J. Japan Statist. Soc.
Vol. 28 No. 1 1998 1–19

ASYMPTOTIC MEAN SQUARE ERRORS OF VARIANCE ESTIMATORS FOR U-STATISTICS AND THEIR EDGEWORTH EXPANSIONS

Yoshihiko Maesono*

This paper studies variance estimators for a class of U-statistics. We obtain asymptotic representations of jackknife, Hinkley's (1978) corrected jackknife, unbiased, Sen's (1960) and new variance estimators. And we investigate asymptotic mean square errors of them, theoretically. The Edgeworth expansions of the estimators with remainder term $o(n^{-1})$ are also established. We show that the normalized Hinkley's corrected estimator coincides the normalized unbiased estimator until the order $n^{-1/2}o_p(n^{-1})$.

Key words and phrases: Edgeworth expansions, estimation of variance, jackknife estimator, mean square errors, U-statistics.

1. Introduction

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

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

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

Several variance estimators for the U-statistic are proposed. Sen (1960) has discussed an estimator of the dominant term $r^2 E[E\{h(X_1, \cdots, X_r) | X_1\} - \theta]^2$ of the variance $n\sigma_n^2 = nVar(U_n)$ in the case of degree 2 and Sen (1977) extended it to general degree r. He also proved the law of large numbers. 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 complicated, and Efron and Stein (1981) have showed that the jackknife variance estimator has positive bias. The bias reduction for the jackknife variance estimator has been studied by Hinkley (1978), and Efron and Stein (1981).

Received June 1996. Revised February 4, 1998. Accepted November, 1997.
*Faculty of Economics, Kyushu University 27 Hakozaki 6-19-1, Higashi-ku, Fukuoka 812-81, Japan.

In the case of small sample, using computer simulation, Schucany and Bankson (1989) discuss biases and mean square errors of Sen's (1960) estimator, the jackknife estimator and an unbiased estimator which is constituted from unbiased estimators of each term of the variance expression. It is easy to see that all above estimators have first order consistency, which means that the normalized estimators converge to the dominant term $r^2\xi_1^2$ of the variance. Shirahata and Sakamoto (1992) have compared several estimators (unbiased estimator, jackknife estimator, bias modified estimator, and iterated bootstrap and bootstrap estimators) by computer simulations. They have also discussed exact representations of the estimators and reduction of the order of summands to compute the variance estimators.

Using the asymptotic representation of the jackknife variance estimator with the residual term $o_p(n^{-1/2})$, Maesono (1995b) has obtained an Edgeworth expansion with remainder term $o(n^{-1/2})$ for the studentized U-statistic. Obtaining the asymptotic representation of the variance estimator with residual term $o_p(n^{-1})$, where

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

Maesono (1996a) has investigated the Edgeworth expansion of the studentized U-statistic with remainder term $o(n^{-1})$. He has also proved the Edgeworth expansion with remainder term $o(n^{-1/2})$ for the jackknife variance estimator $\hat{\sigma}_J^2$. Further, Maesono (1996b) has discussed the expansion for a linear combination of U-statistics.

In this paper we will study the variance estimators more precisely and obtain asymptotic representations of the normalized estimators with residual terms $n^{-1/2}o_p(n^{-1})$. We show that the unbiased estimator of Schucany and Bankson (1989) coincides with the Hinkley's (1978) corrected jackknife estimator until the order $n^{-1/2}o_p(n^{-1})$. Using the asymptotic representations we obtain asymptotic mean square errors of the variance estimators. We also propose a new variance estimator and obtain its mean square error. We establish Edgeworth expansions of those variance estimators with remainder term $o(n^{-1})$.

In Section 2, we will review the variance estimators and propose the new estimator. In Section 3, we will obtain the asymptotic representations of the estimators and discuss the asymptotic mean square errors. The Edgeworth expansions of them are established in Section 4.

Hereafter for the sake of simplicity, we will consider the kernel of degree 2. The generalization to the kernel with arbitrary degree will be obtained with notational complications and tedious calculations.

2. Variance estimators

At first we will obtain the H-decomposition or the $ANOVA$-decomposition for the U-statistic. Under the assumption that $E|h(X_1, X_2)| < \infty$, let us define

$$g_1(x) = E[h(x, X_2)] - \theta, \qquad g_2(x, y) = h(x, y) - \theta - g_1(x) - g_1(y),$$
$$A_1 = \sum_{i=1}^{n} g_1(X_i) \qquad \textit{and} \qquad A_2 = \sum_{C_{n,2}} g_2(X_i, X_j).$$

Then we have

$$U_n - \theta = \frac{2}{n}A_1 + \frac{2}{n(n-1)}A_2.$$

Note that

$$E[g_2(X_1, X_2)|X_1] = 0 \qquad a.s.$$

Then if one of $\{i_1, i_2\}$ is not contained in $\{j_1, \cdots, j_m\}$, for any m-variate function ν which satisfies $E|\nu g_2| < \infty$, we get

$$E[g_k(X_{i_1}, X_{i_2})\nu(X_{j_1}, \cdots, X_{j_m})] = 0. \tag{2.1}$$

Using this equation we have the variance σ_n^2 of U_n

$$\sigma_n^2 = \frac{4}{n}\xi_1^2 + \frac{2}{n(n-1)}\xi_2^2$$

where

$$\xi_1^2 = E[g_1^2(X_1)] \quad \text{ and } \quad \xi_2^2 = E[g_2^2(X_1, X_2)].$$

Since we discuss the asymptotic properties, we will study the estimation of $n\sigma_n^2$. Then we consider the jackknife variance estimator $V_J = n\hat{\sigma}_J^2$. From the viewpoint of estimation for $4\xi_1^2$, Sen (1960, 1977) has proposed the variance estimator V_S

$$V_S = \frac{4}{n-1}\sum_{i=1}^{n}(S_i - U_n)^2$$

where

$$S_i = \frac{1}{n-1}\sum_{j=1,\neq i}^{n} h(X_i, X_j).$$

Sen (1977) also showed that

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

As pointed out by Efron (1987, p.200), changing coefficients of the estimators will have significantly different effects on the small sample performance of the estimators. Since V_J has positive bias and $(n-2)^2/(n-1)^2 = 1 - 2/n + O(n^{-2})$, we consider the new variance estimator V_α given by

$$V_\alpha = \left(1 - \frac{\alpha}{n}\right)V_J \quad \text{ for } \quad \alpha \geq 0.$$

Note that V_2 and V_S are asymptotically equivalent and $V_0 = V_J$. If we choose α properly, we can reduce the bias and the mean square error, which we will discuss in Section 3.

Hinkley (1978) has discussed the bias correction of V_J. 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 from the sample. Then the bias corrected jackknife estimator is given by

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

where $\bar{Q} = \sum_{C_{n,2}} Q_{i,j}/[n(n-1)]$.

