Resampling Estimators for Generalized L-Statistics

by

Jun Shao
Purdue University

Technical Report #88-59

Department of Statistics
Purdue University

November 1988

RESAMPLING ESTIMATORS FOR GENERALIZED L-STATISTICS

Jun Shao

Department of Statistics

Purdue University

West Lafayette, IN 47907

ABSTRACT

A wide class of statistics, the generalized L-statistics, was introduced in Serfling (1984). The generalized L-statistics are asymptotically normal under weak conditions. This report consists of two parts. In part I, we show that the jackknife estimators of the asymptotic variances of generalized L-statistics are consistent. In part II, bootstrap methods for generalized L-statistics are studied. The results provide methods for large sample statistical analysis based on generalized L-statistics.

PART I

JACKKNIFE VARIANCE ESTIMATORS FOR GENERALIZED L-STATISTICS

1. Introduction

The generalized L-statistics was introduced by Serfling (1984). It generalizes the classes of U-statistics and L-statistics and consists of other types of statistics such as trimmed sample variance, trimmed U-statistics and Winsorized U-statistics. See Serfling (1984, 1985) for other examples. Let $X_1,..., X_n$ be independent and identically distributed samples from an unknown population distribution F, m be a fixed positive integer and $h(x_1,...,x_m)$ be a given symmetric function. Denote the distribution function of $h(X_1,...,X_m)$ by $H(y)$, i.e.,

$$H(y) = P_F\{ h(X_1,...,X_m) \leq y\}, \quad y \in \mathbf{R}.$$

Let

$$H_n(y) = n_{(m)}^{-1} \sum_{C_m} I[h(X_{i_1},...,X_{i_m}) \leq y], \tag{1.1}$$

where $I[A]$ is the indicator function of the set A, $n_{(m)}=n(n-1) \cdots (n-m+1)$ and $\sum_{C_m}$ is the summation taken over the $n_{(m)}$ m-tuples $(i_1,...,i_m)$ of distinct elements from $\{ 1, \ldots, n \}$. We consider a class of smooth generalized L-statistics defined by $T(H_n)$, where T is defined to be

$$T(G) = \int yJ[G(y)]dG(y), \quad \text{for any distribution function } G, \tag{1.2}$$

and J is a function on [0,1] (Serfling, 1984). When $h = x$, H_n reduces to the ordinary empirical distribution and $T(H_n)$ reduces to the ordinary L-statistics. When $J \equiv 1$, $T(H_n)$ is a U-statistic. It was shown in Serfling (1984) that the influence function of $T(H_n)$ is

$$\phi(z) = -m\int[g(y,z)-H(y)]J[H(y)]dy,$$

where

$$g(y,z) = \int \cdots \int I[h(x_1,...,x_{m-1},z) \leq y].$$

Furthermore, under either condition A or condition B stated below, the generalized L-statistics are asymptotically normal, i.e.,

$$n^{1/2}[T(H_n) - T(H)] \to N(0, \sigma^2) \quad \text{in distribution},$$

where $\sigma^2 = E_F \phi^2(X_1)$ and is assumed to be finite.

Condition A. $J(t) = 0$ for $0 \le t < \alpha$ or $\beta < t \le 1$, where $0<\alpha<\beta<1$ are constants, J is continuous on $[\alpha,\beta]$ and H is continuous.

Condition B. J is continuous on $[0,1]$ and H is continuous and satisfies

$$\int [H(y)(1-H(y))]^{1/2} dy < \infty. \tag{1.3}$$

The statistics with J functions satisfying condition A are referred to as trimmed statistics in the literature and they usually provide robust estimators (Huber, 1981). Condition (1.3) is equivalent to $E_F h^2(X_1,...,X_m) < \infty$ if the distribution H has regularly varying tails (see Feller, 1966, p.268) with a finite exponent. It is implied by $E_F |h(X_1,...,X_m)|^{2+\delta} < \infty$ for a $\delta>0$.

