Scand J Statist 17: 65–77, 1990

Berry-Esséen and Bootstrap Results for Generalized L-Statistics

R. HELMERS
Centre for Mathematics and Computer Science, Amsterdam

P. JANSSEN
Limburgs Universitair Centrum, Diepenbeek

R. SERFLING
Johns Hopkins University, Baltimore

ABSTRACT. 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 asymptotic accuracy of the bootstrap approximation of generalized L-statistics is also obtained. Studentized versions of the above results are also considered.

Key words: bootstrap, generalized L-statistics, U-statistics, Berry-Esséen theorem, rate of convergence, empirical measures.

1. Introduction

Let $X_1, \ldots, X_n$ be independent random variables having common probability distribution F. Let $h(x_1, \ldots, x_m)$ be a kernel of degree m (i.e. a real-valued measurable function symmetric in m arguments), and let

$$W_{n,1} \leq W_{n,2} \leq \ldots \leq W_{n,\binom{n}{m}}$$

denote the ordered evaluations $h(X_{i_1}, \ldots, X_{i_m})$ taken over the $\binom{n}{m}$ m-tuples in $C_{n,m} = \{(i_1, \ldots, i_m): 1 \leq i_1 < \ldots < i_m \leq n\}$. Many statistics of interest can be represented in the form

$$T_n = \sum_{i=1}^{\binom{n}{m}} C_{ni} W_{n,i} \tag{1.1}$$

for a suitable choice of h and constants C_{ni} generated by some weight function J on $(0, 1)$, i.e.,

$$C_{ni} = \int_{(i-1)/\binom{n}{m}}^{i/\binom{n}{m}} J(t)\, dt.$$

Let H_F denote the *df* of the random variable $h(X_1, \ldots, X_m)$ and let $H_F(y)$ be estimated by

$$H_n(y) = \binom{n}{m}^{-1} \sum_{C_{n,m}} 1\{h(X_{i_1}, \ldots, X_{i_m}) \leq y\}, \; y \in \mathbb{R},$$

the empirical distribution function of U-statistic structure. Since $T_n = T(H_n)$, with $T(\cdot)$ an

L-functional

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

with $G^{-1}(t)=\inf\{y:G(y)\geq t\}$ for any d.f. G, T_n is a generalized L-statistic estimating the parameter

$$\tau=T(H_F).$$

The form (1.1) is quite general and provides a unifying concept to various classes of statistics. For $J\equiv 1$, T_n becomes the U-statistic based on the kernel h. For $m=1$ and $h(x)=x$, T_n is an ordinary L-statistic. Further specific examples, covered by our results, are generalized L-statistics with smoothly trimmed weight functions, introduced by Stigler (1973) for the case $m=1$ and $h(x)=x$. A further choice is $J(t)=6t(1-t)$, which in the case of ordinary L-statistics provides an efficient estimator of location in the case of a logistic distribution.

Under various sets of regularity conditions on J and H_F, asymptotic normality for generalized L-statistics has been recently investigated. Primary references are Serfling (1984), Silverman (1983), Helmers & Ruymgaart (1988) and Gijbels, Janssen & Veraverbeke (1988). The basic asymptotic normality result for T_n is (e.g. see (3.7) in Helmers & Ruymgaart (1988)):

$$n^{1/2}(T_n-\tau)\xrightarrow{d}\mathcal{N}\{0,\sigma^2(T,H_F)\} \tag{1.2}$$

where

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

Note that given J and G, the quantity $\sigma^2(T,G)$ is readily computed.

In applications one often wishes to establish a confidence interval for τ and a studentized version of (1.2) is required. The desired result is easily obtained by showing (see section 2) that, with $\sigma_n^2=\sigma^2(T,H_n)$,

$$\sigma_n\to\sigma(T,H_F),\quad n\to\infty,\quad \text{a.s. } [P] \tag{1.3}$$

where P denotes probability under F. Together (1.2) and (1.3) directly yield an approximate two-sided confidence interval

$$(T_n-\sigma_n n^{-1/2}u_{\alpha/2},\ T_n+\sigma_n n^{-1/2}u_{\alpha/2}) \tag{1.4}$$

for τ based on the normal approximation. Here $u_{\alpha/2}=\Phi^{-1}(1-\alpha/2)$ and Φ denotes the standard normal distribution.