Schucany and Bankson (1989) proposed the unbiased estimator of $n\sigma_n^2$, which is constituted from unbiased estimators of each term of the variance expression. Another variance expression of $n\sigma_n^2$ is

$$n\sigma_n^2 = \frac{4(n-2)}{n-1}a_1^2 + \frac{2}{n-1}a_2^2 \tag{2.3}$$

where

$$a_1^2 = E[h(X_1,X_2)h(X_1,X_3)] - \theta^2 \qquad \textit{and} \qquad a_2^2 = E[h^2(X_1,X_2)] - \theta^2.$$

Let us define

$$\begin{aligned}\zeta_0(x_1,x_2,x_3,x_4) &= \frac{1}{3}\{h(x_1,x_2)h(x_3,x_4) + h(x_1,x_3)h(x_2,x_4) \\ &\qquad + h(x_1,x_4)h(x_2,x_3)\}, \\ \zeta_1(x_1,x_2,x_3) &= \frac{1}{3}\{h(x_1,x_2)h(x_1,x_3) + h(x_1,x_2)h(x_2,x_3) \\ &\qquad + h(x_1,x_3)h(x_2,x_3)\}\end{aligned}$$

and

$$\zeta_2(x_1,x_2) = h^2(x_1,x_2).$$

The unbiased estimators of θ^2, $E[h(X_1,X_2)h(X_1,X_3)]$ and $E[h^2(X_1,X_2)]$ are given by

$$\begin{aligned}\hat{\theta^2} &= \binom{n}{4}^{-1} \sum_{C_{n,4}} \zeta_0(X_{i_1},\cdots,X_{i_4}), \\ \hat{\lambda}_1 &= \binom{n}{3}^{-1} \sum_{C_{n,3}} \zeta_1(X_{i_1},X_{i_2},X_{i_3})\end{aligned}$$

and

$$\hat{\lambda}_2 = \binom{n}{2}^{-1} \sum_{C_{n,2}} \zeta_2(X_{i_1},X_{i_2})$$

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

$$V_U = \frac{4(n-2)}{n-1}\hat{a_1^2} + \frac{2}{n-1}\hat{a_2^2}.$$

Schucany and Bankson (1989) compared the estimators V_J, V_S and V_U by simulation in small samples $n = 10$. We can see that all these estimators converge to $4\xi_1^2$ almost surely. We will study the asymptotic properties of the estimator more precisely.

3. Asymptotic representations and mean square errors

Maesono (1995a) has obtained the asymptotic representations of the variance estimators V_J, V_S, V_C and V_U with residual terms $o_p(n^{-1})$. Here we will consider the asymptotic representations more precisely. Let us define

$$\begin{aligned}\delta(x) &= E[g_2^2(x,X_2)] - \xi_2^2, \\ f_1(x) &= g_1^2(x) - \xi_1^2 + 2E[g_1(X_2)g_2(x,X_2)], \\ f_2(x,y) &= -g_1(x)g_1(y) + g_2(x,y)\{g_1(x)+g_1(y)\} \\ &\qquad + E[g_2(x,X_3)g_2(y,X_3) - g_2(x,X_3)g_1(X_3) - g_2(y,X_3)g_1(X_3)], \\ f_3(x,y,z) &= g_2(x,y)g_2(x,z) + g_2(x,y)g_2(y,z) + g_2(x,z)g_2(y,z) \\ &\qquad - E[g_2(x,X_3)g_2(y,X_3) + g_2(y,X_3)g_2(z,X_3) \\ &\qquad\qquad + g_2(x,X_3)g_2(z,X_3)] \\ &\qquad - 2\{g_1(x)g_2(y,z) + g_1(y)g_2(x,z) + g_1(z)g_2(x,y)\end{aligned}$$

and

$$V_n = \frac{4}{n}\sum_{i=1}^{n} f_1(X_i) + \frac{8}{n(n-1)}\sum_{C_{n,2}} f_2(X_i, X_j) + \frac{8}{n(n-1)(n-2)}\sum_{C_{n,3}} f_3(X_i, X_j, X_k).$$

Note that V_n has already decomposed. For the variance estimators, we have the following representations.

THEOREM 1. *If $E|h(X_1, X_2)|^{4+\varepsilon} < \infty$ for some $\varepsilon > 0$, we have*

$$V_J = V_n + \frac{8}{n^2}\sum_{i=1}^{n}\delta(X_i) + n\sigma_n^2 + \frac{b_J}{n} + R_{1;n}, \tag{3.1}$$

$$V_S = V_n + \frac{8}{n^2}\sum_{i=1}^{n}\{\delta(X_i) - f_1(X_i)\} + n\sigma_n^2 + \frac{b_S}{n} + R_{2;n}, \tag{3.2}$$

$$V_\alpha = V_n + \frac{4}{n^2}\sum_{i=1}^{n}\{2\delta(X_i) - \alpha f_1(X_i)\} + n\sigma_n^2 + \frac{b_\alpha}{n} + R_{3;n}, \tag{3.3}$$

$$V_C = V_n + \frac{4}{n^2}\sum_{i=1}^{n}\delta(X_i) + n\sigma_n^2 + R_{4;n} \tag{3.4}$$

and

$$V_U = V_n + \frac{4}{n^2}\sum_{i=1}^{n}\delta(X_i) + n\sigma_n^2 + R_{5;n} \tag{3.5}$$

where

$$b_J = 2\xi_2^2, \qquad b_S = 2\xi_2^2 - 8\xi_1^2, \quad b_\alpha = 2\xi_2^2 - 4\alpha\xi_1^2$$

and

$$E|R_{k;n}|^{2+\frac{\varepsilon}{2}} = O(n^{-4-\varepsilon}) \qquad (k = 1, \cdots, 5). \tag{3.6}$$

PROOF. See appendix.

b_J, b_S and b_α are n^{-1} biases of the jackknife, the Sen's estimator and the new estimator respectively. Since $R_{k,n} = n^{-1/2}o_p(n^{-1})$, the unbiased estimator V_U coincides the Hinkley's (1978) corrected jackknife estimator V_C until the order $n^{-1/2}o_p(n^{-1})$. It is easy to see that

$$E[f_2(X_1, X_2)|X_1] = E[f_3(X_1, X_2, X_3)|X_1, X_2] = 0 \qquad a.s. \tag{3.7}$$

and $E[f_1(X_1)] = E[\delta(X_1)] = 0$.

