MEAN SQUARED ERRORS OF BOOTSTRAP VARIANCE ESTIMATORS FOR U-STATISTICS

Mizuno, Masayuki
Graduate School of Mathematics, School of Mathematics Kyushu University

Maesono, Yoshihiko
Faculty of Mathematics, Kyushu University

https://doi.org/10.5109/1434312

出版情報：Bulletin of informatics and cybernetics. 43, pp.67-82, 2011-12. 統計科学研究会
バージョン：
権利関係：

MEAN SQUARED ERRORS OF BOOTSTRAP VARIANCE ESTIMATORS FOR *U*-STATISTICS

by

Masayuki Mizuno and **Yoshihiko Maesono**

Reprinted from the Bulletin of Informatics and Cybernetics
Research Association of Statistical Sciences, Vol.43

FUKUOKA, JAPAN
2011

Bulletin of Informatics and Cybernetics, Vol. 43, 2011

MEAN SQUARED ERRORS OF BOOTSTRAP VARIANCE ESTIMATORS FOR U-STATISTICS

By

Masayuki Mizuno* and **Yoshihiko Maesono†**

Abstract

In this paper, we obtain an asymptotic representation of the bootstrap variance estimator for a class of U-statistics. Using the representation of the estimator, we will obtain a mean squared error of the variance estimator until the order n^{-2}. Also we compare the bootstrap and the jackknife variance estimators, theoretically.

Key Words and Phrases: U-statistics, Bootstrap, Jackknife, Mean squared error, Variance estimator.

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, e.g. Sen's(1960) estimator, jackknife, unbiased, bootstrap etc. Maesono(1998) obtained the asymptotic representations of them and discussed mean squared errors. 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 a corresponding U-statistic computed from a sample of $n-1$ points with X_i left out.

Another competitor is a bootstrap variance estimator. Many papers show that the bootstrap estimator has good properties in most case, and especially, the bootstrap

* Graduate School of Mathematics, School of Mathematics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan.

† Faculty of Mathematics, Kyushu University, Motooka, Fukuoka 819–0395, Japan. tel +81–92–802–4480 maesono@math.kyushu-u.ac.jp

variance estimator is superior to the jackknife for the quantile estimator, like the sample median. The theoretical value of the bootstrap estimator is given by

$$\hat{\sigma}_B^2 = \iint \Big[\binom{n}{r}^{-1} \sum_{C_{n,r}} h(x_{i_1}, \cdots, x_{i_r}) - \iint h(y_1, \cdots, y_r) \prod_{j=1}^{r} d\hat{F}(y_j)\Big]^2 \prod_{i=1}^{n} d\hat{F}(x_i).$$

where $\hat{F}$ is an estimator of the distribution function F. In this paper we consider the case of the empirical distribution

$$\hat{F}(x) = F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x)$$

where $I(\cdot)$ is an indicator function.

Many papers discuss properties of $\hat{\sigma}_B^2$ by simulations, and show that the bootstrap estimator has the good properties. On the other hand, the properties of $\hat{\sigma}_J^2$ have been precisely studied. Efron and Stein(1981) have showed that the jackknife variance estimator has positive bias, and 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, using computer simulation, Schucany and Bankson(1989) discussed biases and mean squared errors of the variance estimators. 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, iterated bootstrap and bootstrap estimators) by computer simulations, and pointed out that the bootstrap variance estimator has smallest mean squared error.

Using the asymptotic representation of the jackknife variance estimator $n\hat{\sigma}_J^2$ with the residual term $O_p(n^{-2})$, Maesono(1998) has obtained the mean squared errors until the order n^{-2}. In this paper we will obtain an asymptotic representation of the bootstrap variance estimator $n\hat{\sigma}_B^2$ with residual term $O_p(n^{-2})$ and discuss a mean squared error.

In Section 2, we will review the Hoeffding decomposition and moment evaluations of the decomposition. In Section 3, we will discuss the asymptotic representations and the mean squared errors of the bootstrap variance estimators, and compare the mean squared errors for the variance and covariance estimation.