A different approach for construction of confidence intervals for τ is based on the bootstrap approximation for the d.f. of $n^{1/2}(T_n-\tau)$ or $n^{1/2}\sigma_n^{-1}(T_n-\tau)$. With F_n the empirical d.f., i.e., $F_n(x)=n^{-1}\sum_{i=1}^n 1\{X_i\leq x\}$, the bootstrap method uses resampling with replacement from the observations $X_1,\ldots,X_n$. Conditionally, given $X_1,\ldots,X_n$, we obtain in each resample a collection of random variables $X_1^*,\ldots,X_n^*$ which are conditionally independent with common distribution F_n. For any $n\geq m$, we define

$$H_n^*(y)=\binom{n}{m}^{-1}\sum_{C_{n,m}}1\{h(X_{i_1}^*,\ldots,X_{i_m}^*)\leq y\},\quad y\in\mathbf{R},$$

the empirical d.f. of U-statistic structure based on the bootstrap sample $X_1^*,\ldots,X_n^*$. With

$T_n^*=T(H_n^*)$ and $\sigma_n^*=\sigma(T,H_n^*)$, define for $n\geq m$ and $x\in\mathrm{R}$,

$$F_{T_n}(x)=P\{n^{1/2}(T_n-\tau)\leq x\}$$

$$F_{T_n}^*(x)=P^*\{n^{1/2}(T_n^*-T_n)\leq x\}$$

$$G_{T_n}(x)=P\{n^{1/2}(T_n-\tau)/\sigma_n\leq x\}$$

$$G_{T_n}^*(x)=P^*\{n^{1/2}(T_n^*-T_n)/\sigma_n^*\leq x\}$$

where P^* denotes probability under F_n. In section 4 we show that a.s. $[P]$

$$\lim_{n\to\infty}\sup_x|F_{T_n}(x)-F_{T_n}^*(x)|=0 \tag{1.5}$$

and

$$\lim_{n\to\infty}\sup_x|G_{T_n}(x)-G_{T_n}^*(x)|=0. \tag{1.6}$$

On the basis of these results one easily establishes approximate two-sided confidence intervals for τ:

$$(T_n-n^{-1/2}C_{1-\alpha/2}^*,\ T_n-n^{-1/2}C_{\alpha/2}^*) \tag{1.7}$$

and

$$(T_n-n^{-1/2}\sigma_n C_{1-\alpha/2}^{*s},\ T_n-n^{-1/2}\sigma_n C_{\alpha/2}^{*s}) \tag{1.8}$$

where C_p^* and C_p^{*s} denote the p-th quantile of the bootstrap d.f. $F_{T_n}^*$ and $G_{T_n}^*$. These bootstrap based confidence intervals for τ provide alternatives to the more familiar confidence interval (1.4) based on the normal approximation.

To prove the validity of the bootstrap approximations (1.5) and (1.6) we assume J Lipschitz and we rely on a Berry-Esséen rate associated with the convergence in (1.2). Our proof resembles that of Singh (1981), who employs the classical Berry-Esséen theorem as his main tool to establish the asymptotic validity of the bootstrap approximation for the d.f. of the sample mean. Similarly, our proof will rely on a Berry-Esséen result for generalized L-statistics. Also instrumental in our proof will be certain Glivenko-Cantelli results for the empirical d.f. H_n, established in Helmers, Janssen & Serfling (1988), and parallel results for H_n^*.

In section 2 we derive strong laws for σ_n and σ_n^*; the Glivenko-Cantelli result for H_n^*, which we will also need, is proved in the appendix. In section 3 we obtain a Berry-Esséen bound for generalized L-statistics. Our bootstrap results (1.5) and (1.6) are derived in section 4. Some refinements and possible extensions are briefly discussed in section 5.

We conclude this section by noting that the bootstrap results (1.5) and (1.6) extend the work of Bickel & Freedman (1981) on the asymptotic validity of bootstrap approximations for the distribution function of non-degenerate U-statistics and L-statistics. Further relevant background includes Bretagnolle (1983), treating von Mises statistics, and Boos, Janssen & Veraverbeke (1988), where resampling plans for two-sample U-statistics with estimated parameters are studied.

2. Strong consistency for σ_n and σ_n^*

In this section we show that σ_n and σ_n^* are consistent estimators of σ. Recall that $\sigma_n=\sigma(T,H_n)$, $\sigma_n^*=\sigma(T,H_n^*)$ and $\sigma=\sigma(T,H_F)$.

Theorem 1
Suppose that

(*i*) *J is bounded on* $(0, 1)$;
(*ii*) $E|h(X_1, \ldots, X_m)|^{2+\delta}<\infty$, *for some* $\delta>0$;
(*iii*) $\sigma^2=\sigma^2(T, H_F)>0$.

