Ann. Inst. Statist. Math.
Vol. 50, No. 3, 451–470 (1998)

ASYMPTOTIC COMPARISONS OF SEVERAL VARIANCE ESTIMATORS AND THEIR EFFECTS FOR STUDENTIZATIONS

YOSHIHIKO MAESONO*

Centre for Mathematics and Its Applications, School of Mathematical Sciences, The Australian National University, Canberra, ACT 0200, Australia

(Received December 21, 1995; revised June 6, 1997)

Abstract. In this paper we obtain asymptotic representations of several variance estimators of U-statistics and study their effects for studentizations via Edgeworth expansions. Jackknife, unbiased and Sen's variance estimators are investigated up to the order $o_p(n^{-1})$. Substituting these estimators to studentized U-statistics, the Edgeworth expansions with remainder term $o(n^{-1})$ are established and inverting the expansions, the effects on confidence intervals are discussed theoretically. We also show that Hinkley's corrected jackknife variance estimator is asymptotically equivalent to the unbiased variance estimator up to the order $o_p(n^{-1})$.

Key words and phrases: Edgeworth expansion, Gini's mean difference, jackknife variance estimator, studentized U-statistics, variance estimation.

1. Introduction

Let $X_1, \ldots, X_n$ be independently and identically distributed random variables with distribution function $F(x)$. Let $h(x_1, \ldots, x_r)$ be a real valued function which is symmetric in its arguments. For $n \geq r$ let us define U-statistic by

$$U_n = \binom{n}{r}^{-1} \sum_{C_{n,r}} h(X_{i_1}, \ldots, X_{i_r})$$

where $\sum_{C_{n,r}}$ indicates that the summation is taken over all integers $i_1, \ldots, i_r$ satisfying $1 \leq i_1 < \cdots < i_r \leq n$. U_n is a minimum variance unbiased estimator of $\theta = E[h(X_1, \ldots, X_r)]$ and many statistics in common use are members of U-statistics or approximated by them.

For the variance estimation of U-statistics, several estimators are proposed. In the case of degree 2, Sen (1960) has discussed an estimator of the dominant term $r^2\xi_1^2$ of the variance $n\sigma_n^2 = n\,\mathrm{Var}(U_n)$ where $\xi_1^2 = E[E\{h(X_1, \ldots, X_r) \mid X_1\}]^2$.

* Now at Faculty of Economics, Kyushu University, Fukuoka 812-8581, Japan.

Sen (1977) extended it to general degree r. The jackknife variance estimator $\hat{\sigma}_J^2$ is given by

$$\hat{\sigma}_J^2 = \frac{n-1}{n}\sum_{i=1}^{n}(U_n^{(i)} - U_n)^2$$

where $U_n^{(i)}$ denotes U-statistic computed from a sample of $n-1$ points with X_i left out. The properties of $\hat{\sigma}_J^2$ are precisely studied. Arvesen (1969) has obtained the exact representation of $\hat{\sigma}_J^2$, which is very complicated. Efron and Stein (1981) have showed that the jackknife variance estimator always has positive bias. Further Maesono (1994) has obtained an asymptotic representation of $\hat{\sigma}_J^2$ with residual term $o_p(n^{-1})$ which means

$$P\{|o_p(n^{-1})| \geq n^{-1}(\log n)^{-1}\} = o(n^{-1}).$$

He also established an Edgeworth expansion of $\hat{\sigma}_J^2$ with remainder term $o(n^{-1/2})$. The bias reduction for the jackknife variance estimator has been studied by Hinkley (1978) and Efron and Stein (1981). In the case of small sample and $r = 2$, Schucany and Bankson (1989) discussed biases and mean square errors of Sen's (1960) estimator, the jackknife estimator and an unbiased estimator. They compared the above estimators by simulation. It is easy to see that those estimators have first order consistency, which means that the estimators converge to the dominant term $r^2\xi_1^2$ of the variance.

On the other hand, the asymptotic distribution of U-statistic has been studied. For a standardized U_n, Hoeffding (1948) has proved the asymptotic normality under the conditions that $E[h^2(X_1, \ldots, X_r)]$ exists and $\xi_1^2 > 0$. Callaert *et al.* (1980), and Bickel *et al.* (1986) obtained Edgeworth expansions with remainder term $o(n^{-1})$ for the standardized U_n. The asymptotic distribution of a studentized U-statistic is also studied. The studentization is to substitute an estimator $\hat{\sigma}_n^2$ for σ_n^2. Callaert and Veraverbeke (1981) obtained the Berry-Esséen bound of the studentized U-statistic substituting a jackknife estimator $\hat{\sigma}_J^2$. For degree two ($r = 2$), Helmers (1991) obtained its Edgeworth expansion with remainder term $o(n^{-1/2})$ and Maesono (1995) obtained the expansion for an arbitrary degree r.

In this paper we study the variance estimators more precisely and obtain asymptotic representations of them with residual terms $o_p(n^{-1})$. We show that up to the order $o_p(n^{-1})$, the differences between those estimators are n^{-1} constant term and the unbiased estimator of Schucany and Bankson (1989) is asymptotically equivalent to the Hinkley's (1978) corrected jackknife estimator. Using the asymptotic representations, we also obtain Edgeworth expansions of the studentized U-statistics, in which we substitute each variance estimator for σ_n^2.

In Section 2, we review the variance estimators. In Section 3, the asymptotic representations of the estimators and their biases are studied. The Edgeworth expansion of each studentized U-statistic is established in Section 4 and we study the effects of several studentizations in the case of the variance estimation. In Section 5, the effects on confidence intervals are discussed theoretically. Proofs are given in Section 6.

2. Variance estimators

At first we will obtain H-decomposition or $ANOVA$-decomposition for U-statistics. Let us define

$$\begin{aligned}
&\theta = E[h(X_1,\ldots,X_r)], \qquad \sigma_n^2 = \mathrm{Var}(U_n),\\
&g_1(x_1) = E[h(x_1,X_2,\ldots,X_r)] - \theta\\
&g_2(x_1,x_2) = E[h(x_1,x_2,\ldots,X_r)] - \theta - g_1(x_1) - g_1(x_2),\ldots,g_r(x_1,\ldots,x_r)\\
&\qquad = h(x_1,\ldots,x_r) - \theta - \sum_{k=1}^{r-1}\sum_{C_{r,k}} g_k(x_{i_1},\ldots,x_{i_k}),\\
&A_k = \sum_{C_{n,k}} g_k(X_{i_1},\ldots,X_{i_k}),\\
&\xi_1^2 = E[g_1^2(X_1)] \qquad \text{and} \qquad \xi_2^2 = E[g_2^2(X_1,X_2)].
\end{aligned}$$

Then we have

$$U_n - \theta = \binom{n}{r}^{-1}\sum_{k=1}^{r}\binom{n-k}{r-k}A_k$$

and

$$E[g_k(X_1,\ldots,X_k)\mid X_1,\ldots,X_{k-1}] = 0 \qquad \text{a.s.}$$

Using the above equations, the variance σ_n^2 is given by

$$\begin{aligned}
\sigma_n^2 &= \sum_{k=1}^{r}\binom{r}{k}^2\binom{n}{k}^{-1}E[g_k^2(X_1,\ldots,X_k)] \\
&= \frac{r^2}{n}\xi_1^2 + \frac{r^2(r-1)^2}{2n(n-1)}\xi_2^2 + \cdots \\
&\quad + \frac{r!}{n(n-1)\cdots(n-r+1)}E[g_r^2(X_1,\ldots,X_r)].
\end{aligned} \tag{2.1}$$