2. Hoeffding decomposition and moment evaluation

At first we will obtain a representation of U_n with a sum of forward martingales. This representation is due to Hoeffding(1961) and plays an important role in the asymptotic theory of U-statistics. Under the assumption that $E|h(X_1, \cdots, X_r)| < \infty$, let us define the following notations:

$$a_k(x_1, \cdots, x_k) = E[h(X_1, \cdots, X_r)|X_1 = x_1, \cdots, X_k = x_k] - \theta \quad (1 \le k \le r),$$

$$g_1(x_1) = a_1(x_1), \qquad g_2(x_1, x_2) = a_2(x_1, x_2) - g_1(x_1) - g_1(x_2), \cdots,$$

$$g_r(x_1, \cdots, x_r) = a_r(x_1, \cdots, x_r) - \sum_{j=1}^{r-1} \sum_{C_{r,j}} g_j(x_{i_1}, \cdots, x_{i_j}),$$

$$A_k = \sum_{C_{n,k}} g_k(X_{i_1}, \cdots, X_{i_k}), \qquad \sigma_n^2 = Var(U_n),$$

$$\xi_1^2 = Eg_1^2(X_1), \qquad \xi_2^2 = E[g_2^2(X_1, X_2)].$$

Then we have

$$U_n - E(U_n) = \begin{pmatrix} n \\ r \end{pmatrix}^{-1} \sum_{k=1}^{r} \begin{pmatrix} n-k \\ r-k \end{pmatrix} A_k$$

and

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

Note that $g_k (1 \leq k \leq r)$ are orthogonal in the sense that their covariances are zero, and $\{A_k\}_{n \geq k}$ is a forward martingale for $k = 1, \cdots, r$. This method is called as H-decomposition or $ANOVA$- decomposition, and is familiar in the studies of the analysis of variance, the jackknife inference, and so on. Using the equation (1) we have the variance σ_n^2 of U_n

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

From von Bahr and Esséen(1965) and Dharmadhikari, Fabian and Jogdeo(1968), we have the upper bounds of the absolute moments of A_k as follows. Since we will use moment evaluations for another kernel, we state the Lemma in general form.

LEMMA 2.1. *Let $\rho(x_1, x_2, \cdots, x_k)$ be a real valued function which is symmetric in its arguments. Assume that $E[\rho(X_1, X_2, \cdots, X_k)|X_1, X_2, \cdots, X_{k-1}] = 0$ a.s. Then we have the following inequalities.*
(i) *For $1 \leq p < 2$, if $E|\rho(X_1, \cdots, X_k)|^p < \infty$, there exists a positive constant c_h, which may depend on ρ and F but not on n, such that*

$$E|\sum_{C_{n,k}} \rho(X_{i_1}, \cdots, X_{i_k})|^p \leq c_h n^k. \tag{3}$$

(ii) *For $2 \leq p$, if $E|\rho(X_1, \cdots, X_k)|^p < \infty$, there exists a positive constant c_h', which may depend on ρ and F but not on n, such that*

$$E|\sum_{C_{n,k}} \rho(X_{i_1}, \cdots, X_{i_k})|^p \leq c_h' n^{pk/2}. \tag{4}$$

PROOF. See Maesono(1998).

Hereafter for the sake of simplicity, we will consider the kernel of degree $r = 2$. The generalization to the kernel of arbitrary degree will be obtained with notational complications and tedious calculations. We consider the bootstrap variance estimator $\hat{\sigma}_B^2$ for the U-statistic

$$\hat{\sigma}_B^2 = \iint \Big[\binom{n}{2}^{-1} \sum_{C_{n,2}} h(x_i, x_j) - \iint h(y_1, y_2) dF_n(y_1) dF_n(y_2) \Big]^2 \prod_{i=1}^{n} dF_n(x_k).$$

For the kernel $h(x, y)$, we have

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

and

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

Note that

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

Thus 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

$$E[g_k(X_{i_1}, X_{i_2}) \nu(X_{j_1}, \cdots, X_{j_m})] = 0. \tag{5}$$

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 of the variance estimators, we will study the estimation of $n\sigma_n^2$. Thus we consider the bootstrap variance estimator $V_B = n\hat{\sigma}_B^2$ and the jackknife estimator $V_J = n\hat{\sigma}_J^2$.