Then

$$\sigma_n \to \sigma,\ n\to\infty,\ \text{a.s. } [P] \tag{2.1}$$

and, with P-probability 1,

$$\sigma_n^* \to \sigma,\ n\to\infty,\ \text{a.s. } [P^*]. \tag{2.2}$$

Proof. To prove (2.1), note that

$$\sigma_n^* - \sigma^2 = A_n + B_n \tag{2.3}$$

where

$$A_n = m^2 \int\int J\{H_F(y)\}J\{H_F(z)\}[\{H_n(y)\wedge H_n(z) - H_n(y)H_n(z)\} -$$

$$\{H_F(y)\wedge H_F(z) - H_F(y)H_F(z)\}]\ dy\ dz$$

$$B_n = m^2 \int\int [J\{H_n(y)\}J\{H_n(z)\} - J\{H_F(y)\}J\{H_F(z)\}][H_n(y)\wedge H_n(z) -$$

$$H_n(y)H_n(z)]\ dy\ dz.$$

Since J is bounded we obtain $A_n \to 0$, $n\to\infty$, a.s. $[P]$ by showing that

$$\int\int |\{H_n(y)\wedge H_n(z) - H_n(y)H_n(z)\} - \{H_F(y)\wedge H_F(z) - H_F(y)H_F(z)\}|\ dy\ dz \to 0, \tag{2.4}$$

$n\to\infty$, a.s. $[P]$. To check (2.4) note that by the strong law for U-statistics

$$H_n(y)H_n(z) \to H_F(y)H_F(z),\quad n\to\infty,\quad \text{a.s. } [P]$$

and, using the inequality $u\wedge v - uv \leqslant \{u(1-u)\}^{1/2}\{v(1-v)\}^{1/2}$, for $0<u, v<1$, that the integrand in (2.4) 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 strengthened Glivenko-Cantelli theorem for the empirical d.f. of U-statistic structure (Theorem 2.2 in Helmers, Janssen & 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 $n\geqslant n_0$ and all $x\in \mathrm{R}$,

$$H_n(y) \leqslant H_F(y) + \eta q\{H_F(y)\} \tag{2.5}$$

$$1 - H_n(y) \leqslant 1 - H_F(y) + \eta q\{H_F(y)\} \tag{2.6}$$

where $q(t)=\{t(1-t)\}^{1-2\varepsilon}$, $0<t<1$ and $\varepsilon=\delta/2(4+\delta)$. Note that the condition on q required in theorem 2.2 in Helmers, Janssen & Serfling (1988) is satisfied. Further note that Condition (ii) and a result similar to lemma 2.2.1 in Helmers (1982) implies

$$\int H_F(y)\{1-H_F(y)\}\,dy<\infty,\quad \int q\{H_F(y)\}\,dy<\infty \quad\text{and}\quad \int q^{1/2}\{H_F(y)\}\,dy<\infty. \tag{2.7}$$

From (2.5) and (2.6) it is easily seen that, for a particular realization and the n_0 just defined, we have for $n\geq n_0$

$$H_n(y)\{1-H_n(y)\}\leq H_F(y)\{1-H_F(y)\}+\eta^{1/2}q^{1/2}\{H_F(y)\}+\eta q\{H_F(y)\}. \tag{2.8}$$

By (2.7) we have integrability of the r.h.s. of (2.8), so we can apply Lebesgue's dominated convergence theorem to get the a.s. $[P]$ convergence to zero of A_n.

To show that $B_n\to 0$, $n\to\infty$, a.s. $[P]$ note that (use Condition (i) and arguments as used above to deal with A_n) it is sufficient to show the a.s. $[P]$ convergence to zero of

$$\sup_y |H_n(y)-H_F(y)| \int\int [H_n(y)\{1-H_n(y)\}]^{1/2}[H_n(z)\{1-H_n(z)\}]^{1/2}\,dy\,dz. \tag{2.9}$$

From the discussion concerning A_n we know that that the integral in (2.9) can be bounded. Therefore $B_n\to 0$, $n\to\infty$, a.s. $[P]$ since

$$\sup_y |H_n(y)-H_F(y)|\to 0,\quad n\to\infty,\quad \text{a.s. } [P]$$

(see corollary 2.1 in Helmers, Janssen & Serfling (1988)). So we have established that $\sigma_n^2\to\sigma^2$, $n\to\infty$, a.s. $[P]$, which proves (2.1)