Using the asymptotic representations of Theorem 1, we can study the asymptotic properties of the variance estimators. Here we will obtain asymptotic mean square errors of V_J, V_S, V_α, V_C and V_U up to the order n^{-2}. Let us define

$$mse(V_J) = \frac{16}{n}E[f_1^2(X_1)] + \frac{1}{n^2}\{b_J^2 + 64E[f_1(X_1)\delta(X_1)] + 32E[f_2^2(X_1, X_2)]\},$$

$$mse(V_S) = \frac{16}{n}E[f_1^2(X_1)] + \frac{1}{n^2}\{b_S^2 + 64E[f_1(X_1)(\delta(X_1) - f_1(X_1))]$$

$$+ 32E[f_2^2(X_1, X_2)]\},$$
$$mse(V_\alpha) = \frac{16}{n}E[f_1^2(X_1)] + \frac{1}{n^2}\{b_\alpha^2 + 32E[f_1(X_1)(2\delta(X_1) - \alpha f_1(X_1))] + 32E[f_2^2(X_1, X_2)]\}$$

and

$$mse(V_C) = \frac{16}{n}E[f_1^2(X_1)] + \frac{1}{n^2}\{32E[f_1(X_1)\delta(X_1)] + 32E[f_2^2(X_1, X_2)]\}.$$

Note that $mse(V_J) = mse(V_0)$ and $mse(V_S) = mse(V_2)$. We have the following theorem.

THEOREM 2. *If $E|h(X_1, X_2)|^{4+\varepsilon} < \infty$ for some $\varepsilon > 0$, we have*

$$E(V_J - n\sigma_n^2)^2 = mse(V_J) + O(n^{-\frac{5}{2}}),$$
$$E(V_S - n\sigma_n^2)^2 = mse(V_S) + O(n^{-\frac{5}{2}}),$$
$$E(V_\alpha - n\sigma_n^2)^2 = mse(V_\alpha) + O(n^{-\frac{5}{2}}),$$
$$E(V_C - n\sigma_n^2)^2 = mse(V_C) + O(n^{-\frac{5}{2}})$$

and

$$E(V_U - n\sigma_n^2)^2 = mse(V_C) + O(n^{-\frac{5}{2}}).$$

PROOF. It follows from (3.6) and (A.2) in Lemma 1 (see Appendix) that under the moment condition, for $1 \leq k \leq 5$,

$$E|n^{-1}R_{k;n}\sum_{i=1}^n f_1(X_i)| \leq n^{-1}\left\{E\left|\sum_{i=1}^n f_1(X_i)\right|^{2+\frac{\varepsilon}{2}} E|R_{k;n}|^{2+\frac{\varepsilon}{2}}\right\}^{\frac{2}{4+\varepsilon}} = O(n^{-\frac{5}{2}}),$$

$$E|n^{-2}R_{k;n}\sum_{i=1}^n \delta(X_i)| \leq n^{-2}\left\{E\left|\sum_{i=1}^n \delta(X_i)\right|^{2+\frac{\varepsilon}{2}} E|R_{k;n}|^{2+\frac{\varepsilon}{2}}\right\}^{\frac{2}{4+\varepsilon}} = O(n^{-3}),$$

$$E\left|n^{-2}R_{k;n}\sum_{C_{n,2}} f_2(X_i, X_j)\right| \leq n^{-2}\left\{E\left|\sum_{C_{n,2}} f_2(X_i, X_j)\right|^{2+\frac{\varepsilon}{2}} E|R_{k;n}|^{2+\frac{\varepsilon}{2}}\right\}^{\frac{2}{4+\varepsilon}} = O(n^{-3})$$

and

$$E|R_{k;n}|^2 \leq \{E|R_{k;n}|^{2+\frac{\varepsilon}{2}}\}^{\frac{4}{4+\varepsilon}} = O(n^{-4}).$$

Thus, using these equations and (3.7), we can obtain the equalities.

REMARK 1. It is possible to improve the equations with remainder terms of the order $O(n^{-3})$. But it needs more calculation, then we leave the equations as they are.

Let us define

$$e_1 = E[g_1^4(X_1)], \quad e_2 = E[g_1^2(X_1)g_2^2(X_1, X_2)],$$
$$e_3 = E[g_1^2(X_1)g_1(X_2)g_2(X_1, X_2)],$$
$$e_4 = E[g_1(X_1)g_1(X_2)g_2^2(X_1, X_2)],$$
$$e_5 = E[g_1(X_1)g_1(X_2)g_2(X_1, X_3)g_2(X_2, X_3)],$$
$$e_6 = E[g_1(X_1)g_2(X_1, X_2)g_2(X_1, X_3)g_2(X_2, X_3)]$$

and

$$e_7 = E[g_2(X_1, X_3)g_2(X_2, X_3)g_2(X_1, X_4)g_2(X_2, X_4)].$$

Then, using the equation (2.1), it follows from direct computations that

$$E[f_1^2(X_1)] = e_1 - \xi_1^4 + 4e_3 + 4e_5,$$
$$E[f_1(X_1)\delta(X_1)] = e_2 - \xi_1^2\xi_2^2 + 2e_3$$

and

$$E[f_2^2(X_1, X_2)] = \xi_1^4 + 2e_2 - 4e_3 + 2e_4 - 4e_5 + 4e_6 + e_7.$$

Here we will study the asymptotic mean square errors for the variance and the covariance estimation problems. Also, the asymptotic mean square error of the Wilcoxon's signed rank test will be discussed.

Example 1. *Variance estimation;*

Let us consider the kernel $h(x, y) = (x - y)^2/2$. Then if $Var(X_1) = \sigma^2$, 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. It is easy to see that

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

For the sake of simplicity, we will consider the case that the distribution $F(x)$ is symmetric about the origin. Let us define $m_k = E[X_1^k]$. Then because of symmetry of F, if k is odd number, $m_k = 0$. Using this fact, from direct computations, we can show that

$$\xi_1^2 = \frac{1}{4}(m_4 - \sigma^4), \qquad \xi_2^2 = \sigma^4, \quad e_1 = \frac{1}{16}(m_8 - 4\sigma^2 m_6 + 6\sigma^4 m_4 - 3\sigma^8),$$
$$e_2 = \frac{\sigma^2}{4}(m_6 - 2\sigma^2 m_4 + \sigma^6), \qquad e_4 = \frac{1}{4}(m_4 - \sigma^4)^2, \quad e_6 = -\frac{\sigma^4}{2}(m_4 - \sigma^4),$$
$$e_7 = \sigma^8 \quad \textit{and} \quad e_3 = e_5 = 0.$$

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

$$b_J = 2\sigma^4, \qquad b_S = -2\sigma^4, \quad b_\alpha = 2\sigma^4 - 2\alpha\sigma^4,$$
$$mse(V_J) = \sigma^8 \left\{ \frac{56}{n} + \frac{268}{n^2} \right\}, \qquad mse(V_S) = \sigma^8 \{ \frac{56}{n} + \frac{44}{n^2} \},$$
$$mse(V_\alpha) = \sigma^8 \left\{ \frac{56}{n} + \frac{4(\alpha^2 - 30\alpha + 67)}{n^2} \right\} \textit{ and } mse(V_C) = \sigma^8 \left\{ \frac{56}{n} + \frac{200}{n^2} \right\}.$$

In the case of $\sigma^2 = 1$ and $n = 10$, Schucany and Bankson (1989) discussed the mean square errors of $V_J/n, V_S/n$ and V_C/n by simulation. Corresponding asymptotic mean square errors are given by

$$\frac{mse(V_J)}{10^2} = 0.0828, \qquad \frac{mse(V_S)}{10^2} = 0.0604 \quad \textit{and} \quad \frac{mse(V_C)}{10^2} = 0.0760.$$

Their estimated mean square errors are close to these values.