From the view-point of estimation for $r^2\xi_1^2$, Sen (1960) proposed the variance estimator

$$V_S^2 = r^2(n-1)^{-1}\sum_{i=1}^{n}(S_i - U_n)^2$$

where

$$S_i = \binom{n-1}{r-1}^{-1}\sum_{C_{n-1,r-1}}^{(i)} h(X_i,X_{j_1},\ldots,X_{j_{r-1}})$$

and $\sum_{C_{n-1,r-1}}^{(i)}$ denotes the sum of all possible $r-1$ combinations from $n-1$ indices $\{1,\ldots,i-1,i+1,\ldots,n\}$. Sen (1977) also showed that

$$V_J^2 = n\hat{\sigma}_J^2 = \frac{(n-1)^2}{(n-r)^2}V_S^2. \tag{2.2}$$

Hinkley (1978) has discussed the bias correction of V_J^2. Let us define

$$Q_{i,j} = nU_n - (n-1)(U_n^{(i)} + U_n^{(j)}) + (n-2)U_n^{(i,j)}$$

where $U_n^{(i,j)}$ denotes the value of U_n when X_i and X_j are deleted out from the sample. The bias corrected jackknife variance estimator is given by

$$V_C^2 = V_J^2 - \frac{1}{n+1} \sum_{C_{n,2}} (Q_{i,j} - \bar{Q})^2$$

where $\bar{Q} = \sum_{C_{n,2}} Q_{i,j}/[n(n-1)]$. For $r = 2$, Schucany and Bankson (1989) proposed the unbiased estimator of $n\sigma_n^2$, which is constituted from unbiased estimators of each terms of the variance expression. Another expression of the variance $n\sigma_n^2$ is

$$n\sigma_n^2 = n\binom{n}{r}^{-1} \sum_{k-1}^{r} \binom{r}{k}\binom{n-r}{r-k} a_k^2 \tag{2.3}$$

where

$$a_k^2 = E[h(X_1, \ldots, X_k, X_{k+1}, \ldots, X_r)h(X_1, \ldots, X_k, X_{r+1}, \ldots, X_{2r-k})] - \theta^2.$$

Let us define

$$\begin{aligned}\zeta_k(x_1, \ldots, x_{2r-k}) \\ &= \binom{2r-k}{k}^{-1}\binom{2r-2k}{r-k}^{-1} \\ &\quad \times \sum\nolimits^{*} h(x_{i_1}, \ldots, x_{i_k}, x_{i_{k+1}}, \ldots, x_{i_r})h(x_{i_1}, \ldots, x_{i_k}, x_{i_{r+1}}, \ldots, x_{i_{2r-k}})\end{aligned}$$

where $\sum^*$ denotes the sum extending over all $\binom{2r-k}{k}\binom{2r-2k}{r-k}$ pairs. The unbiased estimator of $E[h(X_1, \ldots, X_k, X_{k+1}, \ldots, X_r)h(X_1, \ldots, X_k, X_{r+1}, \ldots, X_{2r-k})]$ is given by

$$\hat{\lambda}_k = \binom{n}{2r-k}^{-1} \sum_{C_{n,2r-k}} \zeta_k(X_{i_1}, \ldots, X_{i_{2r-k}}).$$

Similarly let us define

$$\zeta_0(x_1, \ldots, x_{2r}) = \binom{2r}{r}^{-1} \sum\nolimits^{**} h(x_{i_1}, \ldots, x_{i_r})h(x_{i_{r+1}}, \ldots, x_{i_{2r}})$$

where $\sum^{**}$ denotes the sum extending over all $\binom{2r}{r}$ pairs. Then

$$\hat{\theta}^2 = \binom{n}{2r}^{-1} \sum_{C_{n,2r}} \zeta_0(X_1, \ldots, X_{2r})$$

is an unbiased estimator of θ^2. Substituting $\hat{a}_k^2 = \hat{\lambda}_k - \hat{\theta}^2$ for a_k^2 in (2.3), we obtain an unbiased estimator V_U^2 of $n\sigma_n^2$ as

$$V_U^2 = n\binom{n}{r}^{-1}\sum_{k=1}^{r}\binom{r}{k}\binom{n-r}{r-k}\hat{a}_k^2.$$

Schucany and Bankson (1989) compared the estimators V_J^2, V_S^2 and V_U^2 by simulation in small samples $n = 10$. It is easy to see that all estimators converge to $r^2\xi_1^2$.

3. Asymptotic representations of variance estimators

To discuss asymptotic properties of a statistic, it is convenient to obtain an asymptotic representation with remainder term $o_p(n^{-1})$ which means

$$P\{|o_p(n^{-1})| \geq n^{-1}(\log n)^{-1}\} = o(n^{-1}).$$

Let $\tilde{T}$, R and $T = \tilde{T} + R$ be random variables, $H(\cdot)$ be a bounded function, and γ be a positive constant. Then

$$\text{(3.1)} \quad \sup_x |P\{T \leq x\} - H(x)| \leq \sup_x |P\{\tilde{T} \leq x\} - H(x)| + P\{|R| \geq \gamma\} + \max\{H(x-\gamma) - H(x), H(x+\gamma) - H(x)\}.$$

Thus the representation of a statistic with remainder term $o_p(n^{-1})$ is very useful for discussing asymptotic properties, especially in the case of Edgeworth expansion. For the jackknife variance estimator V_J^2, Maesono (1994) proved the following asymptotic representation.

THEOREM 1. *If $E|h(X_1, \ldots, X_r)|^{4+\varepsilon} < \infty$ for some $\varepsilon > 0$, we have*

$$\text{(3.2)} \quad V_J^2 = \tilde{V}_n^2 + r^2\xi_1^2 + \frac{r^2(r-1)^2\xi_2^2}{n} + o_p(n^{-1})$$

where

$$\tilde{V}_n^2 = \frac{2r^2}{n}\sum_{i=1}^{n} f_1(X_i) + \frac{2r^2}{n^2}\sum_{C_{n,2}} f_2(X_i, X_j),$$

$$f_1(x) = \frac{1}{2}[g_1^2(x) - \xi_1^2] + (r-1)E[g_1(X_2)g_2(x, X_2)]$$

and

$$\begin{aligned} f_2(x,y) = &-g_1(x)g_1(y) + (r-1)g_2(x,y)(g_1(x) + g_1(y)) \\ &- (r-1)E[g_2(x,X_3)g_1(X_3)] - (r-1)E[g_2(y,X_3)g_1(X_3)] \\ &+ (r-1)^2E[g_2(x,X_3)g_2(y,X_3)] \\ &+ (r-1)(r-2)E[g_3(x,y,X_3)g_1(X_3)]. \end{aligned}$$

Similarly we can obtain asymptotic representations of V_S^2, V_C^2 and V_U^2.

THEOREM 2. *If* $E|h(X_1, \ldots, X_r)|^{4+\varepsilon} < \infty$ *for some* $\varepsilon > 0$, *the asymptotic representations are*

$$V_S^2 = \tilde{V}_n^2 + r^2\xi_1^2 + \frac{r^2\{(r-1)^2\xi_2^2 - 2(r-1)\xi_1^2\}}{n} + o_p(n^{-1}), \tag{3.3}$$

$$V_C^2 = \tilde{V}_n^2 + r^2\xi_1^2 + \frac{r^2(r-1)^2\xi_2^2}{2n} + o_p(n^{-1}) \tag{3.4}$$

and

$$V_U^2 = \tilde{V}_n^2 + r^2\xi_1^2 + \frac{r^2(r-1)^2\xi_2^2}{2n} + o_p(n^{-1}). \tag{3.5}$$

The differences between the variance estimators are the constant order n^{-1} until $o_p(n^{-1})$. Especially the unbiased estimator V_U^2 is asymptotically equivalent to the Hinkley's (1978) corrected jackknife estimator V_C^2. It is easy to see that

$$E[f_1(X_1)] = E[f_2(X_1, X_2)] = 0 \quad \text{and} \quad E[f_2(X_1, X_2)|X_1] = 0 \quad \text{a.s.}$$

Using Theorems 1 and 2, we can study the asymptotic properties of the variance estimators. Maesono (1994) established the Edgeworth expansion of $(V_J^2 - n\sigma_n^2)/\sqrt{\mathrm{Var}(V_J^2)}$ with remainder term $o(n^{-1/2})$. Similar expansions are easily obtained for another estimators. Here we will study the biases for variance estimation and Gini's mean difference.