Next we consider (2.2). To prove this we follow the argument leading to (2.1), with some modifications. First replace H_n by H_n^* in the defining equations of A_n and B_n and denote the resulting expressions by A_n^* and B_n^*. We first prove that, with P-probability 1, $A_n^*\to 0$, $n\to\infty$, a.s. $[P^*]$. To check this we employ the strong law of large numbers for bootstrapped U-statistics, (Athreya *et al.* (1984)) and we need a bootstrap version of theorem 2.2. in Helmers, Janssen & Serfling (1988). More precisely, we must show that, for any q satisfying the assumptions of that theorem, with P-probability 1,

$$\|(H_n^*-H_F)/q\circ H_F\|_\infty\to 0,\quad n\to\infty,\quad \text{a.s. } [P^*] \tag{2.10}$$

where $\|f\|_\infty$ denotes $\sup_x|f(x)|$. A proof of this assertion is in the appendix. Combining the preceding results with the argument leading to (2.1) we see that, with P-probability 1, $A_n^*\to 0$, $n\to\infty$, a.s. $[P^*]$. Verification finally that, with P-probability 1, also $B_n^*\to 0$, $n\to\infty$, a.s. $[P^*]$ is straightforward. We follow the proof given for B_n and employ the modifications given above. This completes the proof of (2.2). □

3. Berry-Esséen 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 $\theta=Eh(X_1,\dots,X_m)$,

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

and define for $n \geq m$

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

Finally recall that $\Phi(x)$ denotes the standard normal distribution.

Theorem 2
Suppose that

(*i*) *J is Lipschitz of order* 1;
(*ii*) $E|g(X_1)|^3 < \infty$ *and* $Eh^2(X_1, \ldots, X_m) < \infty$;
(*iii*) $\sigma^2 = \sigma^2(T, H_F) > 0$.

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

$$\sup_x |\tilde{F}_{T_n}(x) - \Phi(x)| \leq Cn^{-1/2}$$

$$\left[\frac{M^3}{\sigma^3}\{E|g|^3 + (E|h|)^3\} + \frac{Eh^2}{\sigma^2}(2m-1)^2(M^2 + \gamma_m K^2 n^{-1/2}) + 1 + \frac{KE|h|}{\sigma}\right] \tag{3.1}$$

where γ_m is a constant depending only on m.

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

$$T(G_2) - T(G_1) = -\int [\psi\{G_2(y)\} - \psi\{G_1(y)\}]\,dy \tag{3.2}$$

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

$$|\psi\{G_2(y)\} - \psi\{G_1(y)\} - (G_2(y) - G_1(y))J\{G_1(y)\}| \leq K\{G_2(y) - G_1(y)\}^2. \tag{3.3}$$

From (3.2) and (3.3) follows

$$|T_n - \tau + \int \{H_n(y) - H_F(y)\}J\{H_F(y)\}\,dy| \leq K\int \{H_n(y) - H_F(y)\}^2\,dy. \tag{3.4}$$

Now note that

$$-\int \{H_n(y) - H_F(y)\}J\{H_F(y)\}\,dy = U_n \tag{3.5}$$

where

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

with

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

For the r.h.s. of (3.4) we have

$$K\int \{H_n(y) - H_F(y)\}^2\,dy = W_n + R_n \tag{3.6}$$

with

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

and

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

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

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

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

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

with

$$h_{2n}(x_{k_1},\ldots,x_{k_{2m}})=\binom{n}{2m}\binom{n}{m}^{-2}\sum{}^{+}\, h_2(x_{i_1},\ldots,x_{i_m},x_{j_1},\ldots,x_{j_m}) \tag{3.7}$$

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 $\{i_1,\ldots,i_m,j_1,\ldots,j_m\{=\{k_1,k_2,\ldots,k_{2m}\}$. For the second term R_n in the r.h.s. of (3.6) we will exploit that its magnitude is of lower order.

As in Helmers (1981), see also Helmers (1982), we now use (3.4)–(3.6) 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

$$\begin{aligned}&P\{n^{1/2}\sigma^{-1}(U_n+W_n+ER_n)\leq x-n^{-1/2}\}+P\{|R_n-ER_n|\geq\sigma n^{-1}\}\leq\tilde{F}_{T_n}(x)\\&\quad\leq P\{n^{1/2}\sigma^{-1}(U_n-W_n-ER_n)\leq x+n^{-1/2}\}+P\{|R_n-ER_n|\geq\sigma n^{-1}\}.\end{aligned}\tag{3.8}$$