3. Asymptotic representations and mean squared errors

For the variance estimators V_J, Maesono(1998) has obtained the asymptotic representations with residual terms $O_p(n^{-2})$. Let us define

$$\begin{aligned} h^*(x,y) &= h(x,y)-\theta \\ \delta(x) &= E[g_2^2(x,X_2)]-\xi_2^2, \end{aligned}$$

$$\begin{aligned} \delta^*(x) &= -16E[h^*(X_2,X_2)h^*(x,X_3)]+8E[h^*(x,x)h^*(x,X_2)] \\ &\quad +8E[h^*(x,X_2)h^*(X_2,X_2)]-16E[h^*(X_1,X_1)h^*(X_1,X_2)], \end{aligned}$$

$$f_1(x) = g_1^2(x)-\xi_1^2+2E[g_1(X_2)g_2(x,X_2)],$$

$$\begin{aligned} f_2(x,y) &= -g_1(x)g_1(y)+g_2(x,y)\{g_1(x)+g_1(y)\} \\ &\quad +E[g_2(x,X_3)g_2(y,X_3)-g_2(x,X_3)g_1(X_3)-g_2(y,X_3)g_1(X_3)], \end{aligned}$$

$$\begin{aligned} f_3(x,y,z) = {}& g_2(x,y)g_2(x,z)+g_2(x,y)g_2(y,z)+g_2(x,z)g_2(y,z) \\ & -E[g_2(x,X_3)g_2(y,X_3)+g_2(y,X_3)g_2(z,X_3)+g_2(x,X_3)g_2(z,X_3)] \\ & -2\{g_1(x)g_2(y,z)+g_1(y)g_2(x,z)+g_1(z)g_2(x,y)\} \end{aligned}$$

and

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

Note that V_n has already decomposed. For the jackknife variance estimator, Maesono (1998) obtained the following asymptotic representation

$$V_J = V_n+\frac{8}{n^2}\sum_{i=1}^n \delta(X_i)+n\sigma_n^2+\frac{1}{n}b_J+R_{1;n} \tag{6}$$

where

$$b_J = 2\xi_2^2 \tag{7}$$

and $E|R_{1;n}|^{2+\varepsilon/2}=O(n^{-4-\varepsilon})$ for some $\varepsilon>0$. Similarly to V_J, we have an asymptotic representation of V_B as follows.

Theorem 3.1. *If $E|h(X_i,X_j)|^{4+\varepsilon}<\infty$ $(i,j=1,2)$ for some $\varepsilon>0$, we have*

$$V_B = V_n+\frac{1}{n^2}\sum_{i=1}^n\{12\delta(X_i)-20f_1(X_i)+\delta^*(X_i)\}+n\sigma_n^2+\frac{1}{n}b_B+R_{2;n} \tag{8}$$

where

$$b_B = 4\xi_2^2-20\xi_1^2+8E[h^*(X_1,X_1)h^*(X_1,X_2)]$$

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

Proof. See appendix.

b_J and b_B are n^{-1} biases of the jackknife and the bootstrap variance estimators, respectively. It is easy to see that $E[f_1(X_1)] = E[\delta(X_1)] = E[\delta^*(X_1)] = 0$ and

$$E[f_2(X_1, X_2)|X_1] = E[f_3(X_1, X_2, X_3)|X_1, X_2] = 0 \ \ a.s. \tag{9}$$

Using the asymptotic representations of Theorem 3.1, we can study the asymptotic properties of the variance estimators. Here we will obtain mean squared errors of V_J and V_B up to the order n^{-2}. Let us define