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

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

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

$$b_J = 2\sigma^4, \quad b_S = -\frac{22}{5}\sigma^4, \quad b_\alpha = 2\sigma^4 - \frac{16\alpha}{5}\sigma^4,$$

$$mse(V_J) = \sigma^8\left\{\frac{538.33}{n} + \frac{1002.95}{n^2}\right\}, \quad mse(V_S) = \sigma^8\left\{\frac{538.33}{n} - \frac{1135.02}{n^2}\right\},$$

$$mse(V_\alpha) = \sigma^8\left\{\frac{538.33}{n} - \frac{1}{n^2}(10.24\alpha^2 - 1089.46\alpha + 1002.95)\right\}$$

and

$$mse(V_U) = \sigma^8\left\{\frac{538.33}{n} + \frac{764.89}{n^2}\right\}.$$

EXAMPLE 2. *Covariance estimation;*

Let $\{\boldsymbol{X}_i\}_{i\geq 1}$ be two dimensional random vectors. And putting $\boldsymbol{X}_i = (Y_i, Z_i)$, we denote

$$Var(\boldsymbol{X}_1) = Var\{(Y_1, Z_1)\} = \begin{pmatrix} \sigma_y^2 & \rho\sigma_y\sigma_z \\ \rho\sigma_y\sigma_z & \sigma_z^2 \end{pmatrix}.$$

Let us consider a symmetric kernel $h(\boldsymbol{x}_1, \boldsymbol{x}_2) = (y_1 - y_2)(z_1 - z_2)/2$. Then corresponding U-statistic is an unbiased estimator of $\rho\sigma_y\sigma_z = Cov(Y_1, Z_1)$. It is easy to see that

$$\theta = \rho\sigma_y\sigma_z, \quad g_1(\boldsymbol{x}_1) = \frac{1}{2}(y_1 z_1 - \rho\sigma_y\sigma_z) \quad and \quad g_2(\boldsymbol{x}_1, \boldsymbol{x}_2) = -\frac{1}{2}(y_1 z_2 + z_1 y_2).$$

Further we assume that $\boldsymbol{X}_i$ is bivariate normal distribution

$$\boldsymbol{X}_i = (Y_i, Z_i) \sim N\left(\mu, \begin{pmatrix} \sigma_y^2 & \rho\sigma_y\sigma_z \\ \rho\sigma_y\sigma_z & \sigma_z^2 \end{pmatrix}\right).$$

From direct computations we can get

$$\xi_1^2 = \frac{1+\rho^2}{4}\sigma_y^2\sigma_z^2, \quad \xi_2^2 = \frac{1+\rho^2}{2}\sigma_y^2\sigma_z^2, \quad e_1 = \frac{3}{16}(3\rho^4 + 14\rho^2 + 3)\sigma_y^4\sigma_z^4,$$

$$e_2 = \frac{1}{8}(3\rho^4 + 14\rho^2 + 3)\sigma_y^4\sigma_z^4, \quad e_3 = 0, \quad e_4 = \frac{1}{8}(\rho^4 + 6\rho^2 + 1)\sigma_y^4\sigma_z^4,$$

$$e_5 = 0, \quad e_6 = -\frac{1}{8}(\rho^4 + 6\rho^2 + 1)\sigma_y^4\sigma_z^4 \quad and \quad e_7 = \frac{1}{8}(\rho^4 + 6\rho^2 + 1)\sigma_y^4\sigma_z^4.$$

Thus we have

$$b_J = \sigma_y^2\sigma_z^2(1+\rho^2), \quad b_S = -\sigma_y^2\sigma_z^2(1+\rho^2), \quad b_\alpha = (1-\alpha)(1+\rho^2)\sigma_y^2\sigma_z^2,$$

$$mse(V_J) = \sigma_y^4\sigma_z^4\left\{\frac{8}{n}(\rho^4 + 5\rho^2 + 1) + \frac{1}{n^2}(39\rho^4 + 190\rho^2 + 39)\right\},$$

$$mse(V_S) = \sigma_y^4\sigma_z^4\left\{\frac{8}{n}(\rho^4 + 5\rho^2 + 1) + \frac{1}{n^2}(7\rho^4 + 30\rho^2 + 7)\right\}$$

$$mse(V_\alpha) = \sigma_y^4\sigma_z^4\left\{\frac{8}{n}(\rho^4 + 5\rho^2 + 1) + \frac{1}{n^2}[\alpha^2(\rho^2+1)^2 - 6\alpha(3\rho^4 + 14\rho^2 + 3) + 39\rho^4 + 190\rho^2 + 39]\right\}$$

and

$$mse(V_C) = \sigma_y^4\sigma_z^4\{\frac{8}{n}(\rho^4+5\rho^2+1)+\frac{1}{n^2}(30\rho^4+140\rho^2+30)\}.$$

REMARK 2. In the cases of the above tow examples, $mse(V_S) < mse(V_C) < mse(V_J)$. As discussed in Schucany and Bankson (1989), though Sen's estimator V_S has small mean square error, it has substantial negative bias. V_C and V_U are asymptotically unbiased and have smaller mean square error than V_J. But V_C and V_U sometimes take negative values in small sample case. Schucany and Bankson (1989) also pointed out by simulation that from the viewpoint of Pitman closeness V_J is closer to $n\sigma_n^2$ than V_U. If we take $\alpha = 1$ of V_α, both biases and mean square errors are relatively small. Especially in the case of the normal distribution, the biases of V_1 are 0. Note that V_1 is asymptotically equivalent to $(V_J + V_S)/2$ and always takes a non-negative value.