Since both U_n-W_n and U_n+W_n are U-statistics of degree $2m$ with varying kernels $h_n^{\pm}=h_{1n}\pm h_{2n}$ where h_{2n} is defined by (3.7) and

$$h_{1n}(x_{k_1},\ldots,x_{k_{2m}})=\binom{n}{2m}\binom{n}{m}^{-1}\binom{n-m}{m}^{-1}\sum{}^{*}1/2[h_1(x_{i_1},\ldots,x_{i_m})+h_1(x_{j_1},\ldots,x_{j_m})]$$

with Σ^+ as in (3.7), the Berry-Esséen result for U-statistics due to van Zwet (1984) applies. So, using also the inequality,

$\sup_x|\Phi(x+q)-\Phi(x)|\leq|q|$ we get for $U_n-W_n-ER_n$

$$\begin{aligned}&\sup_x|P\{n^{1/2}\sigma^{-1}(U_n-W_n-ER_n)\leq x+n^{-1/2}\}-\Phi(x)|\\&\leq\sup_x|P\{n^{1/2}\sigma^{-1}(U_n-W_n)\leq x\}-\Phi(x)|+\sup_x|\Phi(x+n^{-1/2}+n^{1/2}\sigma^{-1}ER_n)-\Phi(x)|\\&\leq Cn^{-1/2}\left[\frac{E|g_n(X_1)|^3}{\{Eg_n^2(X_1)\}^{3/2}}+\frac{(2m-1)^2E\{h_n^-(X_1,\ldots,X_{2m})\}^2}{Eg_n^2(X_1)}\right]+n^{-1/2}(1+\sigma^{-1}|ER_n|)\end{aligned}\tag{3.9}$$

where

$$g_n(x)=\int \dots \int h_n^-(x, x_2, \dots, x_{2m})\, dF(x_2) \dots dF(x_{2m}). \tag{3.10}$$

The discussion just given deals with the first term of the upper bound for $\tilde{F}_{T_n}(x)$ given in (3.8). 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

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

Hence h_n^+ and h_n^- have the same projection.

Now it can be shown, by some elementary calculations, that

$$|ER_n| \leqslant 2K \frac{m}{n-m+1} E|h(X_1, \dots, X_m)| \tag{3.11}$$

$$E|g_n(X_1)|^3 \leqslant C_1 M^3 \{E|g(X_1)|^3 + (E|h(X_1, \dots, X_m)|)^3\} \tag{3.12}$$

$$E\{h_n^{\pm}(X_1, \dots, X_{2m})\}^2 \leqslant C_2(M^2+K^2) Eh^2(X_1, \dots, X_m). \tag{3.13}$$

Finally note that

$$Eg_n^2(X_1)=\sigma^2(T, H_F). \tag{3.14}$$

The latter equality follows since h_{1n} is obtained by rewriting the kernel h_1 of U_n in 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 (3.9) to (3.14) the appropriate bound for the first term in the r.h.s. of (3.8) is obtained.

It remains to show that the second term in the r.h.s. of (3.8) is of the right order. To verify this note that

$$P\{|R_n - ER_n| \geqslant \sigma n^{-1}\} \leqslant n^2 \sigma^{-2} \operatorname{var} R_n. \tag{3.15}$$

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

$$\operatorname{var} R_n \leqslant K^2 \frac{\gamma_m}{n^3} Eh^2(X_1, \dots, X_m), \tag{3.16}$$

where γ_m is a constant depending only on m. From (3.15) and (3.16) we get that

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

The following extension of theorem 2 provides the appropriate order bound for application in section 4.

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* $0<\delta\leq 1$. *Then there exists a positive constant* C_δ, *depending only on* δ, *such that for all* $n\geq 2m$

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

$$\leq C_\delta n^{-\delta/2}\left[\frac{M^{2+\delta}}{\sigma^{2+\delta}}\{E|g|^{2+\delta}+(E|h|)^{2+\delta}\}+\frac{Eh^2}{\sigma^2}(2m-1)^2(M^2+\gamma_m K^2 n^{(\delta-2)/2})+1+\frac{KE|h|}{\sigma}\right], \tag{3.17}$$

with M, K *and* γ_m *as in theorem 2.*

Proof. First note that only minor changes are needed to obtain the appropriate modification of the Berry-Esséen 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 (3.9). Finally note that the inequality (3.12) is now replaced by

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

with C_3 a positive constant depending only on δ.