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

and

$$\begin{aligned} mse(V_B) &= \frac{16}{n} E[f_1^2(X_1)] + \frac{1}{n^2}\{b_B^2 + 96E[f_1(X_1)\delta(X_1)] \\ &-160E[f_1^2(X_1)] + 8E[f_1(X_1)\delta^*(X_1)] + 32E[f_2^2(X_1, X_2)]\}. \end{aligned}$$

Then Maesono(1998) showed that

$$E(V_J - n\sigma_n^2)^2 = mse(V_J) + O(n^{-5/2}). \tag{10}$$

For the bootstrap estimator V_B, we have the following theorem.

Theorem 3.2. *If* $E|h(X_i, X_j)|^{4+\varepsilon} < \infty$ $(i, j = 1, 2)$ *for some* $\varepsilon > 0$, *we have*

$$E(V_B - n\sigma_n^2)^2 = mse(V_B) + O(n^{-5/2}). \tag{11}$$

Proof. It follows from Theorem 3.1 that under the moment condition, for $1 \le k \le 5$,

$$\begin{aligned} E|n^{-1}R_{2;n}\sum_{i=1}^{n} f_1(X_i)| &\le n^{-1}\{E|\sum_{i=1}^{n} f_1(X_i)|^{2+\varepsilon/2} E|R_{2;n}|^{2+\varepsilon/2}\}^{2/(4+\varepsilon)} \\ &= O(n^{-5/2}), \\ E|n^{-2}R_{2;n}\sum_{i=1}^{n} \delta(X_i)| &\le n^{-2}\{E|\sum_{i=1}^{n} \delta(X_i)|^{2+\varepsilon/2} E|R_{2;n}|^{2+\varepsilon/2}\}^{2/(4+\varepsilon)} \\ &= O(n^{-3}), \\ E|n^{-2}R_{2;n}\sum_{C_{n,2}} f_2(X_i, X_j)| &\le n^{-2}\{E|\sum_{C_{n,2}} f_2(X_i, X_j)|^{2+\varepsilon/2} E|R_{2;n}|^{2+\varepsilon/2}\}^{2/(4+\varepsilon)} \\ &= O(n^{-3}) \end{aligned}$$

and

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

Thus, using these equations and (9), we can obtain the desired result.

REMARK. 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 equation as it is.

Let us define

$$\begin{aligned} e_1 &= E[g_1^4(X_1)], \qquad e_2 = E[g_1^2(X_1)g_2^2(X_1, X_2)], \\ e_3 &= E[g_1(X_1)g_2(X_1, X_2)g_2^2(X_2, X_3)], \\ e_4 &= E[g_1^2(X_1)g_1(X_2)g_2(X_1, X_2)], \\ e_5 &= E[g_1(X_1)g_1(X_2)g_2^2(X_1, X_2)], \\ e_6 &= E[g_1(X_1)g_1(X_2)g_2(X_1, X_3)g_2(X_2, X_3)], \\ e_7 &= E[g_1(X_1)g_2(X_1, X_2)g_2(X_1, X_3)g_2(X_2, X_3)] \end{aligned}$$

and

$$e_8 = E[g_2(X_1, X_3)g_2(X_2, X_3)g_2(X_1, X_4)g_2(X_2, X_4)].$$

Then, using the equation (1), it follows from direct computations that

$$\begin{aligned} E[f_1^2(X_1)] &= e_1 - \xi_1^4 + 4e_4 + 4e_6, \\ E[f_1(X_1)\delta(X_1)] &= e_2 - \xi_1^2\xi_2^2 + 2e_3 \end{aligned}$$

and

$$E[f_2^2(X_1, X_2)] = \xi_1^4 + 2e_2 - 4e_4 + 2e_5 - 4e_6 + 4e_7 + e_8.$$

Here we will study the mean squared errors for the variance and the covariance estimation problems.