EXAMPLE 3. *Wilcoxon's signed rank test;*

In order to compare the mean square errors of the variance estimators, let us discuss the variance estimation of the Wilcoxon's signed rank statistic. Let $X_1, \cdots, X_n$ be a random sample from the distribution $F(x-\eta)$, where $F(x)$ satisfies $F(-x) = 1 - F(x)$ for any x. So, the distribution $F(x)$ is symmetric about origin. The Wilcoxon's signed rank statistic is very popular to test or to estimate η. For the sake of simplicity, we consider the following statistic

$$M_n = \binom{n}{2}^{-1}\sum_{C_{n,2}}\Psi(X_i+X_j)$$

where $\Psi(x) = 1,\ 0$ if $x \geq 0,\ < 0$. M_n is asymptotically equivalent to the Wilcoxon's statistic. Let us assume $\eta = 0$ and $F(x)$ has a density function. From direct computation, we can show that

$$\begin{aligned}
&\theta = \frac{1}{2}, \quad \sigma_n^2 = \frac{2n-1}{6n(n-1)}, \quad g_1(x) = F(x) - \frac{1}{2},\ \xi_1^2 = \frac{1}{12},\ \xi_2^2 = \frac{1}{12},\\
&e_1 = \frac{1}{80}, \quad e_2 = \frac{1}{144}, \quad e_3 = -\frac{1}{360},\ e_4 = -\frac{1}{360},\ e_5 = \frac{1}{720},\\
&e_6 = -\frac{1}{1440} \quad \textit{and} \quad e_7 = \frac{1}{720}.
\end{aligned}$$

Thus we have

$$\begin{aligned}
&b_J = \frac{1}{6}, \quad b_S = -\frac{1}{2}, \quad b_\alpha = \frac{1}{6} - \frac{\alpha}{3},\\
&mse(V_J) = \frac{53}{180n^2}, \quad mse(V_S) = \frac{31}{60n^2},\\
&mse(V_\alpha) = \frac{1}{n^2}\left[\frac{1}{9}\left(\alpha - \frac{1}{2}\right)^2 + \frac{4}{15}\right] \quad \textit{and} \quad mse(V_C) = \frac{4}{9n^2}.
\end{aligned}$$

It follows from the above calculation that

$$mse(V_J) < mse(U_C) < mse(V_S).$$

When $\alpha = 1/2$, $mse(V_{1/2})$ takes a minimum value $4/(15n^2)$ and $mse(V_{1/2}) < mse(V_J)$.

REMARK 3. Example 1 and Example 2 give us the same conclusion. But it is different in the case of the Wilcoxon's statistic. So, we had better to check the mean square errors of the variance estimators using Theorem 2 in each case.

4. Edgeworth expansions

From Theorem 1, we can regard the variance estimators as sum of U-statistics and n^{-1} term. For asymptotic U-statistics, Lai and Wang (1993) have established the Edgeworth expansion with remainder term $o(n^{-1})$. Applying their result, we can get Edgeworth expansions of the variance estimators. Let us assume the following conditions.

(C_1) $E|h(X_1, X_2)|^8 < \infty$

(C_2) $\limsup_{|t|\to\infty} |E[\exp\{itf_1(X_1)\}]| < 1$

(C_3) $E|f_2(X_1, X_2)|^s < \infty$ $(s > 0)$ and there exist K Borel functions $\psi_\nu:\boldsymbol{R} \to \boldsymbol{R}$ such that $E[\psi_\nu^2(X_1)] < \infty(\nu = 1, \cdots, K)$, $K(s-2) > 4s+(28s-40)I_{\{E|f_3(X_1,X_2,X_3)|>0\}}$, and the covariance matrix of $(W_1, \cdots, W_K)$ is positive definite, where $W_\nu = (L\psi_\nu)(X_1)$ and $(L\psi_\nu)(y) = E[f_2(y, X_2)\psi_\nu(X_2)]$, and $I_{\{.\}}$ is an indicator function.

The condition C_3 is concerned with the number of nonzero eigen function of $f_2(x, y)$. 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 $f_2(X_1, X_2) = \sum_{\nu=1}^K c_\nu w_\nu(X_1)w_\nu(X_2) a.s.$; moreover, for some $0 < \gamma < \min\{1, 2(1 - 11/(3s))\}$,

$$\limsup_{|t|\to\infty} \sup_{|u_1|+\cdots+|u_K|\leq|t|^{-\gamma}} \left| E\left[\exp\left(it\left\{f_1(X_1) + \sum_{\nu=1}^K u_\nu w_\nu(X_1)\right\}\right)\right]\right| < 1.$$

Let us define

$$\begin{aligned}
&\tau^2 = E[f_1^2(X_1)], \quad d_1 = E[f_2^2(X_1, X_2)], \quad d_2 = E[f_1(X_1)\delta(X_1)],\\
&d_3 = E[f_1^3(X_1)], \quad d_4 = E[f_1(X_1)f_1(X_2)f_2(X_1, X_2)],\\
&d_5 = E[f_1^4(X_1)], \quad d_6 = E[f_1^2(X_1)f_1(X_2)f_2(X_1, X_2)],\\
&d_7 = E[f_1(X_1)f_1(X_2)f_2(X_1, X_3)f_2(X_2, X_3)],\\
&d_8 = E[f_1(X_1)f_1(X_2)f_1(X_3)f_3(X_1, X_2, X_3)],\\
&\kappa_3 = \tau^{-3}(d_3 + 12d_4), \quad \kappa_4 = \tau^{-4}(d_5 - 3\tau^4 + 24d_6 + 48d_7 + 8d_8),\\
&P_{1C}(x) = \frac{x^2-1}{6}\kappa_3, \quad P_{1J}(x) = P_{1C}(x) - \frac{b_J}{4\tau},\\
&P_{1S}(x) = P_{1C}(x) - \frac{b_S}{4\tau}, \quad P_{1\alpha}(x) = P_{1C}(x) - \frac{b_\alpha}{4\tau},\\
&P_{2J}(x) = \frac{x}{\tau^2}\left(d_1 + 2d_2 + \frac{b_J^2}{32}\right) + \frac{\kappa_4 - b_J\kappa_3}{24}(x^3 - 3x)\\
&\qquad + \frac{\kappa_3^2}{72}(x^5 - 10x^3 + 15x),\\
&P_{2S}(x) = \frac{x}{\tau^2}\left(d_1 + 2d_2 - 2\tau^2 + \frac{b_S^2}{32}\right) + \frac{\kappa_4 - b_S\kappa_3}{24}(x^3 - 3x)\\
&\qquad + \frac{\kappa_3^2}{72}(x^5 - 10x^3 + 15x),\\
&P_{2\alpha}(x) = \frac{x}{\tau^2}\left(d_1 + 2d_2 - \alpha\tau^2 + \frac{b_\alpha^2}{32}\right) + \frac{\kappa_4 - b_S\kappa_3}{24}(x^3 - 3x)\\
&\qquad + \frac{\kappa_3^2}{72}(x^5 - 10x^3 + 15x)
\end{aligned}$$

and

$$P_{2C}(x) = \frac{x}{\tau^2}(d_1 + d_2) + \frac{\kappa_4}{24}(x^3 - 3x) + \frac{\kappa_3^2}{72}(x^5 - 10x^3 + 15x).$$