4. Bootstrapping generalized L-statistics

Let T_n, τ, $\sigma^2(T,G)$ and F_{T_n}, $F^*_{T_n}$, G_{T_n}, $G^*_{T_n}$ denote the quantities as defined in section 1 and $\tilde{F}_{T_n}$ as in section 3. Also define, with $T^*_n=T(H^*_n)$ and for $n\geq m$ and $x\in\mathrm{R}$,

$$\tilde{F}^*_{T_n}(x)=P^*\{n^{1/2}(T^*_n-T_n)/\sigma_n\leq x\}.$$

Our main result is the following theorem.

Theorem 3
Suppose that

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

Then with P-probability 1,

$$\lim_{n\to\infty}\sup_x|\tilde{F}_{T_n}(x)-\tilde{F}^*_{T_n}(x)|=0 \tag{4.1}$$

$$\lim_{n\to\infty}\sup_x|F_{T_n}(x)-F^*_{T_n}(x)|=0 \tag{4.2}$$

$$\lim_{n\to\infty}\sup_x|G_{T_n}(x)-G^*_{T_n}(x)|=0. \tag{4.3}$$

Remark. As indicated in the introduction (see (1.5)–(1.8)) the bootstrap results (4.2) and (4.3) may be used to establish bootstrap confidence intervals for τ based on $n^{1/2}(T_n-\tau)$ or $n^{1/2}(T_n-\tau)/\sigma_n$. The bootstrap result (4.1) is of little practical value: computation of a bootstrap confidence interval for τ based on $n^{1/2}(T_n-\tau)/\sigma$ would require *a priori* knowledge of

$\sigma=\sigma(T, H_F)$ which is usually not available. On the other hand relation (4.1) is the *bootstrap counterpart* to the Berry-Esséen result (3.1).

Proof of theorem 3. We first prove (4.2). Note that

$$\sup_x |F_{T_n}(x)-F^*_{T_n}(x)|\leq C_n+D_n+E_n$$

where $C_n=\sup_x|F_{T_n}(x)-\Phi(x\sigma^{-1})|$, $D_n=\sup_x|F^*_{T_n}(x)-\Phi(x\sigma_n^{-1})|$, with $\sigma_n^2=\sigma^2(T, H_n)$, and $E_n=\sup_x|\Phi(x\sigma^{-1})-\Phi(x\sigma_n^{-1})|$. The term C_n is not random, whereas D_n and E_n are random.

To see that $C_n\rightarrow 0$, $n\rightarrow\infty$, apply theorem 3.2 in Serfling (1984).

To show that $D_n\rightarrow 0$, $n\rightarrow\infty$, a.s. $[P]$ a more refined argument is needed. Let E^* denote expectation under F_n and define

$$\theta_n=E^*h(X_1^*, \ldots, X_m^*)=n^{-m}\sum_{i_1=1}^{n}\ldots\sum_{i_m=1}^{n} h(X_{i_1}, \ldots, X_{i_m}),$$

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

$$g_n^*(x)=\int\ldots\int h(x, x_2, \ldots, x_m)\, dF_n(x_2)\ldots dF_n(x_m)-\theta_n.$$

An application of corollary 1 yields

$$D_n\leq C n^{-\delta/2}\left\{\frac{E^*|g_n^*(X_1^*)|^{2+\delta}+(E^*|h(X_1^*, \ldots, X_m^*)|)^{2+\delta}}{\sigma_n^{2+\delta}}+\frac{E^*h^2(X_1^*, \ldots, X_m^*)}{\sigma_n^2}\right.$$
$$\left.+\frac{E^*|h(X_1^*, \ldots, X_m^*)|}{\sigma_n}\right\}, \tag{4.4}$$

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 (4.4) An application of Jensen's inequality for conditional expectations in combination with the inequality $|a-b|^{2+\delta}\leq 2^{2+\delta}(|a|^{2+\delta}+|b|^{2+\delta})$ yields

$$E^*|g_n^*(X_1^*)|^{2+\delta}\leq 2^{2+\delta}\{E^*|h(X_1^*, \ldots, X_m^*)|^{2+\delta}+|\theta_n|^{2+\delta}\}. \tag{4.5}$$

Clearly

$$E^*|h(X_1^*, \ldots, X_m^*)|^{2+\delta}=n^{-m}\sum_{i_1=1}^{n}\ldots\sum_{i_m=1}^{n}|h(X_{i_1}, \ldots, X_{i_m})|^{2+\delta}. \tag{4.6}$$

By condition (ii) the strong law of large numbers for von Mises statistics applies. Therefore the r.h.s. of (4.6) converges to $E|h(X_1, \ldots, X_m)|^{2+\delta}$ a.s. $[P]$ and

$$\theta_n\rightarrow\theta, \quad n\rightarrow\infty. \quad \text{a.s. } [P]. \tag{4.7}$$

From (4.5) to (4.7) 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

$$E^*|h(X_1^*, \ldots, X_m^*)|^r=n^{-m}\sum_{i_1=1}^{n}\ldots\sum_{i_m=1}^{n}|h(X_{i_1}, \ldots, X_{i_m})|^r\rightarrow E^*|h(X_1, \ldots, X_m)|^r,$$
$$n\rightarrow\infty, \quad \text{a.s. } [P]. \tag{4.8}$$

Note that for (4.8) condition (ii) becomes effective again.