EXAMPLE 3.3. *Variance estimation;*
Let us consider the kernel $h(x, y) = (x - y)^2/2$. Then if $E(X_1) = \mu$ and $Var(X_1) = \sigma^2$ exist, 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. Note that $h(x_1, x_1) = 0$ and then $\delta^*(x) = 0$. It is easy to see that

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

For the sake of simplicity, we will consider the case that the distribution $F(x)$ is symmetric about the mean μ. Let us define $m_k = E\{(X_1 - \mu)^k\}$. Then because of symmetry of F, if k is odd number, $m_k = 0$. Using this fact, it follows from direct computations that

$$\begin{aligned} &\xi_1^2 = \frac{1}{4}(m_4 - \sigma^4), \quad \xi_2^2 = \sigma^4, \quad e_1 = \frac{1}{16}(m_8 - 4\sigma^2 m_6 + 6\sigma^4 m_4 - 3\sigma^8), \\ &e_2 = \frac{\sigma^2}{4}(m_6 - 2\sigma^2 m_4 + \sigma^6), \quad e_5 = \frac{1}{4}(m_4 - \sigma^4)^2, \quad e_7 = -\frac{\sigma^4}{2}(m_4 - \sigma^4), \\ &e_8 = \sigma^8 \quad \text{and} \quad e_3 = e_4 = e_6 = 0. \end{aligned}$$

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

$$b_J = 2\sigma^4 \quad \text{and} \quad b_B = -6\sigma^4.$$

Then we have the mean squared errors

$$mse(V_J) = \sigma^8\Big\{\frac{56}{n} + \frac{268}{n^2}\Big\} \quad \text{and} \quad mse(V_B) = \sigma^8\Big\{\frac{56}{n} - \frac{196}{n^2}\Big\}$$

In the case of $\sigma^2 = 1$ and $n = 10$, Schucany and Bankson(1989) discussed the mean squared error of V_J/n, by simulation. Corresponding asymptotic mean squared error is given by

$$\frac{mse(V_J)}{10^2} = 0.0828.$$

Their estimated mean squared error is close to this value.
[*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 $Var(X_1) = \sigma^2$,

$$b_J = 2\sigma^4, \qquad b_B = -8\sigma^4,$$
$$mse(V_J) = \sigma^8\{\frac{538.33}{n} + \frac{1002.95}{n^2}\}$$

and

$$mse(V_B) = \sigma^8\{\frac{538.33}{n} - \frac{4086.31}{n^2}\}.$$

Example 3.4. *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

$$\theta = \rho\sigma_y\sigma_z, \quad g_1(\boldsymbol{x}_1) = \frac{1}{2}\{(y_1 - \mu_y)(z_1 - \mu_z) - \rho\sigma_y\sigma_z\}$$

and

$$g_2(\boldsymbol{x}_1, \boldsymbol{x}_2) = -\frac{1}{2}\{(y_1 - \mu_y)(z_2 - \mu_z) + (z_1 - \mu_z)(y_2 - \mu_y)\}$$

where $\mu_y = E(Y_1)$ and $\mu_z = E(Z_1)$. Note that $h(\boldsymbol{x}_1, \boldsymbol{x}_2) = 0$. Here we assume that $\boldsymbol{X}_i$ is the bivariate normal distribution

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

From direct computations we can get

$$\xi_1^2 = \frac{1+\rho^2}{4}\sigma_y^2\sigma_z^2, \quad \xi_2^2 = \frac{1+\rho^2}{2}\sigma_y^2\sigma_z^2, \quad e_1 = \frac{3}{16}(3\rho^4 + 14\rho^2 + 3)\sigma_y^4\sigma_z^4,$$
$$e_2 = \frac{1}{8}(3\rho^4 + 14\rho^2 + 3)\sigma_y^4\sigma_z^4, \quad e_3 = e_4 = 0, \quad e_5 = \frac{1}{8}(\rho^4 + 6\rho^2 + 1)\sigma_y^4\sigma_z^4,$$
$$e_6 = 0, \quad e_7 = -\frac{1}{8}(\rho^4 + 6\rho^2 + 1)\sigma_y^4\sigma_z^4 \quad \text{and} \quad e_8 = \frac{1}{8}(\rho^4 + 6\rho^2 + 1)\sigma_y^4\sigma_z^4.$$