Example 1. Let us define

$$b_J = \frac{1}{2}r^2(r-1)^2\xi_2^2 \quad \text{and} \quad b_S = \frac{1}{2}r^2\{(r-1)^2\xi_2^2 - 4(r-1)\xi_1^2\}.$$

Then from (2.1), (3.2) and (3.3), b_J and b_S are n^{-1} biases of V_J^2 and V_S^2 respectively.

(i) *Variance estimation*;

Let us consider the kernel $h(x, y) = (x - y)^2/2$. Then if $\mathrm{Var}(X_1) = \sigma^2$ exists, the U-statistic

$$U_n = \binom{n}{2}^{-1} \sum_{C_{n,2}} h(X_i, X_j)$$

is an unbiased estimator of σ^2. Without loss of generality, we assume $E(X_1) = 0$. It is easy to see that

$$\theta = \sigma^2, \quad g_1(x) = \frac{1}{2}(x^2 - \sigma^2) \quad \text{and} \quad g_2(x, y) = -xy. \tag{3.6}$$

So we can get

$$\xi_1^2 = \frac{1}{4}\{E(X_1^4) - \sigma^4\} \quad \text{and} \quad \xi_2^2 = \sigma^4.$$

(*Normal distribution:*) If the underlying distribution is normal, that is $X_i \sim N(0, \sigma^2)$, we can show that

$$b_J = 2\sigma^4 \quad \text{and} \quad b_S = -2\sigma^4.$$

(*Logistic distribution:*) We consider the logistic distribution which has the density function

$$\frac{\pi e^{-\pi x/\sqrt{3}\sigma}}{\sqrt{3}\sigma(1 + e^{-\pi x/\sqrt{3}\sigma})}.$$

In this case we have that $\mathrm{Var}(X_1) = \sigma^2$,

$$b_J = 2\sigma^4 \quad \text{and} \quad b_S = -\frac{22}{5}\sigma^4.$$

(*Laplace distribution:*) Finally we consider the Laplace distribution which has the density function

$$\frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}|x|/\sigma}.$$

Then we have that $\mathrm{Var}(X_1) = \sigma^2$,

$$b_J = 2\sigma^4 \quad \text{and} \quad b_S = -8\sigma^4.$$

(ii) *Gini's mean difference*;

Let us consider the kernel $h(x, y) = |x - y|$. The corresponding U-statistic is Gini's mean difference. We study the normal case, that is $X_i \sim N(0, \sigma^2)$. From direct computations, we can show that

$$\theta = E|X_1 - X_2| = \frac{2}{\sqrt{\pi}}\sigma, \qquad \xi_1^2 = \left\{\frac{1}{3} - \frac{4 - 2\sqrt{3}}{\pi}\right\}\sigma^2 \quad \text{and}$$

$$\xi_2^2 = \left\{\frac{4}{3} + \frac{4 - 4\sqrt{3}}{\pi}\right\}\sigma^2.$$

Thus we have

$$b_J = 8\left\{\frac{1}{3} + \frac{1 - \sqrt{3}}{\pi}\right\}\sigma^2 \quad \text{and} \quad b_S = \frac{8\{5 - 3\sqrt{3}\}}{\pi}\sigma^2 (< 0).$$

The biases of V_J^2 are always positive, and in many cases the Sen's estimator V_S^2 has negative biases. Though V_C^2 and V_U^2 are unbiased until the order $o(n^{-1})$, they take sometimes negative values, especially in small sample case. V_J^2 and V_S^2 are always positive.

4. Edgeworth expansions of studentized U-statistics

Maesono (1994) obtained the Edgeworth expansion of the studentized U-statistics $\sqrt{n}(U_n - \theta)/V_J$ with remainder term $o(n^{-1})$. Let us define

$$V_n^2 = \tilde{V}_n^2 + r^2\xi_1^2 + \frac{r^2\delta}{n} + o_p(n^{-1}).$$

Since V_S^2, V_C^2 and V_U^2 can be expressed as the form V_n^2, we will obtain an Edgeworth expansion of the studentized U-statistic $\sqrt{n}(U_n - \theta)/V_n$.

LEMMA 1. *If* $E|h(X_1, \ldots, X_r)|^9 < \infty$ *and* $\xi_1^2 > 0$, *we have*

$$\begin{aligned} r\xi_1 V_n^{-1} = 1 - n^{-1}\sum_{i=1}^{n}\frac{1}{\xi_1^2}f_1(X_i) - n^{-2}\sum_{C_{n,2}}\frac{1}{\xi_1^2}\left[f_2(X_i, X_j) - \frac{3}{\xi_1^2}f_1(X_i)f_1(X_j)\right] \\ + n^{-1}\left\{\frac{3E[f_1^2(X_1)]}{2\xi_1^4} - \frac{\delta}{2\xi_1^2}\right\} + o_p(n^{-1}). \end{aligned}$$

Using the above lemma, we can obtain an asymptotic representation of the studentized U-statistic $\sqrt{n}(U_n - \theta)/V_n$. Let us define

$$\begin{aligned} \tau &= \frac{3E[f_1^2(X_1)]}{2\xi_1^4} - \frac{\delta}{2\xi_1^2}, \qquad \eta = E[f_1(X_1)g_1(X_1)] \\ q_1(x) &= \tau r g_1(x) - \frac{r}{\xi_1^2}\Big\{(f_1(x)g_1(x) - \eta) + E[f_2(x, X_2)g_1(X_2)] \\ &\qquad - \frac{3\eta}{\xi_1^2}f_1(x) + (r-1)E[g_2(x, X_2)f_1(X_2)]\Big\}, \\ q_2(x, y) &= r(r-1)g_2(x, y) - \frac{r}{\xi_1^2}[f_1(x)g_1(y) + f_1(y)g_1(x)] \end{aligned}$$

and

$$\begin{aligned} q_3(x, y, z) &= r(r-1)(r-2)g_3(x, y, z) \\ &\quad - \frac{r}{\xi_1^2}\Big\{(r-1)[f_1(x)g_2(y, z) + f_1(y)g_2(x, z) + f_1(z)g_2(x, y)] \\ &\quad + g_1(x)\left[f_2(y, z) - \frac{3}{\xi_1^2}f_1(y)f_1(z)\right] \\ &\quad + g_1(y)\left[f_2(x, z) - \frac{3}{\xi_1^2}f_1(x)f_1(z)\right] \\ &\quad + g_1(z)\left[f_2(x, y) - \frac{3}{\xi_1^2}f_1(x)f_1(y)\right]\Big\}. \end{aligned}$$

Then we can get the following lemma.

LEMMA 2. *If* $E|h(X_1, \ldots, X_r)|^9 < \infty$ *and* $\xi_1^2 > 0$, *we have*

$$\sqrt{n}V_n^{-1}(U_n - \theta) = \frac{\sqrt{n}}{r\xi_1}U_n^* - \frac{\eta}{\sqrt{n}\xi_1^3} + o_p(n^{-1})$$

where

$$U_n^* = n^{-1}\sum_{i=1}^{n}\left\{rg_1(X_i) + \frac{q_1(X_i)}{n}\right\} + n^{-2}\sum_{C_{n,2}} q_2(X_i, X_j) + n^{-3}\sum_{C_{n,3}} q_3(X_i, X_j, X_k).$$

Thus the studentized U-statistic is a sum of a U-statistic U_n^* with degree three and $n^{-1/2}$ term. For asymptotic U-statistics, Lai and Wang (1993) established the Edgeworth expansion with remainder term $o(n^{-1})$. Applying their result to U_n^*, we can obtain an Edgeworth expansion of the studentized U-statistic. Let us assume the following conditions.