For various purposes in statistical analysis, we need a consistent estimator of the unknown asymptotic variance σ^2. In this paper, we prove that the estimators of σ^2 obtained by using the jackknife method (Quenouille, 1956; Tukey, 1958) are strongly consistent. For $i=1,...,n$, let H_{ni} be defined as in (1.1) corresponding to $n-1$ samples $X_1,..., X_{i-1}, X_{i+1},..., X_n$. The jackknife estimator of σ^2 is defined to be

$$s_J^2 = (n-1)\sum_{i=1}^n [T(H_{ni}) - \bar{T}_n]^2, \tag{1.4}$$

where $\bar{T}_n = n^{-1}\sum_{i=1}^n T(H_{ni})$.

In Section 2, the strong consistency of s_J^2 is proved for trimmed generalized L-statistics. The case of untrimmed generalized L-statistics is treated in Section 3. Since U- and L-statistics are special cases of generalized L-statistics, our result includes the existing results in jackknifing U- and L-statistics (see

Arvesen, 1969; Parr and Schucany, 1982) as special cases.

2. Trimmed generalized L-statistics

Let U_n be a U-statistic (see Hoeffding, 1948) defined to be

$$U_n = n_{(m)}^{-1} \sum_{C_m} k(X_{i_1},...,X_{i_m}),$$

where $k(x_1,...,x_m)$ is a symmetric kernel. For each i, let

$$U_{ni} = [(n-1)_{(m)}]^{-1} \sum_{C_m^i} k(X_{i_1},...,X_{i_m}),$$

where $(n-1)_{(m)} = (n-1) \cdots (n-m)$ and $\sum_{C_m^i}$ is the summation taken over the $(n-1)_{(m)}$ m-tuples $(i_1,...,i_m)$ of distinct elements from the integers $\{1, \ldots, i-1, i+1, \ldots, n\}$. The jackknife estimator of the asymptotic variance of U_n is

$$s_U^2 = (n-1) \sum_{i=1}^{n} (U_{ni} - U_n)^2.$$

Lemma 1. Assume that $E_F k^2(X_1,...,X_m) < \infty$. Then

$$s_U^2 \to m^2 \int \phi^2(y) dF(y) \quad a.s.,$$

where $\phi(y) = E_F[k(X_1,...,X_m) | X_1 = y] - E_F k(X_1,...,X_m)$.

This result was proved in Arvesen (1969, Theorem 5), although he stated a weaker version of this result (the weak consistency). The following lemmas are also needed for the proof of the main results.

Lemma 2. Let H, H_n and H_{ni} be defined as in Section 1. Then

(i) $\sum_{i=1}^{n} [H_n(y) - H_{ni}(y)] = 0$ for any y.

(ii) $\| H_n - H_{ni} \|_\infty \le m(n-m)^{-1}$, where $\| \ \|_\infty$ is the sup norm.

Proof. Let

$$A_{ni}(y) = [(n-1)_{(m-1)}]^{-1}\sum_{c^i_{m-1}} I[h(X_i, X_{i_1}, \ldots, X_{i_{m-1}}) \le y] \tag{2.1}$$

where $(n-1)_{(m-1)} = (n-1) \cdots (n-m+1)$ and $\sum_{c^i_{m-1}}$ is the summation taken over the $(n-1)_{(m-1)}$ $m-1$-tuples $(i_1, \ldots, i_{m-1})$ of distinct elements from the integers $\{1, \ldots, i-1, i+1, \ldots, n\}$. A straightforward calculation shows that

$$H_n(y) - H_{ni}(y) = m(n-m)^{-1}[A_{ni}(y) - H_n(y)].$$

Then (i) follows from $n^{-1}\sum_{i=1}^n A_{ni}(y) = H_n(y)$ and (ii) follows from both $\| A_{ni} \|_\infty$ and $\| H_n \|_\infty$ are bounded by one. □

Lemma 3. Assume that H is continuous. Then

$$\| H_n - H \|_\infty \to 0 \quad a.s.$$

Proof. For each y, $H_n(y)$ is a U-statistic. From theory of U-statistic, $H_n(y) \to H(y)$ $a.s.$ Let $\mathbf{D} = \{$ all rational numbers in $\mathbf{R} \}$. Then almost surely, $H_n(y) \to H(y)$ for all $y \in \mathbf{D}$. Let $\omega = (X_1, X_2, \ldots)$ be fixed such that $H_n(y) \to H(y)$ for all $y \in \mathbf{D}$. Since $\mathbf{D}$ is a dense subset of $\mathbf{R}$ and H_n is a distribution function, H_n converges weakly to H. From the continuity of H, we have $\| H_n - H \|_\infty \to 0$. This completes the proof. □

We now establish the strong consistency of s_J^2 given by (1.4) for trimmed generalized L-statistics.