From (4.4) to (4.8) it is clear that to obtain $D_n \to 0$, $n \to \infty$, a.s. [P] it remains to verify that $\sigma_n^2 \to \sigma^2$, $n \to \infty$, a.s. [P]. But the latter result was already obtained in theorem 1.

To complete the proof of (4.2) we must check $E_n \to 0$, $n \to \infty$, a.s. [P]. Since

$$\sup_x |\Phi(x\sigma^{-1}) - \Phi(x\sigma_n^{-1})| \leq \left(\frac{\sigma}{\sigma_n} \vee \frac{\sigma_n}{\sigma}\right) - 1 \tag{4.9}$$

the proof follows from the a.s. convergence of σ_n^2 to σ^2 (theorem 1).

The proof of (4.1) is similar, now starting with

$$\sup_x |\tilde{F}_{T_n}(x) - \tilde{F}^*_{T_n}(x)| \leq \sup_x |\tilde{F}_{T_n}(x) - \Phi(x)| + \sup_x |\tilde{F}^*_{T_n}(x) - \Phi(x)|. \tag{4.10}$$

To check (4.3), we note that $G_n(x) = F_{T_n}(x\sigma_n)$ and $G_n^*(x) = F^*_{T_n}(x\sigma_n^*)$. From the proof of (4.2) we infer that a.s. [P]

$$\sup_x |F_{T_n}(x\sigma_n) - \Phi(x\sigma_n\sigma^{-1})| \to 0, \quad n \to \infty$$

and

$$\sup_x |F^*_{T_n}(x\sigma_n^*) - \Phi(x\sigma_n^*\sigma^{-1})| \to 0, \quad n \to \infty.$$

An argument like (4.9) ensures that it remains to prove that $\sigma_n \to \sigma$, $n \to \infty$, a.s. [P], and that, with P-probability 1, $\sigma_n^* \to \sigma$, $n \to \infty$, a.s. [P^*]. Both these a.s. statements were already established in theorem 1. This completes the proof of the theorem. □

5. Refinements and possible extensions

Going through the proof of relation (4.1) it is easy to see that the result can be strengthened to

$$\sup_x |\tilde{F}_{T_n}(x) - \tilde{F}^*_{T_n}(x)| = 0(n^{-1/2}) \text{ a.s. } [P]$$

provided that we replace condition (ii) in theorem 3 by

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

We need only apply theorem 2 to each of the terms appearing in the r.h.s. of (4.10) and argue as in the proof given for (4.2). A similar a.s. order bound for

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

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

Better a.s. rates, typically of order $o(n^{-1/2})$, can be obtained for

$$\sup_x |G_{T_n}(x) - G^*_{T_n}(x)|.$$

Results of this type have been established by Babu & Singh (1983) for studentized smooth functions of sums of i.i.d. vectors and by Helmers (1990) for studentized U-statistics. Extensions of the present results to the case of studentized generalized L-statistics, i.e., to $n^{1/2}(T_n - \tau)/\sigma_n$, will be developed in a separate paper. Though the result by Helmers (1990) appears to be a crucial step, validating such an extension will be quite laborious, and therefore outside the scope of this paper.

Extending the results of Theorem 3 to the case of generalized L-statistics based on trimmed weightfunctions (hence J is not Lipschitz) is presently under investigation. It requires a different methodology based on bootstrap results for quantile processes of U-statistic structure.