(C_1) $E|h(X_1,\ldots,X_r)|^9 < \infty$

(C_2) $\limsup_{|t|\to\infty} |E[\exp\{itg_1(X_1)\}]| < 1$

(C_3) $E|g_2(X_1,X_2)|^s < \infty$ $(s > 0)$ and there exist $K+2$ Borel functions $\psi_\nu : \boldsymbol{R} \to \boldsymbol{R}$ such that $E[\psi_\nu^2(X_1)] < \infty (\nu = 1,\ldots,K+2)$, $(K+2)(s-2) > 4s + (28s - 40)I_{\{E|g_3(X_1,X_2,X_3)|>0\}}$, and the covariance matrix of $(W_1,\ldots,W_{K+2})$ is positive definite, where $W_\nu = (L\psi_\nu)(X_1)$ and $(L\psi_\nu)(y) = E[g_2(y,X_2)\psi_\nu(X_2)]$, and $I_{\{\cdot\}}$ is an indicator function.

The condition (C_3) is concerned with the number of nonzero eigen function of $g_2(x,y)$. Alternatively Lai and Wang (1993) have proved the validity of the Edgeworth expansion under the following condition $(\tilde{C}_3)$.

$(\tilde{C}_3)$ There exist constants c_ν and Borel functions $w_\nu : \boldsymbol{R} \to \boldsymbol{R}$ such that $E[w_\nu(X_1)] = 0$, $E|w_\nu(X_1)|^s < \infty$ for some $s \geq 5$ and $q_2(X_1,X_2) = \sum_{\nu=1}^{K} c_\nu w_\nu(X_1)w_\nu(X_2)$ a.s.; moreover, for some $0 < \gamma < \min\{1, 2[1 - 11/(3s)]\}$,

$$\limsup_{|t|\to\infty} \sup_{|u_1|+\cdots+|u_K|\leq|t|^{-\gamma}} \left|E\left[\exp\left(it\left\{g_1(X_1) + \sum_{\nu=1}^{K} u_\nu w_\nu(X_1)\right\}\right)\right]\right| < 1.$$

Let us define

$$\begin{aligned}
e_1 &= E[g_1^3(X_1)], \quad e_2 = (r-1)E[g_1(X_1)g_1(X_2)g_2(X_1,X_2)],\\
e_3 &= E[g_1^4(X_1)], \quad e_4 = (r-1)E[g_1^2(X_1)g_1(X_2)g_2(X_1,X_2)],\\
e_5 &= (r-1)^2E[g_1(X_2)g_1(X_3)g_2(X_1,X_2)g_2(X_1,X_3)],\\
e_6 &= (r-1)(r-2)E[g_1(X_1)g_1(X_2)g_1(X_3)g_3(X_1,X_2,X_3)],\\
\omega_1 &= \xi_1^{-3}(2e_1 + 3e_2), \quad \omega_2 = \xi_1^{-3}(e_1 + 3e_2),\\
\omega_3 &= 6\xi_1^{-4}(e_3 - 6\xi_1^4 + 12e_4 + 6e_5 + 4e_6) - 2\xi_1^{-6}(2e_1 + 3e_2)(2e_1 + 9e_2),\\
\omega_4 &= 3\xi_1^{-6}(4e_1^2 + 12e_1e_2 + 3e_2^2)\\
&\quad + 18\xi_1^{-4}(\{2\delta - (r-1)^2\xi_2^2\}\xi_1^2 - e_3 + 2\xi_1^4 - 4e_4 - 2e_5)
\end{aligned}$$

and

$$Q_n(x) = \Phi(x) + \frac{\phi(x)}{6\sqrt{n}}(\omega_1 x^2 + \omega_2) + \frac{\phi(x)}{72n}(-\omega_1^2 x^5 + \omega_3 x^3 + \omega_4 x).$$

We have the following theorem.

THEOREM 3. *Assume that the conditions* (C_1) *and* (C_2) *hold. If either condition* (C_3) *or* $(\tilde{C}_3)$ *is satisfied, we have*

$$\sup_x |P\{\sqrt{n}V_n^{-1}(U_n \quad \theta) \le x\} \quad Q_n(x)| = o(n^{-1}).$$

Since the studentizations by V_J, V_S^2, V_C^2 and V_U^2 are special cases of the above studentized U-statistic, we have the Edgeworth expansions of them. The differences between the expansions based on V_J^2, V_S^2, V_C^2 and V_U^2 appear in ω_4. Let ω_{4J}, ω_{4S}, ω_{4C} and ω_{4U} are corresponding terms to ω_4 of the Edgeworth expansions of the U-statistics studentized by V_J^2, V_S^2, V_C^2 and V_U^2, respectively. Then it follows from (3.2)~(3.5) that

$$\begin{aligned}
\omega_{4J} &= 3\xi_1^{-6}(4e_1^2 + 12e_1e_2 + 3e_2^2) \\
&\quad + 18\xi_1^{-4}((r-1)^2\xi_2^2\xi_1^2 - e_3 + 2\xi_1^4 - 4e_4 - 2e_5), \\
\omega_{4S} &= 3\xi_1^{-0}(4e_1^2 + 12e_1e_2 + 3e_2^2) \\
&\quad + 18\xi_1^{-4}(\{(r-1)^2\xi_2^2 - 4(r-1)\xi_1^2\}\xi_1^2 - e_3 + 2\xi_1^4 - 4e_4 - 2e_5)
\end{aligned}$$

and

$$\omega_{4C} = \omega_{4U} = 3\xi_1^{-6}(4e_1^2 + 12e_1e_2 + 3e_2^2) + 18\xi_1^{-4}(-e_3 + 2\xi_1^4 - 4e_4 - 2e_5).$$

Example 2. Let us consider the case of the variance estimation. The U-statistic with the kernel $h(x,y) = (x-y)^2/2$ is an unbiased estimator of $\sigma^2 = \mathrm{Var}(X_1)$. For the sake of simplicity, we will consider the case that the distribution $F(x)$ is symmetric about the origin. Let us define $m_k = E[X_1^k]$. Because of symmetry of F, if k is odd number, $m_k = 0$. It follows from (3.6) and direct computations that

$$g_1(x) = \frac{1}{2}(x^2 - \sigma^2), \qquad g_2(x,y) = -xy, \qquad f_1(x) = \frac{1}{8}(x^4 - 2\sigma^2x^2 - m_4)$$

and

$$f_2(x,y) = -\frac{1}{4}(x^2-\sigma^2)(y^2-\sigma^2) - \frac{xy}{2}(x^2+y^2-2\sigma^2) + \xi_1^2xy.$$

Thus putting

$$\begin{aligned}
c_1 &= -\frac{1}{3}, & w_1(x) &= x, & c_2 &= -\frac{1}{3\xi_1^2}, & w_2(x) &= f_1(x) + g_1(x), \\
c_3 &= \frac{1}{3\xi_1^2}, & w_3(x) &= f_1(x), & c_4 &= \frac{1}{3\xi_1^2} & \text{and} \quad w_4(x) &= g_1(x),
\end{aligned}$$

we have

$$q_2(X_1, X_2) = \sum_{\nu=1}^{4} c_\nu w_\nu(X_1) w_\nu(X_2) \qquad \text{a.s.}$$

Assume that $E|X_1|^{24} < \infty$ and the underlying distribution $F(x)$ has a density function. We can show that

$$\limsup_{|t|\to\infty} \sup_{|u_1|+\cdots+|u_4|<|t|^{-1}} \left| E\left[\exp\left(it\left\{g_1(X_1) + \sum_{\nu=1}^{4} u_\nu w_\nu(X_1)\right\}\right)\right]\right| < 1.$$