Theorem 1. Assume condition A. Then

$$s_J^2 \to \sigma^2 \quad a.s.$$

Proof. Define

$$W_{ni}(y) = [H_{ni}(y) - H_n(y)]^{-1}\int_{H_n(y)}^{H_{ni}(y)} J(t)\mathrm{d}t - J[H(x)] \tag{2.2}$$

for $H_{ni}(y) \ne H_n(y)$ and $W_{ni}(y) = 0$ if $H_{ni}(y) = H_n(y)$. From Lemma 8.1.1B in Serfling (1980),

$$T(H_{ni}) - T(H_n) = \int [H_n(y) - H_{ni}(y)] J[H(y)] \mathrm{d}y \qquad (2.3)$$
$$+ \int W_{ni}(y)[H_n(y) - H_{ni}(y)] \mathrm{d}y.$$

Let $U_{ni} = \int [H_{ni}(y) - H(y)] J[H(y)] \mathrm{d}y$, $U_n = \int [H_n(y) - H(y)] J[H(y)] \mathrm{d}y$, $R_{ni} = \int W_{ni}(y)[H_n(y) - H_{ni}(y)] \mathrm{d}y$ and $\bar{R}_n = n^{-1} \sum_{i=1}^n R_{ni}$. From Lemma 2, $U_n = n^{-1} \sum_{i=1}^n U_{ni}$. Then

$$s_J^2 = (n-1) \sum_{i=1}^n (U_{ni} - U_n)^2 \qquad (2.4)$$
$$+ (n-1) \sum_{i=1}^n (R_{ni} - \bar{R}_n)^2 + 2(n-1) \sum_{i=1}^n R_{ni}(U_{ni} - U_n).$$

Note that U_n is a U-statistic with $\int \{ I[h(x_1,...,x_m) \le y] - H(y) \} J[H(y)] \mathrm{d}y$ as the kernel. Hence from Lemma 1,

$$(n-1) \sum_{i=1}^n (U_{ni} - U_n)^2 \to \sigma^2 \ \ a.s.$$

Using Cauchy-Schwarz inequality, the result follows from

$$(n-1) \sum_{i=1}^n R_{ni}^2 \to 0 \ \ a.s. \qquad (2.5)$$

Let a and b be two constants such that $H(a) < \alpha$ and $H(b) > \beta$. From Lemma 2(ii) and Lemma 3, for almost all $\omega = (X_1, X_2, ...)$, there is an $n_\omega > 0$ such that $H_{ni}(a) < \alpha$, $H_n(a) < \alpha$, $H_{ni}(b) > \beta$ and $H_n(b) > \beta$ hold for all $i \le n$ and $n \ge n_\omega$. Then $R_{ni} = \int_a^b W_{ni}(y)[H_n(y) - H_{ni}(y)] \mathrm{d}y$, since $J(t) = 0$ if $t < \alpha$ or $t > \beta$. Thus,

$$\max_{i \le n} R_{ni}^2 \le (b-a)^2 \max_{i \le n} (\| W_{ni} \|_\infty \ \| H_n - H_{ni} \|_\infty)$$
$$\le C n^{-2} \max_{i \le n} \| W_{ni} \|_\infty,$$

where C is a constant. Since J is a continuous function on $[\alpha, \beta]$, $\| H_{ni} - H_n \|_\infty \le m(n-m)^{-1}$ and $\| H_n - H \|_\infty \to 0 \ a.s.$, $\max_{i \le n} \| W_{ni} \|_\infty \to 0 \ a.s.$ Hence (2.5) holds and the result follows. □

3. Untrimmed generalized L-statistics

For untrimmed generalized L-statistics, we prove the following similar result.

