of (3.4). An application of Jensen's inquality for conditional expectations in combination with the inequality

$$|a-b|^{2+\delta}\leqslant 2^{2+\delta}(|a|^{2+\delta}+|b|^{2+\delta})\ \text{yields}$$

$$E^*|g_n^*(X_1^*)|^{2+\delta} \leq 2^{2+\delta}(E^*|h(X_1^*,...,X_m^*)|^{2+\delta}+|\theta_n|^{2+\delta}). \tag{3.5}$$

Clearly

$$E^*|h(X_1^*,...,X_m^*)|^{2+\delta} = n^{-m}\sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n}|h(X_{i_1},...,X_{i_m})|^{2+\delta}. \tag{3.6}$$

By Condition (ii) the strong law of large numbers for von Mises statistics applies. Therefore the r.h.s. of (3.6) converges to $E|h(X_1,...,X_m)|^{2+\delta}$ a.s. and

$$\theta_n \to \theta,\ n\to\infty,\ \text{a.s.} \tag{3.7}$$

From (3.5)-(3.7) we obtain, with probability one, that $E^*|g_n^*(X_1^*)|^{2+\delta}$ is bounded by some finite constant. Again by the strong law for von Mises statistics we have for $r=1,2$ that

$$E^*|h(X_1^*,...,X_m^*)|^r = n^{-m}\sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n}|h(X_{i_1},...,X_{i_m})|^r. \tag{3.8}$$

$$\to E|h(X_1,...,X_m)|^r,\ n\to\infty,\ \text{a.s.}$$

Note that for (3.8) Condition (ii) becomes effective again.
From (3.4)-(3.8) it is clear that to obtain $B_n\to 0$, $n\to\infty$, a.s., it remains to prove that

$$\sigma_n^2 \to \sigma^2,\ n\to\infty,\ \text{a.s.}$$

To verify this, note that

$$\sigma_n^2-\sigma^2 = D_n+E_n$$

where

$$D_n = \int\int J(H_F(y))J(H_F(z))[(H_n(y)\wedge H_n(z)-H_n(y)H_n(z))-(H_F(y)\wedge H_F(z)-H_F(y)H_F(z))]dydz$$

$$E_n = \int\int J(H_n(y))J(H_n(z))-J(H_F(y))J(H_F(z))][H_n(y)\wedge H_n(z)-H_n(y)H_n(z)]dydz.$$

Since J is bounded we obtain $D_n\to 0$, $n\to\infty$, a.s. by showing that

$$\int\int|(H_n(y)\wedge H_n(z)-H_n(y)H_n(z))-(H_F(y)\wedge H_F(z)-H_F(y)H_F(z))|dydz\to 0 \tag{3.9}$$

as $n\to\infty$, a.s. To check (3.9) note that by the strong law for U-statistics

$$H_n(y)H_n(z)\to H_F(y)H_F(z),\ n\to\infty,\ a.s.$$

and, using the inequality $u\wedge v-uv\leq(u(1-u))^{1/2}(v(1-v))^{1/2}$, for $0<u,v<1$, that the integrand in (3.9) can be bounded by

$$(H_n(y)(1-H_n(y)))^{1/2}(H_n(z)(1-H_n(z))^{1/2}+(H_F(y)(1-H_F(y)))^{1/2}(H_F(z)(1-H_F(z)))^{1/2}.$$

The strenghtened Glivenko-Cantelli theorem for the empirical distribution function of U-statistic structure (Theorem 2.2 in HELMERS, JANSSEN and SERFLING (1988)) implies for every $\eta>0$ the existence of a natural number n_0, depending only on η and the particular realization, such that for all $n\geq n_o$ and all $y\in\mathbb{R}$

$$H_n(y)\leq H_F(y) + \eta q(H_F(y)) \tag{3.10}$$

$$1-H_n(y)\leq 1-H_F(y) + \eta q(H_F(y)) \tag{3.11}$$

where $q(t)=(t(1-t))^{1-2\epsilon}, 0<t<1$ and $\epsilon=\delta/2(4+\delta)$. Note that the condition on q required in Theorem 2.2 in HELMERS, JANSSEN and SERFLING (1988) is satisfied. Further note that Condition (ii) and a result similar to Lemma 2.2.1 in HELMERS (1982) implies

$$\int H_F(y)(1-H_F(y))dy<\infty,\ \int q(H_F(y))dy<\infty \text{ and } \int q^{1/2}(H_F(y))dy<\infty \tag{3.12}$$

From (3.10) and (3.11) it is easily seen that, for a particular realization and the n_0 just defined, we have for $n \geqslant n_0$

$$H_n(y)(1-H_n(y)) \leqslant H_F(y)(1-H_F(y)) + \eta^{1/2} q^{1/2}(H_F(y)) + \eta q(H_F(y)) \qquad (3.13)$$

By (3.12) we have integrability of the r.h.s of (3.13), so we can apply Lesbesque's dominated convergence theorem to get the a.s. convergence to zero of D_n.
To show that $E_n \to 0, n \to \infty$, a.s. note that (use condition (i) and arguments as used above to deal with D_n) $|E_n|$ is bounded by

$$2MK \sup_y |H_n(y) - H_F(y)| \int\int (H_n(y)(1-H_n(y)))^{1/2} (H_n(z)(1-H_n(z)))^{1/2} dydz.$$

From the discussion concerning D_n we know that the integral in the upper bound can be bounded. Therefore $E_n \to 0, n \to \infty$, a.s. since $\sup_y |H_n(y) - H_F(y)| \to 0, n \to \infty$, a.s. (see Corollary 2.1 in HELMERS, JANSSEN and SERFLING (1988)). So we have established that $\sigma_n^2 \to \sigma^2, n \to \infty$, a.s. which was the remaining convergence needed to obtain the a.s. convergence of B_n to zero.