Hence the conditions (C_1), (C_2) and (C_3) are satisfied. Further we can obtain that

$$\xi_1^2 = \frac{1}{4}(m_4 - \sigma^4), \quad \xi_2^2 = \sigma^4, \quad e_1 = \frac{1}{8}(m_6 - 3\sigma^2 m_4 + 2\sigma^6),$$
$$e_3 = \frac{1}{16}(m_8 - 4\sigma^2 m_6 + 6\sigma^4 m_4 - 3\sigma^8) \quad \text{and} \quad e_2 = e_4 = e_5 = e_6 = 0.$$

(*Normal distribution:*) If the underlying distribution is normal, that is $X_i \sim N(0, \sigma^2)$, we can show that

$$e_1 = \sigma^6, \quad e_3 = \frac{15}{4}\sigma^8, \quad \xi_1^2 = \frac{1}{2}\sigma^4,$$
$$\omega_1 = 4\sqrt{2}, \quad \omega_2 = 2\sqrt{2}, \quad \omega_3 = -10,$$
$$\omega_{4J} = -102, \quad \omega_{4S} = -174 \quad \text{and} \quad \omega_{4C} = \omega_{4U} = -138.$$

(*Logistic distribution:*) We consider the logistic distribution which has the density function $\pi e^{-\pi x/(\sqrt{3}\sigma)}/\sqrt{3}\sigma(1 + e^{-\pi x/\sqrt{3}\sigma})$. In this case we have

$$e_1 = \frac{128}{35}\sigma^6, \quad e_3 = \frac{240}{7}\sigma^8, \quad \xi_1^2 = \frac{4}{5}\sigma^4,$$
$$\omega_1 = 10.22, \quad \omega_2 = 5.11, \quad \omega_3 = 76.45,$$
$$\omega_{4J} = -592.32, \quad \omega_{4S} = -664.32 \quad \text{and} \quad \omega_{4C} = \omega_{4U} = -614.82.$$

(*Laplace distribution:*) For the Laplace distribution which has the density function $e^{-\sqrt{2}|x|/\sigma}/\sqrt{2}\sigma$, we get

$$e_1 = \frac{37}{4}\sigma^6, \quad e_3 = \frac{2193}{16}\sigma^8, \quad \xi_1^2 = \frac{5}{4}\sigma^4,$$
$$\omega_1 = 13.24, \quad \omega_2 = 6.62, \quad \omega_3 = 136.98,$$
$$\omega_{4J} = -1002.86, \quad \omega_{4S} = -1063.34 \quad \text{and} \quad \omega_{4C} = \omega_{4U} = -1017.26.$$

In all cases, we have $\omega_{4S} < \omega_{4C} < \omega_{4J}$. If the sample size n is small, the Edgeworth expansions are affected by the differences of the studentizations.

5. Confidence intervals

One important application of the Edgeworth expansion is to construct a confidence interval of the parameter θ. Here we will discuss the effects of the studentizations to the confidence interval. From a Cornish-Fisher expansion based on the Edgeworth expansion, an approximation s_α of the α-quantile of $\sqrt{n}(U_n - \theta)/V_n$ is

$$s_\alpha = z_\alpha - \frac{1}{6\sqrt{n}}(\omega_1 z_\alpha^2 + \omega_2) + \frac{1}{72n}\{(4\omega_1^2 - 2\omega_1\omega_2 - \omega_3)z_\alpha^3 + (4\omega_1\omega_2 - \omega_2^2 - \omega_4)z_\alpha\} \tag{5.1}$$

where z_α is an α-quantile of the standard normal distribution and ω_i's are defined in Section 4. The confidence interval $I_{(ST)}$ with coefficient $1-\alpha$ $(0<\alpha<1/2)$ is given by

$$u_n - \frac{v_n}{\sqrt{n}}\hat{s}_{1-\alpha/2} \le \theta \le u_n - \frac{v_n}{\sqrt{n}}\hat{s}_{\alpha/2} \tag{5.2}$$

where u_n and v_n are sample values of U_n and V_n respectively, and $\hat{s}_\alpha$ is a sample value of s_α with estimated values $\hat{\omega}_i (i = 1 \sim 4)$.

Also we can construct a confidence interval based on the Edgeworth expansion of the standardized U-statistic $(U_n - \theta)/\sigma_n$. The confidence interval $I^*_{(SD)}$ with coefficient $1-\alpha$ is given by

$$u_n - \frac{v_n}{\sqrt{n}}\hat{s}^*_{1-\alpha/2} \le \theta \le u_n - \frac{v_n}{\sqrt{n}}\hat{s}^*_{\alpha/2} \tag{5.3}$$

where

$$\hat{s}^*_\alpha = z_\alpha + \frac{\hat{\kappa}_3}{6\sqrt{n}}(z_\alpha^2 - 1) + \frac{\hat{\kappa}_4}{24n}(z_\alpha^3 - 3z_\alpha) - \frac{\hat{\kappa}_3^2}{36n}(2z_\alpha^3 - 5z_\alpha), \tag{5.4}$$

and $\hat{\kappa}_3$ and $\hat{\kappa}_4$ are estimated values of

$$\kappa_3 = \xi_1^{-3}(e_1 + 3e_2) \quad \text{and} \quad \kappa_4 = \xi_1^{-4}(e_3 - 3\xi_1^3 + 12e_4 + 12e_5 + 4e_6).$$

As pointed out by Hall ((1992), Chap. 3), the convergence rate of a coverage probability of the interval $I_{(ST)}$ is $o(n^{-1/2})$, and that of the interval $I^*_{(SD)}$ is $O(n^{-1/2})$. Thus from the theoretical view-point, the confidence interval $I_{(ST)}$ is better than $I^*_{(SD)}$. Here we compare the intervals based on the variance estimators V_J^2, V_S^2, V_C^2 and V_U^2. Comparisons are theoretical and not by simulation. Similarly as Lemma 1, we can show that

$$\begin{aligned} V_n &= r\xi_1 + \frac{1}{2r\xi_1}\left(\tilde{V}_n^2 + \frac{r^2\delta}{n}\right) - \frac{1}{8r^3\xi_1^3}\left(\tilde{V}_n^2 + \frac{r^2\delta}{n}\right)^2 + o_p(n^{-1}) \\ &= r\xi_1 + \frac{r\delta}{2n\xi_1} - \frac{r}{2n\xi_1^3}E[f_1^2(X_1)] \\ &\quad + \frac{1}{2r\xi_1}\tilde{V}_n^2 - \frac{r}{n^2\xi_1^3}\sum_{C_{n,2}} f_1(X_i)f_1(X_j) + o_p(n^{-1}). \end{aligned} \tag{5.5}$$

Since $E[\tilde{V}_n^2] = E[f_1(X_1)f_1(X_2)] = 0$, V_n is asymptotically equal to

$$r\xi_1 + \frac{r\delta}{2n\xi_1} - \frac{r}{2n\xi_1^3}E[f_1^2(X_1)].$$

Also it may be possible to make estimators $\hat{\omega}_i (i = 1 \sim 4)$, $\hat{\kappa}_3$ and $\hat{\kappa}_4$ which converge to $\omega_i (i = 1 \sim 4)$, κ_3 and κ_4 respectively. Replacing $\hat{\omega}_i (i = 1 \sim 4)$, $\hat{\kappa}_3$ and $\hat{\kappa}_4$ by $\omega_i (i = 1 \sim 4)$, κ_3 and κ_4, we can compare the confidence intervals theoretically.

Example 3. Continued from Example 2.