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*

$$P\left\{\frac{\sqrt{n}(V_J - n\sigma_n^2)}{4\tau} \le x\right\} = \Phi(x) - \frac{\phi(x)P_{1J}(x)}{\sqrt{n}} - \frac{\phi(x)P_{2J}(x)}{n} + o(n^{-1}),$$
$$P\left\{\frac{\sqrt{n}(V_S - n\sigma_n^2)}{4\tau} \le x\right\} = \Phi(x) - \frac{\phi(x)P_{1S}(x)}{\sqrt{n}} - \frac{\phi(x)P_{2S}(x)}{n} + o(n^{-1}),$$
$$P\left\{\frac{\sqrt{n}(V_\alpha - n\sigma_n^2)}{4\tau} \le x\right\} = \Phi(x) - \frac{\phi(x)P_{1\alpha}(x)}{\sqrt{n}} - \frac{\phi(x)P_{2\alpha}(x)}{n} + o(n^{-1}),$$
$$P\left\{\frac{\sqrt{n}(V_C - n\sigma_n^2)}{4\tau} \le x\right\} = \Phi(x) - \frac{\phi(x)P_{1C}(x)}{\sqrt{n}} - \frac{\phi(x)P_{2C}(x)}{n} + o(n^{-1})$$

and

$$P\left\{\frac{\sqrt{n}(V_U - n\sigma_n^2)}{4\tau} \le x\right\} = \Phi(x) - \frac{\phi(x)P_{1C}(x)}{\sqrt{n}} - \frac{\phi(x)P_{2C}(x)}{n} + o(n^{-1}).$$

PROOF. It is sufficient to prove the case of V_J. Since $\tilde{U}_n = \sqrt{n}\{V_n + 8\sum_{i=1}^n \delta(X_i)/n^2 + R_{1;n}\}$ is an asymptotic U-statistic, it follows from Lai and Wang (1993) that

$$P\left\{\frac{\tilde{U}_n}{4\tau} \le x\right\} = \Phi(x) - \frac{\phi(x)P_{1C}(x)}{\sqrt{n}} - \frac{\phi(x)\tilde{P}_2(x)}{n} + o(n^{-1})$$

where

$$\tilde{P}_2(x) = \frac{x}{\tau^2}(d_1 + 2d_2) + \frac{\kappa_4}{24}(x^3 - 3x) + \frac{\kappa_3^2}{72}(x^5 - 10x^3 + 15x).$$

Since

$$P\left\{\frac{\sqrt{n}(V_J - n\sigma_n^2)}{4\tau} \le x\right\} = P\left\{\frac{\tilde{U}_n}{4\tau} \le x - \frac{b_J}{4\tau\sqrt{n}}\right\},$$

expanding by $b_J/(4\tau\sqrt{n})$, we have the Edgeworth expansion for V_J.

EXAMPLE 4. Let us consider the case of variance estimation in Example 1. From direct computation, we can show that

$$f_1(x) = \frac{1}{4}(x^2 - \sigma^2) - \frac{1}{4}\xi_1^2$$

and

$$\begin{aligned} f_2(x, y) &= -\frac{1}{4}(x^2 - \sigma^2)(y^2 - \sigma^2) - \frac{xy}{2}(x^2 + y^2 - 2\sigma^2) + \sigma^2 xy \\ &= -\frac{1}{4}(x^2 - \sigma^2)(y^2 - \sigma^2) - \frac{1}{2}(x^3 + x)(y^3 + y) + \frac{1}{2}x^3y^3 \\ &\quad + \left(2\sigma^2 + \frac{1}{2}\right)xy. \end{aligned}$$

Thus putting

$$c_1 = -\frac{1}{4}, \quad w_1(x) = x^2 - \sigma^2, \quad c_2 = -\frac{1}{2}, \ \ w_2(x) = x^3 + x,$$
$$c_3 = \frac{1}{2}, \quad w_3(x) = x^3, \quad c_4 = 2\sigma^2 + \frac{1}{2} \ \ and \ \ w_4(x) = x,$$

we have

$$f_2(X_1, X_2) = \sum_{\nu=1}^{3} c_\nu w_\nu(X_1) w_\nu(X_2) \qquad a.s.$$

Assume that $E|X_1|^{22} < \infty$ and the underlying distribution $F(x)$ has a density function. We can show that

$$\limsup_{|t|\to\infty} \sup_{|u_1|+\cdots+|u_4|\le|t|^{-1}} \left| E\left[\exp\left(it\left\{f_1(X_1) + \sum_{\nu=1}^{4} u_\nu w_\nu(X_1)\right\}\right)\right]\right| < 1.$$

Hence the conditions (C_1), (C_2) and $(\tilde{C}_3)$ are satisfied.

Appendix A

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

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

and

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

Then we can show that

$$E[\rho_k(X_1, \cdots, X_k) | X_1, \cdots, X_{k-1}] = 0 \qquad a.s. \tag{A.1}$$

and

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

where

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

Using the equation (11) and moment evaluations of martingales (Dharmadhikari, Fabian and Jogdeo (1968)), we have the upper bounds of the absolute moments of Λ_k as follows.

LEMMA 1. *For $q \ge 2$, if $E|\nu(X_1, \cdots, X_r)|^q < \infty$, there exists a positive constant c, which may depend on ν and F but not on n, such that*

$$E|\Lambda_k|^q \le c n^{\frac{qk}{2}}. \tag{A.2}$$

For the simplicity we use a symbol $o_p^*(n^{-3/2})$ which may be different in each case but satisfies

$$E|o_p^*(n^{-\frac{3}{2}})|^{2+\frac{\varepsilon}{2}} = O(n^{-4-\varepsilon}).$$

It follows from Markov's inequality that $o_p^*(n^{-3/2}) = n^{-1/2}o_p(n^{-1})$.

From Markov's inequality and (12), we can easily obtain the following lemma which is useful for obtaining the asymptotic representations.

LEMMA 2. *If* $E[\nu(X_1, \cdots, X_r)] = 0$ *and* $E|\nu(X_1, \cdots, X_r)|^{2+\varepsilon} < \infty$ *for* $\varepsilon > 0$*, we have that*

$$n^{-r-1} \sum_{C_{n,r}} \nu(X_{i_1}, \cdots, X_{i_r}) = \frac{1}{n^2(r-1)!}\Lambda_1 + o_p^*(n^{-\frac{3}{2}}). \tag{A.3}$$

and

$$n^{-r} \sum_{k=4}^{r} \binom{n-k}{r-k} \Lambda_k = o_p^*(n^{-\frac{3}{2}}). \tag{A.4}$$

Using the above lemmas, we will prove Theorem 2.