Theorem 2. Assume condition B. Then

$$s_J^2 \to \sigma^2 \quad a.s.$$

Proof. From (2.2)-(2.4), we only need to show (2.5) holds. Using Lemma 2(i), we obtain

$$(n-1)\sum_{i=1}^n R_{ni}^2 = (n-1)m^2(n-m)^{-2}\sum_{i=1}^n \{\int W_{ni}(y)[A_{ni}(y) - H(y)]dy\}^2$$
$$\le Cn^{-1}\sum_{i=1}^n \{\int |A_{ni}(y) - H(y)|dy\}^2 \max_{i\le n} \|W_{ni}\|_\infty,$$

where C is a constant. From the proof of Theorem 1, $\max_{i\le n}\|W_{ni}\|_\infty \to 0$ $a.s.$ Then (2.5) follows from

$$n^{-1}\sum_{i=1}^n [\int |A_{ni}(y) - H_n(y)|dy]^2 = O(1) \quad a.s. \tag{3.1}$$

Let $\xi_n = n^{-1}\sum_{i=1}^n [\int |A_{ni}(y) - H(y)|dy]^2$. Using the notation in (2.1), we have

$$\xi_n \le n^{-1}[(n-1)_{(m-1)}]^{-1}\sum_{i=1}^n \sum_{C_{m-1}^i} \{\int |I[h(X_i, X_{i_1},...,X_{i_{m-1}})\le y] - H(y)|dy\}^2$$
$$= n_{(m)}^{-1}\sum_{C_m} \{\int |I[h(X_{i_1},...,X_{i_m})\le y] - H(y)|dy\}^2, \tag{3.2}$$

which is a U-statistic with a kernel $\{\int |I[h(x_1,...,x_m)\le y] - H(y)|dy\}^2$. Under condition (1.3),

$$E_F\{\int |I[h(X_1,...,X_m)\le y] - H(y)|dy\}^2 \tag{3.3}$$
$$= \iint E_F|I[h(X_1,...,X_m)\le y] - H(y)||I[h(X_1,...,X_m)\le z] - H(z)|dydz$$
$$\le \{\int \{E_F[I[h(X_1,...,X_m)\le y] - H(y)]^2\}^{1/2}dy\}^2$$
$$= \{\int [H(y)(1-H(y))]^{1/2}dy\}^2 < \infty.$$

From the almost sure convergence of U-statistics, the quantity in (3.2) converges almost surely to the quantity in (3.3). Hence $\xi_n = O(1)$ $a.s.$ Similarly, $\int |H_n(y) - H(y)|dy$ is bounded by

$$n_{(m)}^{-1}\sum_{C_m}\int |I[h(X_{i_1},...,X_{i_m})\le y] - H(y)|dy,$$

which converges almost surely to

$$E_F\int |I[(X_1,...,X_m)\le y] - H(y)|dy < \infty$$

under condition (1.3). Then (3.1) follows from

$$n^{-1}\sum_{i=1}^n [\int |A_{ni}(y) - H_n(y)|dy]^2 \le 2\xi_n + 2[\int |H_n(y) - H(y)|dy]^2.$$

This completes the proof. □

4. Remarks

A different type of generalized L-statistic (Serfling, 1984) is $T(K_n)$, where T is given by (1.2) and

$$K_n(y) = n^{-m} \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} I[h(X_{i_1},\ldots,X_{i_m}) \le y].$$

$T(H_n)$ and $T(K_n)$ are closely related and have the same limiting distribution. Note that $K_n(y)$ is a V-statistic. Consistency of jackknife estimators for V-statistics can be established using similar techniques in treating jackknife estimators for U-statistics (e.g., Sen, 1977). Therefore, our results in the previous sections can be extended to the statistics $T(K_n)$ with some modifications.

PART II

BOOTSTRAPPING FOR GENERALIZED L-STATISTICS

1. Introduction

Let $X_1,..., X_n$ be independent and identically distributed (i.i.d.) samples from an unknown population distribution F and $T_n = T_n(X_1,...,X_n)$ be a statistic. The bootstrap (Efron, 1979) is a useful nonparametric method for statistical analysis based on T_n. For example, the bootstrap can be used to approximate the sampling distribution of a function $L_n = L_n(T_F, T_n)$ and its other characteristics for various purposes in statistical inferences for T_F, where T_F depends on F and is an unknown parameter of interest.

Let $X_1^*,..., X_n^*$ be i.i.d. samples drawn from the empirical distribution $F_n(x) = n^{-1}\sum_{i=1}^n I[X_i \le x]$, where $I[A]$ is the indicator function of the set A. X_i^* are called bootstrap samples. A bootstrap analog for T_n is $T_n^* = T_n(X_1^*,...,X_n^*)$. The sampling distribution of L_n, $P_F\{ L_n(T_F, T_n) \le t \}$, is approximated by the bootstrap estimate $P_*\{ L_n(T_n, T_n^*) \le t \}$, where P_* is the probability corresponding to the bootstrap sampling.