To complete the proof of (3.1) we must check that $C_n \to 0, n \to \infty$, a.s. Since

$$\sup_x |\Phi(x\sigma^{-1}) - \Phi(x\sigma_n^{-1})| \leqslant \left(\frac{\sigma}{\sigma_n} \vee \frac{\sigma_n}{\sigma}\right) - 1,$$

the proof follows from the a.s. convergence of σ_n^2 to σ^2. The proof of (3.2) is similar, now starting with

$$\sup_x |\tilde{F}_{T_n}(x) - \tilde{F}^*_{T_n}(x)| \qquad (3.14)$$
$$\leqslant \sup_x |\tilde{F}_{T_n}(x) - \Phi(x)| + \sup_x |\tilde{F}^*_{T_n}(x) - \Phi(x)|.$$

This completes the proof of the theorem. □

4. A REFINEMENT

Going through the proof of relation (3.2) it is easy to see that the result can be strengthened to $\sup_x |\tilde{F}_{T_n}(x) - \tilde{F}^*_{T_n}(x)| = 0(n^{-1/2})$ a.s. provided that we replace Condition (ii) in Theorem 2 by

$$\max_{1 \leqslant i_1 \leqslant \cdots \leqslant i_m \leqslant m} E|h(X_{i_1}, \ldots, X_{i_m})|^3 < \infty.$$

We need only apply Theorem 1 to each of the terms appearing in the r.h.s. of (3.14) and argue as in the proof given for (3.1). A similar a.s. order bound for

$$\sup |F_{T_n}(x) - F^*_{T_n}(x)|$$

requires in view of (3.3) an investigation of the a.s. rate at which σ_n^2 tends to σ^2. In any case we expect an a.s. rate $0(n^{-1/2})$ under an appropriate set of moment conditions. We shall not pursue this point here.

Finally we note that better a.s. rates, typically of the order $o(n^{-1/2})$, for the accuracy of the bootstrap approximation can be obtained when estimating a suitable *studentized* generalized L-statistic rather than normalized generalized L-statistics as studied in the present paper. Results of this type have been established by BABU and SINGH (1983) for studentized smooth functions of sums of i.i.d. vectors and by HELMERS (1987) for studentized U-statistics. Extensions of the present results to the case of studentized generalized L-statistics will be developed in a separate paper.

REFERENCES

G.J. BABU and K. SINGH (1983). Inference on means using the bootstrap. *Ann. Statist. 11,* 999-1003.
P.J. BICKEL and D. FREEDMAN (1981). Some asymptotic theory for the bootstrap. *Ann. Statist. 9,* 1196-1217.

J. BRETAGNOLLE (1983). Lois limites du bootstrap de certaines fonctionelles. *Ann. Inst. H. Poincaré 19,* 281-286.

D. BOOS, P. JANSSEN and N. VERAVERBEKE (1988). Resampling from centered data in the two-sample problem. *J. Statist. Planning Inf. 19* (to appear).

R. HELMERS (1981). A Berry-Esseen theorem for linear combinations of order statistics. *Ann. Probability 9,* 342-347.

R. HELMERS (1982). Edgeworth expansions for linear combinations of order statistics. *Math. Centre Tracts 105,* Amsterdam.

R. HELMERS (1987). On the Edgeworth expansion and the bootstrap approximation for a studentized U-statistic. *Report MS-R8708, Centre for Mathematics and Computer Science,* Amsterdam.

R. HELMERS and F.H. RUYMGAART (1988). Asymptotic normality of generalized L-statistics with unbounded scores. *J. Statist. Planning Inf. 19,* 43-53.

R. HELMERS, P. JANSSEN and R. SERFLING (1988). Glivenko-Cantelli properties of some generalized empirical df's and strong convergence of generalized L-statistics. *Prob. Theory and Rel Fields* (to appear).

P. JANSSEN, R. SERFLING and N. VERAVERBEKE (1984). Asymptotic normality for a general class of statistical functions and application to measure of spread. *Ann. Statist. 12,* 1369-1379.

V.V. PETROV (1975). Sums of independent random variables. Springer-Verlag, Berlin.

R. SERFLING (1980). Approximation theorems of mathematical statistics. J. Wiley, New York.

R. SERFLING (1984). Generalized L-, M- and R-statistics. *Ann Statist. 12,* 76-86.

B. SILVERMAN (1983). Convergence of a class of empirical distribution functions of dependent random variables. *Ann. Probability 11,* 745-751.

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

W.R. VAN ZWET (1984). A Berry-Esseen bound for symmetric statistics. *Z. Wahr. verw. Gebiete 66, 425-440.*