Thus we have

$$b_J = \sigma_y^2\sigma_z^2(1+\rho^2), \quad b_B = -3(1+\rho^2)\sigma_y^2\sigma_z^2,$$
$$mse(V_J) = \sigma_y^4\sigma_z^4\{\frac{8}{n}(\rho^4 + 5\rho^2 + 1) + \frac{1}{n^2}(39\rho^4 + 190\rho^2 + 39)\}$$

and

$$mse(V_B) == \sigma_y^4\sigma_z^4\{\frac{8}{n}(\rho^4 + 5\rho^2 + 1) - \frac{1}{n^2}(25\rho^4 + 146\rho^2 + 25)\}.$$

Remark. In the cases of the above tow examples, $mse(V_B) < mse(V_J)$. We had better to study the properties of both estimators more precisely.

4. Appendix

Proof of Theorem 1 From the definition, we have

$$\begin{aligned}\hat{\sigma}_B^2 &= \iint \Bigl[\binom{n}{2}^{-1}\sum_{C_{n,2}} h(x_i,x_j) - \iint h(y_1,y_2)dF_n(y_1)dF_n(y_2)\Bigr]^2 \prod_{k=1}^n dF_n(x_k) \\ &= \iint \binom{n}{2}^{-2}\Bigl[\sum_{C_{n,2}} h^*(x_i,x_j)\Bigr]^2 \prod_{k=1}^n dF_n(x_k) \\ &\qquad - \Bigl[\iint h^*(y_1,y_2)dF_n(y_1)dF_n(y_2)\Bigr]^2. \end{aligned} \tag{12}$$

For the first term of (12), it is easy to get the following equation

$$\begin{aligned} &\iint \binom{n}{2}^{-2}\Bigl[\sum_{C_{n,2}} h^*(x_i,x_j)\Bigr]^2 \prod_{k=1}^n dF_n(x_k) \\ &= \frac{1}{n^2(n-1)^2}\int\int\Bigl(\sum_{i\neq j} h^*(x_i,x_j)\Bigr)^2 \prod_{k=1}^n dF_n(x_k) \\ &= \frac{1}{n^2(n-1)^2}\int\int\sum_{i\neq j}\sum_{\ell\neq m} h^*(x_i,x_j)h^*(x_\ell,x_m)\prod_{k=1}^n dF_n(x_k). \end{aligned}$$

Let us define

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

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

and $\zeta_2(x_1,x_2)=\{h^*(x_1,x_2)\}^2$. Then for the first term of (12), we have

$$\begin{aligned}&\iint \binom{n}{2}^{-2}\Big[\sum_{C_{n,2}} h^*(x_i,x_j)\Big]^2\prod_{k=1}^{n} dF_n(x_k)\\
=\ &\frac{1}{n^2(n-1)^2}\int\cdots\int\Big(2n(n-1)\zeta_2(x_1,x_2)+4n(n-1)(n-2)\zeta_1(x_1,x_2,x_3)\\
&\qquad +n(n-1)(n-2)(n-3)\zeta_0(x_1,x_2,x_3,x_4)\Big)\prod_{k=1}^{4} dF_n(x_k)\\
=\ &\frac{2}{n^3(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\zeta_2(X_i,X_j)+\frac{4(n-2)}{n^4(n-1)}\sum_{i=1}^{n}\cdots\sum_{k=1}^{n}\zeta_1(X_i,X_j,X_k)\\
&+\frac{(n-2)(n-3)}{n^5(n-1)}\sum_{i=1}^{n}\cdots\sum_{\ell=1}^{n}\zeta_0(X_i,X_j,X_k,X_\ell).\end{aligned}$$