In many situations L_n is $n^{1/2}(T_n - T_F)$ and it can be approximated by an average of i.i.d. random variables, i.e.,

$$T_n - T_F = n^{-1}\sum_{i=1}^n \phi(X_i) + R_n, \qquad (1.1)$$

where ϕ is a function depending on F and T_n and satisfies $E_F\phi(X_i) = 0$ and $0 < E_F\phi^2(X_i) < \infty$. Note that $n^{-1}\sum_{i=1}^n \phi(X_i) = O_p(n^{-1/2})$. Hence usually $R_n = o_p(n^{-1/2})$. More generally, we have

$$T_n - T_F = U_n + R_n, \qquad (1.2)$$

where $U_n = U_n(X_1,...,X_n)$ is a U-statistic (see Hoeffding, 1948) satisfying $E_F U_n = 0$ and $R_n = o_p(n^{-1/2})$. Serfling (1984) gives a wide class of statistics, the generalized L-statistics, which have property (1.2). More details for the generalized L-statistics is given in the next section.

A bootstrap analog of (1.2) is

$$T_n^* - T_n = U_n^* - U_n + R_n^*, \tag{1.3}$$

where $U_n^* = U_n(X_1^*,...,X_n^*)$ and R_n^* satisfies

$$R_n^* = o_p(n^{-1/2}). \tag{1.4}$$

Note that the o_p in (1.4) is with respect to the unconditional probability P defined by $P\{A\} = E_F P_*\{A\}$ for any measurable set A. Equation (1.3) can be called a bootstrap representation for the bootstrap statistic $T_n^* - T_n$. A direct consequence of (1.3)-(1.4) is that the bootstrap estimator of the sampling distribution $P_F\{\, n^{1/2}(T_n - T_F) \le t \,\}$ is weakly consistent, i.e.,

$$\sup_t \left| P_*\{\, n^{1/2}(T_n^* - T_n) \le t \,\} - P_F\{\, n^{1/2}(T_n - T_F) \le t \,\} \right| = o_p(1). \tag{1.5}$$

This follows from (1.4) and a well established bootstrap theory for U-statistics (see Bickel and Freedman, 1981).

For several classes of statistics such as (ordinary) L-statistics and differentiable statistical functionals, (1.1) holds and the bootstrap representation holds with $U_n^* = n^{-1}\sum_{i=1}^n \phi(X_i^*)$ (see Babu and Singh, 1984; Gill, 1987). The purpose of this paper is to show the bootstrap representation (1.3) holds for a wide class of statistics, the generalized L-statistics. The result includes that for ordinary L-statistics since $n^{-1}\sum_{i=1}^n \phi(X_i)$ is a special case of U-statistics.

2. Bootstrap representations

Let $h(x_1,...,x_m)$ be a symmetric function on $\mathbf{R}^m$ and $H_F(x)$ be the distribution function of $h(X_1,...,X_m)$, i.e.,

$$H_F(x) = P_F\{\, h(X_1,...,X_m) \le x \,\}, \quad x \in \mathbf{R}.$$

An empirical version of $H_F(x)$ is

$$H_n(x) = \binom{n}{m}^{-1}\sum_c I[h(X_{i_1},...,X_{i_m})\le x], \tag{2.1}$$

where $\sum_c$ is the summation taken over all combinations of m integers $(i_1,...,i_m)$ chosen from the integers $1,...,n$. Note that $H_n(x)$ is a U-statistics. Let J be a function defined on the interval [0,1], G be a distribution function and

$$T(G) = \int xJ[G(x)]\mathrm{d}G(x).$$

A class of generalized L-statistics is defined to be T_n=$T(H_n)$ (Serfling, 1984). The corresponding T_F is $T(H_F)$. Examples of generalized L-statistics include U-statistics, (ordinary) L-statistics, trimmed variances, trimmed U-statistics and Winsorized U-statistics (see more examples in Serfling, 1984).