6. Appendix: proof of (2.10)

To show (2.10) some changes have to be made in the proof of theorem 2.2 of Helmers, Janssen & Serfling (1988). With q as in that theorem and θ_0 as in their proof of the theorem the details are as follows. Their corollary 2.1 directly yields that, with P-probability 1, $\|H_n^* - H_{F_n}\|_\infty \to 0$, $n \to \infty$. a.s. $[P^*]$, where

$$H_{F_n}(y) = n^{-m} \sum_{i_1=1}^{n} \ldots \sum_{i_m=1}^{n} 1\{h(X_{i_1}, \ldots, X_{i_m}) \leq y\}, \quad y \in \mathrm{R}$$

We also require that $\|H_{F_n} - H_F\|_\infty \to 0$, $n \to \infty$, a.s. $[P]$. This Glivenko-Cantelli result is not explicitly stated in Helmers, Janssen & Serfling (1988), but it is easily obtained from their theorem 2.1 utilizing a representation of H_{F_n} as a linear combination of empirical distribution functions of U-statistic structure, with kernels of degree at most m, and with coefficients which are at most of order 1. Combining these results we directly see that, with P-probability 1, $\|H_n^* - H_F\|_\infty \to 0$, $n \to \infty$, a.s. $[P^*]$, i.e., we have proved (2.10) for the special case $q \equiv 1$. It is clear from theorem 2.2 in Helmers, Janssen & Serfling (1988) that to establish (2.10), the only missing ingredient is to check that, with P-probability 1,

$$\binom{n}{m}^{-1} \sum_{c_{n,m}} \frac{1\{h(X_{i_1}^*, \ldots, X_{i_m}^* \leq \theta_0\}}{q \circ H_F\{h(X_{i_1}^*, \ldots, X_{i_m}^*)\}} \to \int_{-\infty}^{\theta_0} \{q \circ H_F(y)\}^{-1} \, dH_F(y),$$

$n \to \infty$, a.s. $[P^*]$. But this follows directly from the strong law of large numbers for bootstrapped U-statistics (Athreya *et al.* (1984)) and the proof of (2.10) is complete. □

Acknowledgement

The research was supported by the U.S. Department of Navy under Office of Naval Research Contract No. N00014-79-C-0801 and N00014-88-K-0090 and NATO under Research Grant No. 0034/87.

The authors thank the referee and the associate editor for careful reading of the manuscript and for their constructive remarks.

References

Athreya, K. B., Ghosh, M., Low, L. Y. & Sen, P. K. (1984). Law of large numbers for bootstrapped U-statistics. *J. Statist. Plann. Inference* **9**, 185–194

Babu, G. J. & Singh, K. (1983). Inference on means using the bootstrap. *Ann. Statist.* **11**, 999–1003.

Bickel, P. J. & Freedman, D. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9**, 1196–1217.

Bretagnolle, J. (1983). Lois limites du bootstrap de certaines fonctionelles. *Ann. Inst. H. Poincaré*, Sec. B. **19**, 281–296.

Boos, D., Janssen, P. & Veraverbeke, N. (1989). Resampling from centered data in the two-sample problem. *J. Statist. Plann. Inference.* **21**, 326–345.

Gijbels, I., Janssen, P. & Veraverbeke, N. (1988). Weak and strong representations for trimmed U-statistics. *Probab. Theory Rel. Fields* **77**, 179–194.

Helmers, R. (1981). A Berry-Esséen theorem for linear combinations of order statistics. *Ann. Probab.* **9**, 342–347.
Helmers, R. (1982). Edgeworth expansions for linear combinations of order statistics. *Math. Centre Tracts* **105**, Amsterdam.
Helmers, R. (1990). On the Edgeworth expansion and the bootstrap approximation for a studentized *U*-statistic. *Ann. Statist.* (to appear).
Helmers, R. & Ruymgaart, F. H. (1988). Asymptotic normality of generalized *L*-statistics with unbounded scores. *J. Statist. Plann. Inference* **19**, 43–53.
Helmers, R. Janssen, P. & Serfling, R. (1988). Glivenko-Gantelli properties of some generalized empirical d.f.s and strong convergence of generalized *L*-statistics. *Probab. Theory Rel. Fields* **79**, 75–93.
Petrov, V. V. (1975). *Sums of independent random variables.* Springer, Berlin.
Serfling, R. (1980). *Approximation theorems of mathematical statistics.* J. Wiley, New York.
Serfling, R. (1984). Generalized *L*-, *M*- and *R*-statistics. *Ann. Statist.* **12**, 76–86.
Silverman, B. (1983). Convergence of a class of empirical distribution functions of dependent random variables. *Ann. Probab.* **11**, 745–751.
Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9**, 1187–1195.
Stigler, S. M. (1973). The asymptotic distribution of the trimmed mean. *Ann. Statist.* **1**, 462–477.
van Zwet, W. R. (1984). A Berry-Esséen bound for symmetric statistics. *Z. Wahrsch. Verw. Gebiete* **66**, 425–440.

Received February 1988, in final form November 1989

R. Helmers, Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica, Kruislaan 413, NL-1098 SJ Amsterdam, The Netherlands