Let us consider the case of the variance estimation. Since exact lower and upper bounds are available, we discuss the case of the normal distribution. Let us assume that $X_i \sim N(0, \sigma^2)$. The exact confidence interval with coefficient $1-\alpha$ is given by

$$\frac{(n-1)u_n}{\chi^2_{n-1;1-\alpha/2}} \le \sigma^2 \le \frac{(n-1)u_n}{\chi^2_{n-1;\alpha/2}}$$

where $\chi^2_{n-1;\alpha}$ is an α-quantile of the chi-square distribution with $n-1$ degree of freedom. From (5.5) and direct computations, we have

$$v_n \simeq \sqrt{2}\sigma^2 + \frac{\sqrt{2}\delta}{n\sigma^2} - \frac{7\sqrt{2}}{4n}\sigma^2.$$

Since u_n is an estimate of σ^2, we can approximate the estimates v_J, v_S, v_C and v_U

$$v_J \simeq \sqrt{2}\sigma^2 - \frac{3\sqrt{2}}{4n}\sigma^2 \simeq \left(\sqrt{2} - \frac{3\sqrt{2}}{4n}\right)u_n,$$

$$v_S \simeq \sqrt{2}\sigma^2 - \frac{7\sqrt{2}}{4n}\sigma^2 \simeq \left(\sqrt{2} - \frac{7\sqrt{2}}{4n}\right)u_n,$$

$$v_C \simeq \sqrt{2}\sigma^2 - \frac{5\sqrt{2}}{4n}\sigma^2 \simeq \left(\sqrt{2} - \frac{5\sqrt{2}}{4n}\right)u_n$$

and $v_U \simeq v_C$. It follows from Example 2 and (5.1) that the approximations of the α-quantile $\hat{s}_\alpha$ are given by

$$\hat{s}_{\alpha,J} \simeq z_\alpha - \frac{1}{6\sqrt{n}}(4\sqrt{2}z_\alpha^2 + 2\sqrt{2}) + \frac{1}{72n}(106z_\alpha^3 + 158z_\alpha),$$

$$\hat{s}_{\alpha,S} \simeq z_\alpha - \frac{1}{6\sqrt{n}}(4\sqrt{2}z_\alpha^2 + 2\sqrt{2}) + \frac{1}{72n}(106z_\alpha^3 + 230z_\alpha),$$

$$\hat{s}_{\alpha,C} \simeq z_\alpha - \frac{1}{6\sqrt{n}}(4\sqrt{2}z_\alpha^2 + 2\sqrt{2}) + \frac{1}{72n}(106z_\alpha^3 + 194z_\alpha)$$

and $\hat{s}_{\alpha,U} \simeq \hat{s}_{\alpha,C}$. Here $\hat{s}_{\alpha,\cdot}$ denotes the Cornish-Fisher approximation based on each studentization. So the differences of the confidence intervals are the coefficients of u_n. For example, the approximation of $u_n - v_J\hat{s}_{\alpha,J}/\sqrt{n}$ is

$$\left\{1 - \frac{1}{\sqrt{n}}\left[z_\alpha - \frac{1}{6\sqrt{n}}(4\sqrt{2}z_\alpha^2 + 2\sqrt{2}) + \frac{1}{72n}(106z_\alpha^3 + 158z_\alpha)\right]\left[\sqrt{2} - \frac{3\sqrt{2}}{4n}\right]\right\}u_n.$$

Similarly an approximation of $\hat{s}^*_\alpha$ is given by

$$\hat{s}^*_\alpha \simeq z_\alpha + \frac{\sqrt{2}}{3\sqrt{n}}(z_\alpha^2 - 1) + \frac{1}{18n}(z_\alpha^3 - 7z_\alpha).$$

Table 1. Coefficients of u_n.

n	α	V_n^2	exact lower	$I_{(ST)}$ lower	$I^*_{(SD)}$ lower	exact upper	$I_{(ST)}$ upper	$I^*_{(SD)}$ upper
10	0.05	V_J^2	0.4731	0.0882	0.0282	3.3329	2.9827	1.6213
		V_S^2		0.1145	0.1333		2.8407	1.5542
		V_C^2		0.0991	0.0808		2.9139	1.5877
	0.01	V_J^2	0.3815	−0.4602	−0.4109	5.1875	4.2202	1.7159
		V_S^2		−0.3974	−0.2583		3.9671	1.6385
		V_C^2		−0.4317	−0.3346		4.0965	1.6772
20	0.05	V_J^2	0.5783	0.4479	0.3175	2.1333	2.1093	1.5002
		V_S^2		0.4483	0.3530		2.0799	1.4742
		V_C^2		0.4477	0.3352		2.0950	1.4872
	0.01	V_J^2	0.4925	0.2049	0.0360	2.7762	2.7108	1.6024
		V_S^2		0.2090	0.0861		2.6591	1.5711
		V_C^2		0.2064	0.0610		2.6854	1.5868
101	0.05	V_J^2	0.7718	0.7619	0.7081	1.3473	1.3519	1.2547
		V_S^2		0.7615	0.7110		1.3511	1.2521
		V_C^2		0.7617	0.7096		1.3515	1.2534
	0.01	V_J^2	0.7134	0.6911	0.6034	1.4853	1.4959	1.3228
		V_S^2		0.6907	0.6073		1.4945	1.3196
		V_C^2		0.6909	0.6054		1.4952	1.3212

Multiplying the approximations of v_J, v_S, v_C and v_U, we can obtain approximations of lower and upper bounds of the intervals based on the standardized U-statistic.

The Table 1 lists the coefficients of u_n for lower and upper bounds of $I_{(ST)}$ and $I^*_{(SD)}$. The approximations based on the studentizations are better than the standardizations except the lower bounds of the case $n = 10$. This supports the fact that the convergence rate of the coverage probability of the interval based on the studentized U-statistic is better than that of the interval based on the standardized U-statistic. The intervals based on the jackknife studentization are comparable to those based on the Sen's studentization. The studentizations based on the Hinkley's and the unbiased estimators are moderate intervals.

6. Proofs

We first review the moments evaluations of H-decomposition, which is very useful for discussing asymptotic properties. Let $\nu(x_1, \dots, x_r)$ be a function which is symmetric in its arguments and $E[\nu(X_1, \dots, X_r)] = 0$. Let us define

$$\rho_1(x_1) = E[\nu(x_1, X_2, \dots, X_r)],$$
$$\rho_2(x_1, x_2) = E[\nu(x_1, x_2, \dots, X_r)] - \rho_1(x_1) - \rho_1(x_2), \dots,$$

and

$$\rho_r(x_1, x_2, \ldots, x_r) = \nu(x_1, x_2, \ldots, x_r) - \sum_{k=1}^{r-1} \sum_{C_{r,k}} \rho_k(x_{i_1}, x_{i_2}, \ldots, x_{i_k}).$$

Then we can show that

$$E[\rho_k(X_1, \ldots, X_k) \mid X_1, \ldots, X_{k-1}] = 0 \quad \text{a.s.} \tag{6.1}$$

and

$$\sum_{C_{n,r}} \nu(X_{i_1}, \ldots, X_{i_r}) = \sum_{k=1}^{r} \binom{n-k}{r-k} \Lambda_k$$

where

$$\Lambda_k = \sum_{C_{n,k}} \rho_k(X_{i_1}, \ldots, X_{i_k}).$$

Using the equation (6.1) and moment evaluations of martingales (Dharmadhikari *et al.* (1968)), the upper bounds of the absolute moments of Λ_k are given by

$$E|\Lambda_k|^q \leq cn^{(qk)/2} \qquad (q \geq 2, 1 \leq k \leq r). \tag{6.2}$$

It follows from Markov's inequality that if

$$E|R|^\beta = O(n^{-1-\beta-\gamma}) \qquad \text{for some} \quad \beta \geq 1 \ \text{ and } \ \gamma > 0, \tag{6.3}$$

we have

$$P\{|R| \geq n^{-1}(\log n)^{-1}\} = o(n^{-1}). \tag{6.4}$$

It is trivial that $cn^{-1-\gamma} = o_p(n^{-1})$ for a constant c and $\gamma > 0$.

Using the above evaluations, we can easily prove the following lemma.