Approximation of V_J

At first we will obtain the approximation of V_J. Let us define

$$\begin{aligned}
D_1 &= \sum_{i=1}^{n} g_1^2(X_i), \qquad D_2 = \sum_{C_{n,2}} g_1(X_i)g_1(X_j), \\
D_3 &= \sum_{C_{n,2}} \{g_1(X_i) + g_1(X_j)\}g_2(X_i, X_j), \\
D_4 &= \sum_{C_{n,3}} \{g_1(X_i)g_2(X_j, X_k) + g_1(X_j)g_2(X_i, X_k) + g_1(X_k)g_2(X_i, X_j)\}, \\
D_5 &= \sum_{C_{n,2}} g_2^2(X_i, X_j), \\
D_6 &= \sum_{C_{n,3}} \{g_2(X_i, X_j)g_2(X_i, X_k) + g_2(X_i, X_j)g_2(X_j, X_k) \\
&\qquad + g_2(X_i, X_k)g_2(X_j, X_k)\}
\end{aligned}$$

and

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

From Maesono (1995, p.18), we have

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

Note that $V_J = (n-1)\sum_{i=1}^{n}(U_n^{(i)} - U_n)^2$. Using the H-decomposition, Lemma 1 and Lemma 2, we have that

$$\begin{aligned}
\frac{4}{n}D_1 &= 4\xi_1^2 + \frac{4}{n}\sum_{i=1}^{n}\{g_1^2(X_i) - \xi_1^2\},\\
\frac{8}{n(n-1)}D_3 &= \frac{8}{n}\sum_{i=1}^{n}E[g_1(X_0)g_2(X_i,X_0)|X_i]\\
&\quad + \frac{8}{n(n-1)}\sum_{C_{n,2}}\{[g_1(X_i)+g_1(X_j)]g_2(X_i,X_j)\\
&\quad - E[g_1(X_0)g_2(X_i,X_0)|X_i] - E[g_1(X_0)g_2(X_j,X_0)|X_j]\},\\
\frac{8}{n(n-1)(n-2)}D_5 &= \frac{4\xi_2^2}{n} + \frac{8}{n^2}\sum_{i=1}\delta(X_i) + o_p^*(n^{-\frac{3}{2}}),\\
\frac{8(n-4)}{n(n-1)(n-2)^2}D_6 &= \frac{8}{n(n-1)}\sum_{C_{n,2}}E[g_2(X_i,X_0)g_2(X_j,X_0)|X_i,X_j]\\
&\quad + \frac{8}{n(n-1)(n-2)}\sum_{C_{n,3}}\beta(X_i,X_j,X_k) + o_p^*(n^{-\frac{3}{2}})
\end{aligned}$$

and

$$\frac{16}{n(n-1)(n-2)^2}D_7 = o_p^*(n^{-\frac{3}{2}})$$

where X_0 is a random vector with distribution $F(x)$ and is independent of $X_1, \cdots, X_n$ and

$$\begin{aligned}
\beta(x,y,z) &= g_2(x,y)g_2(x,z) + g_2(x,y)g_2(y,z) + g_2(x,z)g_2(y,z)\\
&\quad - E[g_2(x,X_0)g_2(y,X_0) + g_2(y,X_0)g_2(z,X_0) + g_2(x,X_0)g_2(z,X_0)].
\end{aligned}$$

Thus we have the equation (3.1).

Approximation of V_S

It follows from the equation (2.2) that

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

From the equation (A.3), we can show that

$$-\frac{2}{n}V_J = -\frac{8\xi_1^2}{n} - \frac{8}{n^2}\sum_{i=1}^{n}f_1(X_i) + o_p^*(n^{-\frac{3}{2}})$$

and

$$O(n^{-2})V_J = o_p^*(n^{-\frac{3}{2}}).$$

Thus we have the equation (3.2).

Approximation of V_α

Similarly as V_S, we can easily obtain the equation (3.3).

Approximation of V_C

To obtain the equation (3.4), it is sufficient to prove the following lemma which is an improvement of Lemma A4 in Maesono (1995).

LEMMA 3. *If* $E|h(X_1, X_2)|^{4+\varepsilon} < \infty$ *for some* $\varepsilon > 0$, *we have*

$$\frac{1}{n+1} \sum_{1\le i<j\le n} (Q_{i,j} - \bar{Q})^2 = \frac{2\xi_2^2}{n} + \frac{4}{n^2}\sum_{i=1}^{n} \delta(X_i) + o_p^*(n^{-\frac{3}{2}}).$$

PROOF. From the proof of Lemma A4 in Maesono (1995), we have

$$\begin{aligned}
&\frac{1}{n+1} \sum_{1\le i<j\le n} (Q_{i,j} - \bar{Q})^2 \\
&\quad = \frac{4}{(n+1)(n-1)(n-3)} D_5 - \frac{8}{(n+1)(n-1)(n-2)(n-3)} D_6 \\
&\qquad + \frac{16}{(n+1)(n-1)(n-2)(n-3)^2} D_7.
\end{aligned}$$

Using the H-decomposition and the equations (A.3) and (A.4), we get that

$$\frac{4}{(n+1)(n-1)(n-3)} D_5 = \frac{2\xi_2^2}{n} + \frac{4}{n^2}\sum_{i=1}^{n} \delta(X_i) + o_p^*(n^{-\frac{3}{2}}),$$

$$\frac{8}{(n+1)(n-1)(n-2)(n-3)} D_6 = o_p^*(n^{-\frac{3}{2}})$$

and

$$\frac{8}{(n+1)(n-1)(n-2)(n-3)^2} D_7 = o_p^*(n^{-\frac{3}{2}}).$$

Thus we have the equation (3.5).

Approximation of V_U

Finally we will consider the unbiased estimator V_U. We will obtain approximations of $\hat{a_1^2}$ and $\hat{a_2^2}$. Let us consider $\hat{\lambda}_1$. From the definition, we can get

$$E[h(x, X_2)] = g_1(x) + \theta, \qquad h(x, y) = g_2(x, y) + g_1(x) + g_1(y) + \theta.$$

Using these equations and (2.1), we can show that

$$\begin{aligned}
E[\zeta_1(x, y, X_3)] = \frac{1}{3}\{&g_2(x, y)[g_1(x) + g_1(y)] + 3g_1(x)g_1(y) + g_1^2(x) + g_1^2(y) \\
&+ E[g_2(x, X_3)g_2(y, X_3) + (g_2(x, X_3) + g_2(y, X_3))g_1(X_3)] \\
&+ 2\theta g_2(x, y) + 4\theta g_1(x) + 4\theta g_1(y) + \xi_1^2 + 3\theta^2\}.
\end{aligned}$$