It was shown in Serfling (1984) that T_n satisfies (1.2) with $R_n = o_p(n^{-1/2})$ and

$$U_n = \int [H_F(x)-H_n(x)]J[H_F(x)]\mathrm{d}x \tag{2.2}$$

under the following condition.

Condition A. (1) The functions J and H_F are continuous.

(2) The distribution H_F satisfies $\int [H_F(x)(1-H_F(x))]^{1/2}\mathrm{d}x < \infty$.

For any integers $1\le i_1\le i_2\le \cdots \le i_m\le n$, let $H_F^{i_1,...,i_m}$ be the distribution of $h(X_{i_1},...,X_{i_m})$. To establish the bootstrap representation (1.3)-(1.4), we need to assume

Condition B. $\int [H_F^{i_1,...,i_m}(x)(1-H_F^{i_1,...,i_m}(x))]^{1/2}\mathrm{d}x < \infty$ for any integers $i_1\le \cdots \le i_m$.

Note that for a random variable Y with distribution G, the condition $\int [G(x)(1-G(x))]^{1/2}\mathrm{d}x < \infty$ is almost equivalent to the condition $E_G Y^2 < \infty$ (see Serfling, 1980, p.276) and is implied by $E_G |Y|^{2+\delta} < \infty$ for a $\delta>0$. Hence condition B is almost the same as $E_F h^2(X_{i_1},...,X_{i_m}) < \infty$ and implied by

$E_F \, | h(X_{i_1},...,X_{i_m}) |^{2+\delta} < \infty$ for any integers $i_1 \le \cdots \le i_m$.

Lemma 1. Let H_n^* be the bootstrap analog of H_n, i.e., H_n^* is defined by (2.1) with X_{i_j} replaced by the bootstrap samples $X_{i_j}^*$. If H_F is continuous, then

$$\| H_n^* - H_F \| \to 0 \;\; a.s. \quad \text{and} \quad \| H_n^* - H_n \| \to 0 \;\; a.s.,$$

where $\| \;\; \|$ is the sup norm.

Proof. For any fixed x, since $H_n(x)$ is a bounded U-statistic and $H_n^*(x)$ is its bootstrap analog, $H_n^*(x) \to H_F(x)$ *a.s.* (Athreya et al., 1984). Then almost surely, $H_n^*(x) \to H_F(x)$ for all rational x, which implies H_n^* converges weakly to H_F *a.s.* since all rational numbers form a dense set in **R** and H_n^* is a distribution function. Then $\| H_n^* - H_F \| \to 0$ *a.s.* follows from the continuity of H_F. A similar argument shows that $\| H_n - H_F \| \to 0$ *a.s.* Hence the results hold. □

Theorem 1. Assume conditions A and B. For the generalized L-statistics $T_n = T(H_n)$, the bootstrap representation (1.3)-(1.4) holds with U_n given by (2.2).

Proof. Let $W_n^*(x) = M[H_n(x), H_n^*(x)] - J[H_F(x)]$ if $H_n(x) \ne H_n^*(x)$ and $= 0$ if $H_n(x) = H_n^*(x)$, where $M(s, t) = \int_s^t J(u) du/(t-s)$. Then from Lemma 8.1.1B of Serfling (1980),

$$T_n^* - T_n = U_n^* - U_n + \int W_n^*(x)[H_n(x) - H_n^*(x)] dx,$$

where U_n^* is the bootstrap analog of U_n given by (2.2) with H_n replaced by H_n^*. From Lemma 1 and the continuity of J, $\| W_n^* \| \to 0$ *a.s.* It remains to show that

$$\int | H_n^*(x) - H_n(x) | dx = O_p(n^{-1/2}). \tag{2.3}$$

Let E_* and V_* be the expectation and variance taken under the bootstrap probability P_*, respectively. Since $E_*[H_n^*(x)] = K_n(x)$, where

$$K_n(x) = n^{-m}\sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} I[h(X_{i_1},...,X_{i_m}) \le x], \tag{2.4}$$

we have

$$E_F E_*[\,|H_n^*(x) - H_n(x)|\,] \le \{E_F V_*[H_n^*(x)] + E_F[K_n(x) - H_n(x)]^2\}^{1/2}.$$