LEMMA 3. *If* $E[\nu(X_1, \ldots, X_r)] = 0$ *and* $E|\nu(X_1, \ldots, X_r)|^{4+\varepsilon} < \infty$ *for* $\varepsilon > 0$, *we have that*

$$n^{-r-1} \sum_{C_{n,r}} \nu(X_1, \ldots, X_r) = o_p(n^{-1}), \tag{6.5}$$

$$n^{-r} \sum_{k=3}^{r} \Lambda_k = o_p(n^{-1}) \tag{6.6}$$

$$n^{-2}\Lambda_1^2 = n^{-1}E[\rho_1^2(X_1)] + n^{-2} \sum_{C_{n,2}} 2\rho_1(X_i)\rho_1(X_j) + o_p(n^{-1}), \tag{6.7}$$

$$n^{-3}\Lambda_1\Lambda_2 = o_p(n^{-1}) \tag{6.8}$$

and

$$n^{-4}\Lambda_2^2 = o_p(n^{-1}). \tag{6.9}$$

PROOF OF THEOREM 2.

[*Approximation of* V_S^2]

It follows from the equations (2.2), (3.2) and (6.5) that

$$V_S^2 = \left\{1 - \frac{2(r-1)}{n} + O(n^{-2})\right\} V_J^2 = V_J^2 - \frac{2r^2(r-1)}{n}\xi_1^2 + o_p(n^{-1}).$$

Thus we can easily obtain the equation (3.3).

[*Approximation of* V_C^2]

To obtain the equation (3.4), it is sufficient to prove the following lemma.

LEMMA 4. *If* $E|h(X_1,\ldots,X_r)|^{4+\varepsilon} < \infty$ *for some* $\varepsilon > 0$, *we have*

$$\frac{1}{n+1}\sum_{1\le i<j\le n}(Q_{i,j} - \bar{Q})^2 = \frac{r^2(r-1)^2}{2n}\xi_2^2 + o_p(n^{-1}). \tag{6.10}$$

PROOF. For the sake of simplicity we will consider the case $r = 2$. For general r, we can prove Lemma 4 similarly. Since $\sum_{i=1}^n U_n^{(i)} = nU_n$ and $\sum_{j=1,\neq i}^n U_n^{(i,j)} = (n-1)U_n^{(i)}$, we have $\bar{Q} = 0$. Further since $\sum_{i=1,\neq j}^n U_n^{(i)} = nU_n - U_n^{(j)}$,

$$\begin{aligned}
&\sum_{C_{n,2}}(Q_{i,j} - \bar{Q})^2 \\
&\quad = \frac{1}{2}\sum_{i\neq j}\{(n-2)^2(U_n^{(i,j)} - U_n^{(i)})^2 + (n-1)^2(U_n^{(j)} - U_n)^2 + (U_n^{(i)} - U_n)^2 \\
&\qquad\qquad + 2(n-1)(U_n^{(j)} - U_n)(U_n^{(i)} - U_n) \\
&\qquad\qquad - 2(n-2)(U_n^{(i,j)} - U_n^{(i)})(U_n^{(i)} - U_n) \\
&\qquad\qquad - 2(n-1)(n-2)(U_n^{(i,j)} - U_n^{(i)})(U_n^{(j)} - U_n)\} \\
&\quad = \frac{(n-2)^2}{2}\sum_{i\neq j}(U_n^{(i,j)} - U_n^{(i)})^2 - \frac{n(n-1)(n-2)}{2}\sum_{i=1}^n(U_n^{(i)} - U_n)^2.
\end{aligned}$$

From direct computations, we can show that

$$\begin{aligned}
\sum_{i=1}^n(U_n^{(i)} - U_n)^2 &= \frac{4}{n(n-1)}D_1 - \frac{8}{n(n-1)^2}D_2 + \frac{8}{n(n-1)^2}D_3 \\
&\quad \frac{16}{n(n-1)^2(n-2)}D_4 + \frac{8}{n(n-1)^2(n-2)}D_5 \\
&\quad + \frac{8(n-4)}{n(n-1)^2(n-2)^2}D_6 - \frac{16}{n(n-1)^2(n-2)^2}D_7
\end{aligned}$$

where

$$D_1 = \sum_{i=1}^{n} g_1^2(X_i), \qquad D_2 = \sum_{C_{n,2}} g_1(X_i)g_1(X_j),$$

$$D_3 = \sum_{C_{n,2}} \{g_1(X_i) + g_1(X_j)\}g_2(X_i, X_j),$$

$$D_4 = \sum_{C_{n,3}} \{g_1(X_i)g_2(X_j, X_k) + g_1(X_j)g_2(X_i, X_k) + g_1(X_k)g_2(X_i, X_j)\},$$

$$D_5 = \sum_{C_{n,2}} g_2^2(X_i, X_j),$$

$$D_6 = \sum_{C_{n,3}} \{g_2(X_i, X_j)g_2(X_i, X_k) + g_2(X_i, X_j)g_2(X_j, X_k) + g_2(X_i, X_k)g_2(X_j, X_k)\}$$

and

$$D_7 = \sum_{C_{n,4}} \{g_2(X_i, X_j)g_2(X_k, X_\ell) + g_2(X_i, X_k)g_2(X_j, X_\ell) + g_2(X_i, X_\ell)g_2(X_j, X_k)\}.$$

Note that $\sum_{j=1,\neq i}^{n}(U_n^{(i,j)} - U_n^{(i)})^2$ is a value of $\sum_{i=1}^{n}(U_n^{(i)} - U_n)^2$ when X_i is deleted from the sample. Let $\nu(x_1, \ldots, x_s)$ be a function which is symmetric in its arguments. Then we have

$$\sum_{i=1}^{n} \sum\nolimits_{C_{n-1,s}}^{(i)} \nu(X_{j_1}, \ldots, X_{j_s}) = (n-s) \sum_{C_{n,s}} \nu(X_{i_1}, \ldots, X_{i_s})$$

where $\sum_{C_{n-1,s}}^{(i)}$ denotes a sum of all s combinations from $n-1$ indices $\{1, \ldots, i-1, i+1, \ldots, n\}$. Using this equation, we have

$$\begin{aligned}
&\sum_{i=1}^{n} \sum_{j=1,\neq i}^{n} (U_n^{(i,j)} - U_n^{(i)})^2 \\
&\quad = \frac{4}{n-2}D_1 - \frac{8}{(n-1)(n-2)}D_2, \\
&\quad + \frac{8}{(n-1)(n-2)}D_3 - \frac{16}{(n-1)(n-2)^2}D_4 + \frac{8}{(n-1)(n-2)(n-3)}D_5 \\
&\quad + \frac{8(n-5)}{(n-1)(n-2)^2(n-3)}D_6 - \frac{16(n-4)}{(n-1)(n-2)^2(n-3)^2}D_7.
\end{aligned}$$

Comparing the coefficients, it follows from H-decomposition and (6.5) that

$$\frac{1}{n+1} \sum_{C_{n,2}} (Q_{i,j} - \bar{Q})^2 = \frac{2}{n}\xi_2^2 + o_p(n^{-1}).$$

[*Approximation of* V_U^2]

Finally we will consider the unbiased estimator V_U^2. Applying H-decomposition to $\hat{a}_k^2$, it follows from (6.5) that

$$n\binom{n}{r}^{-1}\sum_{k=3}^{r}\binom{r}{k}\binom{n-r}{r-k}\hat{a}_k^2 = o_p(n^{-1}).$$

We will obtain approximations of $\hat{a}_1^2$ and $\hat{a}_2^2$. From the definitions,

$$\begin{aligned}
&E[h(x,y,X_3,\ldots,X_r)h(x,X_{r+1},\ldots,X_{2r-1})]\\
&\quad = \{g_2(x,y)+g_1(x)+g_1(y)+\theta\}\{g_1(x)+\theta\},\\
&E[h(x,y,X_3,\ldots,X_r)h(X_3,X_{r+1},\ldots,X_{2r-1})]\\
&\quad = E[\{g_3(x,y,X_3)+g_2(x,X_3)+g_2(y,X_3)\}g_1(X_3)]+\xi_1^2\\
&\qquad + \theta\{g_2(x,y)+g_1(x)+g_1(y)+\theta\}
\end{aligned}$$

and

$$\begin{aligned}
&E[h(x,X_3,\ldots,X_{r+1})h(y,X_3,X_{r+2},\ldots,X_{2r-1})]\\
&\quad = E[g_2(x,X_3)g_2(y,X_3)+\{g_2(x,X_3)+g_2(y,X_3)\}g_1(X_3)]\\
&\qquad + g_1(x)g_1(y)+\theta\{g_1(x)+g_1(y)+\theta\}+\xi_1^2.
\end{aligned}$$