We also have

$$\begin{aligned}
&E[\zeta_1(x, X_2, X_3)] \\
&\quad = \frac{2}{3}\{E[g_2(x, X_3)g_1(X_3)] + \xi_1^2\} + \frac{4}{3}\theta g_1(x) + \theta^2 + \frac{1}{3}g_1^2(x)
\end{aligned}$$

and $E[\zeta_1(X_1, X_2, X_3)] = \xi_1^2 + \theta^2$. Here we have

$$\begin{aligned}
&E[\zeta_1(x, X_2, X_3)] - \xi_1^2 - \theta^2 \\
&\quad = \frac{2}{3}E[g_2(x, X_2)g_1(X_2)] + \frac{4}{3}\theta g_1(x) + \frac{1}{3}\{g_1^2(x) - \xi_1^2\} \\
&\quad = \tilde{g}_1(x) \qquad (say), \\
&E[\zeta_1(x, y, X_3)] - \xi_1^2 - \theta^2 - \tilde{g}_1(x) - \tilde{g}_1(y) \\
&\quad = \frac{1}{3}E[g_2(x, X_3)g_2(y, X_3) - \{g_2(x, X_3) + g_2(y, X_3)\}g_1(X_3)] \\
&\qquad + g_1(x)g_1(y) + \frac{1}{3}\{g_1(x) + g_1(y) + 2\theta\}g_2(x, y) \\
&\quad = \tilde{g}_2(x, y) \qquad (say)
\end{aligned}$$

and

$$\begin{aligned}
&\zeta_1(x, y, z) - \xi_1^2 - \theta^2 - \tilde{g}_2(x, y) - \tilde{g}_2(x, z) - \tilde{g}_2(y, z) \\
&\qquad - \tilde{g}_1(x) - \tilde{g}_1(y) - \tilde{g}_1(z) \\
&\quad = \frac{1}{3}\{g_2(x, y)g_2(x, z) + g_2(x, y)g_2(y, z) + g_2(x, z)g_2(y, z) \\
&\qquad - E[g_2(x, X_3)g_2(y, X_3) - g_2(x, X_3)g_2(z, X_3) - g_2(y, X_3)g_2(z, X_3)]\} \\
&\qquad + \frac{2}{3}\{g_1(x)g_2(y, z) + g_1(y)g_2(x, z) + g_1(z)g_2(x, y)\} \\
&\quad = \tilde{g}_3(x, y, z) \qquad (say).
\end{aligned}$$

Thus using the H-decomposition, we can show that

$$\begin{aligned}
\text{(A.5)} \qquad \hat{\lambda}_1 = \xi_1^2 + \theta^2 &+ \frac{3}{n}\sum_{i=1}^{n}\tilde{g}_1(X_i) + \frac{6}{n(n-1)}\sum_{C_{n,2}}\tilde{g}_2(X_i, X_j) \\
&+ \frac{6}{n(n-1)(n-2)}\sum_{C_{n,2}}\tilde{g}_3(X_i, X_j, X_k).
\end{aligned}$$

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

$$\begin{aligned}
E[\zeta_0(x, y, z, X_4)] = &\frac{1}{3}\{g_1(x)g_2(y, z) + g_1(y)g_2(x, z) + g_1(z)g_2(x, y) \\
&\quad + \theta g_2(x, y) + \theta g_2(x, z) + \theta g_2(y, z) \\
&\quad + 2g_1(x)g_1(y) + 2g_1(x)g_1(z) + 2g_1(y)g_1(z)\} \\
&+ \theta g_1(x) + \theta g_1(y) + \theta g_1(z) + \theta^2, \\
E[\zeta_0(x, y, X_3, X_4)]& \\
= \frac{1}{3}\{2g_1(x)g_1(y) &+ \theta g_2(x, y)\} + \theta g_1(x) + \theta g_1(y) + \theta^2
\end{aligned}$$

and

$$E[\zeta_0(x, X_2, X_3, X_4)] = \theta\{g_1(x) + \theta\}.$$

Thus from the H-decomposition and the equation (A.4), we have

$$\begin{aligned}
\hat{\theta^2} = \theta^2 &+ \frac{4}{n}\sum_{i=1}^{n}\theta g_1(X_i) \\
&+ \frac{4}{n(n-1)}\sum_{C_{n,2}}\{g_1(X_i)g_1(X_j) + \theta g_2(X_i, X_j)\} \\
&+ \frac{8}{n(n-1)(n-2)}\sum_{C_{n,3}}\{g_1(X_i)g_2(X_j, X_k) + g_1(X_j)g_2(X_i, X_k) \\
&\qquad + g_1(X_k)g_2(X_i, X_j)\} + o_p^*(n^{-\frac{3}{2}}).
\end{aligned} \tag{A.6}$$

Combining the equations (A.5) and (A.6), we have the approximation of $\hat{a_1^2}$ as

$$\frac{4(n-2)}{n-1}\hat{a}_1^2 = 4\xi_1^2 - \frac{4\xi_1^2}{n} - \frac{4}{n^2}\sum_{i=1}^{n} f_1(X_i) + V_n + o_p^*(n^{-\frac{3}{2}}). \tag{A.7}$$

Since

$$E[h^2(X_1, X_2)] = 2\xi_1^2 + \xi_2^2 + \theta^2,$$

using the H-decomposition and the equation (13), we obtain

$$n^{-1}\hat{\lambda}_2 = \frac{2\xi_1^2 + \xi_2^2 + \theta^2}{n} + \frac{2}{n^2}\sum_{i=1}^{n}\{\delta(X_i) + f_1(X_i) + 2\theta g_1(X_i)\} + o_p^*(n^{-\frac{3}{2}}).$$

From the H-decomposition and the equations (A.3) and (A.4), we can show that

$$\frac{2}{n-1}\hat{a_2^2} = \frac{4}{n^2}\sum_{i=1}^{n}\{\delta(X_i) + f_1(X_i)\} + \frac{4\xi_1^2 + 2\xi_2^2}{n} + o_p^*(n^{-\frac{3}{2}}). \tag{A.8}$$

Combining the above evaluations (A.7) and (A.8), we have the desired approximation (3.5).

Acknowledgements

The author wishes to thank the referee for 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

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

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

[3] Efron, B. (1987). Better bootstrap confidence intervals (with discussion). *Jour. Amer. Statist. Assoc.*, **82**, 171–200.

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

[5] Hinkley, D. V. (1978). Improving the jackknife with special reference to correlation estimation. *Biometrika*, **65**, 13–1.

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

[7] Maesono, Y. (1995a). Comparisons of variance estimators and their effects for studentized U-statistics. *The Australian National University Statistics Research Report*, No. SRR 40–95.

[8] Maesono, Y. (1995b). On the normal approximation of a studentized U-statistic. *Jour. Jap. Statist. Soc.*, **25**, 19–33.

[9] Maesono, Y. (1996a). Edgeworth expansions of a studentized and a jackknife estimator of variance. to appear in *Jour. Statist. Plann. Inf.*

[10] Maesono, Y. (1996b). An Edgeworth expansion of a linear combination of U-statistics. *Jour. Jap. Statist. Soc.*, **26**, 189–207.

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

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

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

[14] Shirahata, S. and Sakamoto, Y. (1992). Estimate of variance of U-statistics. *Commun. Statist.-Theory Meth.*, **21**, 2969–2981.