Hence (2.3) follows from

$$\int \{E_F[K_n(x) - H_n(x)]^2\}^{1/2} \mathrm{d}x = O(n^{-1/2}) \tag{2.5}$$

and

$$\int \{E_F V_*[H_n^*(x)]\}^{1/2} \mathrm{d}x = O(n^{-1/2}). \tag{2.6}$$

Let $Z_n(x)$ be the average of all terms $I[h(X_{i_1},...,X_{i_m}) \le x]$ with at least one equality $i_j = i_l$, $j \ne l$. From Serfling (1980, p.206),

$$H_n(x) - K_n(x) = [1 - n_{(m)}/n^m][H_n(x) - Z_n(x)], \tag{2.7}$$

where $n_{(m)} = n(n-1) \cdots (n-m+1)$. Then

$$E_F[H_n(x) - K_n(x)]^2 \le Cn^{-2}\{E_F[H_n(x) - H_F(x)]^2 + E_F[Z_n(x) - Z_F(x)]^2 + [Z_F(x) - H_F(x)]^2\},$$

where C is a constant and $Z_F(x) = E_F[Z_n(x)]$. Then (2.5) follows from condition B. Since for given $X_1,..., X_n$, $H_n^*(x)$ is a U-statistic, we have

$$V_*[H_n^*(x)] \le mn^{-1}K_n(x)[1 - K_n(x)]$$

(see Serfling, 1980, p.183). Then (2.6) follows from

$$A_n = \int \{E_F[K_n(x)(1 - K_n(x))]\}^{1/2} \mathrm{d}x = O(1).$$

Note that A_n is bounded by

$$\int_{-\infty}^{0} \{E_F[K_n(x)]\}^{1/2} \mathrm{d}x + \int_{0}^{\infty} \{E_F[1 - K_n(x)]\}^{1/2} \mathrm{d}x.$$

From (2.7),

$$E_F[K_n(x)] = [n_{(m)}/n^m]H_F(x) + [1 - n_{(m)}/n^m]Z_F(x).$$

Hence A_n is bounded by

$$\int_{-\infty}^{0} [H_F(x)]^{1/2} \mathrm{d}x + \int_{-\infty}^{0} [Z_F(x)]^{1/2} \mathrm{d}x + \int_{0}^{\infty} [1 - H_F(x)]^{1/2} \mathrm{d}x + \int_{0}^{\infty} [1 - Z_F(x)]^{1/2} \mathrm{d}x,$$

which is finite under condition B. This completes the proof. □

If the function J is more smooth (condition C), then we can obtain a stronger result than (1.4) under less requirement on the moment of $h(X_{i_1},\ldots,X_{i_m})$.

Condition C. J is Lipschitz continuous of order δ ($0<\delta\le 1$), i.e., there is a constant $C>0$ such that $|J(t)-J(s)| \le C|s-t|^{\delta}$ for any $s,t \in [0,1]$, and $\int [H_F^{i_1,\ldots,i_m}(x)(1-H_F^{i_1,\ldots,i_m}(x))]^{(1+\delta)/2}dx < \infty$ for any integers $i_1 \le \cdots \le i_m$.

Theorem 2. Assume condition C. Then (1.3) holds with

$$R_n^* = O_p(n^{-(1+\delta)/2}).$$

Proof. Using the same notation as in the proof of Theorem 1, we have $|W_n^*(x)| \le C[|H_n^*(x)-H_n(x)|^{\delta} + |H_n(x)-H_F(x)|^{\delta}]$ by the Lipschitz continuity of J. Then

$$\begin{aligned} |R_n^*| \le C[&\int |H_n^*(x)-H_n(x)|^{1+\delta}dx \\ &+ \int |H_n(x)-H_F(x)|^{\delta}|H_n^*(x)-H_n(x)|dx]. \end{aligned} \tag{2.8}$$

Since $E_F E_* |H_n^*(x)-H_n(x)|^{1+\delta} \le \{E_F E_*[H_n^*(x)-H_n(x)]^2\}^{(1+\delta)/2}$, the first integral on the right hand side of (2.8) can be shown to be $O_p(n^{-(1+\delta)/2})$ by using the same argument as in the proof of Theorem 1. Note that

$$\begin{aligned} &E_F E_* |H_n(x)-H_F(x)|^{\delta}|H_n^*(x)-H_n(x)| \\ &= E_F |H_n(x)-H_F(x)|^{\delta} E_* |H_n^*(x)-H_n(x)| \\ &\le E_F |H_n(x)-H_F(x)|^{\delta}\{E_*[H_n^*(x)-H_n(x)]^2\}^{1/2} \\ &\le [E_F |H_n(x)-H_F(x)|^{2\delta}]^{1/2}\{E_F E_*[H_n^*(x)-H_n(x)]^2\}^{1/2} \end{aligned}$$

and

$$[E_F |H_n(x)-H_F(x)|^{2\delta}]^{1/2} \le (m/n)^{\delta/2}[H_F(x)(1-H_F(x))]^{\delta/2}.$$