It follows from the above equations that

$$\begin{aligned}
&(2r-1)E[\zeta_1(x,y,X_3,\ldots,X_{2r-1})]\\
&\quad = (r-2)E[g_3(x,y,X_3)g_1(X_3)]+(r-1)E[g_2(x,X_3)g_2(y,X_3)]\\
&\qquad + (2r-3)\{E[\{g_2(x,X_3)+g_2(y,X_3)\}g_1(X_3)]+\xi_1^2\}\\
&\qquad + (r-1)g_1(x)g_1(y)+r\theta g_2(x,y)+(2r-1)\theta\{g_1(x)+g_1(y)+\theta\}\\
&\qquad + \{g_2(x,y)+g_1(x)+g_1(y)+\theta\}\{g_1(x)+g_1(y)\}.
\end{aligned}$$

We also have

$$\begin{aligned}
&(2r-1)E[\zeta_1(x,X_2,X_3,\ldots,X_{2r-k})]\\
&\quad = (2r-2)\{E[g_2(x,X_2)g_1(X_2)]+\xi_1^2\}+2r\theta g_1(x)+(2r-1)\theta^2+g_1^2(x).
\end{aligned}$$

Thus using H-decomposition and (6.6), we can show that

$$\hat{\lambda}_1 = \xi_1^2+\theta^2+\frac{1}{n}\sum_{i-1}^{n}\tilde{g}_1(X_i)+\frac{2r-2}{n^2}\sum_{C_{n,2}}\tilde{g}_2(X_i,X_j)+o_p(n^{-1}) \tag{6.11}$$

where

$$\tilde{g}_1(x) = (2r-2)E[g_2(x,X_2)g_1(X_2)]+2r\theta g_1(x)+\{g_1^2(x)-\xi_1^2\}$$

and

$$\begin{aligned}
\tilde{g}_2(x,y) &= (r-2)E[g_3(x,y,X_3)g_1(X_3)]+(r-1)E[g_2(x,X_3)g_2(y,X_3)]\\
&\quad - E[\{g_2(x,X_3)+g_2(y,X_3)\}g_1(X_3)]\\
&\quad + (r+1)g_1(x)g_1(y)+\{g_1(x)+g_1(y)+r\theta\}g_2(x,y).
\end{aligned}$$

Next we will obtain an approximation of $\hat{\theta}^2$. Similarly as $\hat{\lambda}_1$, we can get

$$\hat{\theta}^2 - \theta^2 + \frac{2r}{n}\sum_{i=1}^{n} g_1^*(X_i) + \frac{2r}{n^2}\sum_{C_{n,2}} g_2^*(X_i, X_j) + o_p(n^{-1}) \tag{6.12}$$

where

$$g_1^*(x) = \theta g_1(x) \quad \text{and} \quad g_2^*(x,y) = r g_1(x) g_1(y) + (r-1)\theta g_2(x,y).$$

Combining (6.11) and (6.12), the approximation of $\hat{a}_1^2$ is given by

$$n\binom{n}{r}^{-1}\binom{r}{1}\binom{n-r}{r-1}\hat{a}_1^2 = r^2\xi_1^2 - \frac{r^2(r-1)^2}{n}\xi_1^2 + \tilde{V}_n^2 + o_p(n^{-1}).$$

Similarly we can show that

$$n\binom{n}{r}^{-1}\binom{r}{2}\binom{n-r}{r-2}\hat{a}_2^2 = \frac{r^2(r-1)^2}{2n}\{2\xi_1^2 + \xi_2^2\} + o_p(n^{-1}).$$

Combining the above evaluations, we have the desired approximation (3.5).

PROOF OF LEMMA 1. Using Taylor expansion, we have

$$V_n^{-1} = \frac{1}{r\xi_1} - \frac{1}{2r^3\xi_1^3}\left(\tilde{V}_n^2 + \frac{r^2\delta}{n}\right) + \frac{3}{8r^5\xi_1^5}\left(\tilde{V}_n^2 + \frac{r^2\delta}{n}\right)^2 + o_p(n^{-1}).$$

From (6.7), (6.8) and (6.9), we have an approximation of $\tilde{V}_n^2$ and so Lemma 1 (cf. Maesono (1994), Lemma 3).

PROOF OF LEMMA 2. Combining Lemma 1 and Lemma 3, we can prove the asymptotic representation of the studentized U-statistic (cf. Maesono (1994), Theorem 3).

PROOF OF THEOREM 3. It follows from Lemma 2 that

$$P\{\sqrt{n}V_n^{-1}(U_n - \theta) \le x\} = P\left\{\frac{\sqrt{n}}{r\xi_1}U_n^* - \frac{\eta}{\sqrt{n}\xi_1^3} \le x\right\} + o(n^{-1}).$$

Applying the results of Lai and Wang (1993) to $\sqrt{n}U_n^*/(r\xi_1)$, we have

$$P\left\{\frac{\sqrt{n}}{r\xi_1}U_n^* - \frac{\eta}{\sqrt{n}\xi_1^3} \le x\right\} = P\left\{\frac{\sqrt{n}}{r\xi_1}U_n^* \le x + \frac{\eta}{\sqrt{n}\xi_1^3}\right\} = Q_n(x) + o(n^{-1}).$$

Acknowledgements

The author wishes to thank the referees for their helpful comments. He is also grateful to the hospitality of the Centre for Mathematics and its Applications at the Australian National University, where he carried out a part of this study.

References

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

Bickel, P. J., Goetze, F. and van Zwet W. R. (1986). The Edgeworth expansion for U-statistics of degree two, *Ann. Statist.*, **14**, 1463–1484.

Callaert, H. and Veraverbeke, N. (1981). The order of the normal approximation for a studentized U-statistic, *Ann. Statist.*, **9**, 194–200.

Callaert, H., Janssen, P. and Veraverbeke, N. (1980). An Edgeworth expansion for U-statistics, *Ann. Statist.*, **8**, 299–312.

Dharmadhikari, S. W., Fabian, V. and Jogdeo, K. (1968). Bounds on the moments of martingales, *Ann. Math. Statist.*, **39**, 1719–1723.

Efron, B. and Stein, C. (1981). The jackknife estimate of variance, *Ann. Statist.*, **9**, 586–596.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*, Springer, New York.

Helmers, R. (1991). On the Edgeworth expansion and the bootstrap approximation for a studentized U-statistic, *Ann. Statist.*, **19**, 470–484.

Hinkley, D. V. (1978). Improving the jackknife with special reference to correlation estimation, *Biometrika*, **65**, 13–21.

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

Lai, T. L. and Wang, J. Q. (1993). Edgeworth expansion for symmetric statistics with applications to bootstrap methods, *Statistica Sinica*, **3**, 517–542.

Maesono, Y. (1994). Edgeworth expansions of a studentized U-statistic and a jackknife estimator of variance, *The Australian National University Statistics Research Report*, No. SRR 038-94.

Maesono, Y. (1995). On the normal approximations of studentized U-statistic, *J. Japan Statist. Soc.*, **25**, 19–33.

Schucany, W. R. and Bankson, D. M. (1989). Small sample variance estimators for U-statistics, *Austral. J. Statist.*, **31**, 417–426.

Sen, P. K. (1960). On some convergence properties of U-statistics, *Calcutta Statist. Assoc. Bull.*, **10**, 1–18.

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