Using the same argument as in the proof of Theorem 1, we have

$$\int [H_F(x)(1-H_F(x))]^{\delta/2}\{E_F E_*[H_n^*(x)-H_n(x)]^2\}^{1/2}dx = O(n^{-(1+\delta)/2})$$

under condition C. Hence the result follows. □

3. Complements

(1) From (2.1), U_n defined in (2.2) is a U-statistic with a kernel

$$k(x_1,\dots,x_m) = \int\{H_F(x) - I[h(x_1,\dots,x_m) \le x]\} J[H_F(x)]\mathrm{d}x .$$

Under condition B, $E_F k^2(X_{i_1},\dots,X_{i_m}) < \infty$ for any integers $i_1 \le \cdots \le i_m$ (see Serfling, 1980, Lemma 8.2.5A). Hence (1.5) holds with $T_n = U_n$ and $T_n^* = U_n^*$ (see Bickel and Freedman, 1981). Then Theorem 1 or 2 implies that (1.5) holds for the generalized L-statistics $T_n = T(H_n)$ satisfying condition A and either condition B or condition C.

(2) Under condition A, Serfling (1984) showed that the distribution of $n^{1/2}(T_n - T_F)$ converges weakly to $N(0, \sigma^2)$, where σ^2 is given in (3.3) of Serfling (1984) and is generally unknown. In statistical analysis, we often need a consistent estimator of the asymptotic standard deviation σ. Let Q_n and q be the interquartile ranges of $P_*\{ n^{1/2}(T_n^* - T_n) \le t \}$ and $N(0,1)$, respectively. Then from (1.5), Q_n/q is consistent for σ, i.e.,

$$Q_n/q - \sigma = o_p(1).$$

(3) Serfling (1984) introduced another type of generalized L-statistics $T(K_n)$, where K_n is defined in (2.4). With some minor changes in the proofs of Theorems 1 and 2, we can establish the bootstrap representation (1.3)-(1.4) for $T(K_n)$ with U_n and U_n^* replaced by

$$V_n = \int[H_F(x) - K_n(x)]J[H_F(x)]\mathrm{d}x$$

and the bootstrap analog V_n^*, respectively. Note that V_n is a V-statistic. Since V-statistics are closely related to U-statistics, result (1.5) can be extended to $T_n = T(K_n)$ in a straightforward manner.

References

Arvesen, J. N. (1969). Jackknifing U-statistics. *Ann. Math. Statist.* **40**, 2076-2100.

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

Babu, G. J. and Singh, K. (1984). Asymptotic representations related to jackknifing and bootstrapping L-statistics. *Sankhya* A **46**, 195-206.

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

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**, 1-26.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications.* Vol. II. Wiley, New York.

Gill, R. D. (1987). Non- and semi-parametric maximum likelihood estimators and the von Mises method (Part I). To appear in *Scand. J. Statist.*

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19**, 293-325.

Huber, P. J. (1981). *Robust Statistics*. Wiley, New York.

Parr, W. C. and Schucany, W. R. (1982). Jackknifing L-statistics with smooth weight functions. *J. Amer. Statist. Assoc.* **77**, 629-638.

Quenouille, M. (1956). Notes on bias in estimation. *Biometrika* **43**, 353-360.

Sen, P. K. (1977). Some invariance principles relating to jackknifing and their role in sequential analysis. *Ann. Statist.* **5**, 316-329.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

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

Serfling, R. J. (1985). A class of problems in statistical computation: Generalized L- and related statistics. Proceedings of the sixteenth symposium on the interface, L. Billard (ed), 85-88. North-Holland Publishing Co., Amsterdam, Netherlands.

Tukey, J. (1958). Bias and confidence in not quite large samples. *Ann. Math. Statist.* **29**, 614.