Furthermore, for each ζ_k we can show that

$$\begin{aligned}&\sum_{i=1}^{n}\sum_{j=1}^{n}\zeta_2(X_i,X_j)\\
=\ &\sum_{i=1}^{n}\{h^*(X_i,X_i)\}^2+2\sum_{C_{n,2}}\zeta_2(X_i,X_j),\end{aligned}$$

$$\begin{aligned}&\sum_{i=1}^{n}\cdots\sum_{k=1}^{n}\zeta_1(X_i,X_j,X_k)\\
=\ &\sum_{i=1}^{n}\{h^*(X_i,X_i)\}^2+2\sum_{C_{n,2}}\zeta_2(X_i,X_j)\\
&+2\sum_{C_{n,2}}\{h^*(X_i,X_i)h^*(X_i,X_j)+h^*(X_i,X_j)h^*(X_j,X_j)\}\\
&+6\sum_{C_{n,3}}\zeta_1(X_i,X_j,X_k)\end{aligned}$$

and

$$\begin{aligned}
&\sum_{i=1}^{n}\cdots\sum_{l=1}^{n}\zeta_0(X_i,X_j,X_k,X_l)\\
= &\sum_{i=1}^{n}\{h^*(X_i,X_i)\}^2+4\sum_{C_{n,2}}\zeta_2(X_i,X_j)\\
&+4\sum_{C_{n,2}}\{h^*(X_i,X_i)h^*(X_i,X_j)+h^*(X_i,X_j)h^*(X_j,X_j)\}\\
&+2\sum_{C_{n,2}}h^*(X_i,X_i)h^*(X_j,X_j)+24\sum_{C_{n,3}}\zeta_1(X_i,X_j,X_k)\\
&+4\sum_{C_{n,3}}\{h^*(X_i,X_i)h^*(X_j,X_k)+h^*(X_j,X_j)h^*(X_i,X_k)+h^*(X_k,X_k)h^*(X_i,X_j)\}\\
&+24\sum_{C_{n,4}}\zeta_0(X_i,X_j,X_k,X_\ell).
\end{aligned}$$

Substituting these terms, we get

$$\begin{aligned}
&\iint\binom{n}{2}^{-2}\left[\sum_{C_{n,2}}h^*(x_i,x_j)\right]^2\prod_{k=1}^{n}dF_n(x_k)\\
= &\frac{7n-6}{n^5}\sum_{i=1}^{n}\{h^*(X_i,X_i)\}^2+\frac{4(4n^2-9n+6)}{n^5(n-1)}\sum_{C_{n,2}}\zeta_2(X_i,X_j)\\
&+\frac{24(n-2)}{n^5}\sum_{C_{n,2}}\frac{1}{2}\{h^*(X_i,X_i)h^*(X_i,X_j)+h^*(X_i,X_j)h^*(X_j,X_j)\}\\
&+\frac{2(n-2)(n-3)}{n^5(n-1)}\sum_{C_{n,2}}h^*(X_i,X_i)h^*(X_j,X_j)\\
&+\frac{24(n-2)(2n-3)}{n^5(n-1)}\sum_{C_{n,3}}\zeta_1(X_i,X_j,X_k)\\
&+\frac{12(n-2)(n-3)}{n^5(n-1)}\sum_{C_{n,3}}\frac{1}{3}\{h^*(X_i,X_i)h^*(X_j,X_k)+h^*(X_j,X_j)h^*(X_i,X_k)\\
&\qquad\qquad+h^*(X_k,X_k)h^*(X_i,X_j)\}\\
&+\frac{24(n-2)(n-3)}{n^5(n-1)}\sum_{C_{n,4}}\zeta_0(X_i,X_j,X_k,X_\ell).
\end{aligned}$$

For the second term of (12), we have

$$\left[\iint h^*(y_1,y_2)dF_n(y_1)dF_n(y_2)\right]^2=\frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{\ell=1}^{n}h^*(X_i,X_j)h^*(X_k,X_\ell).$$

Thus we can show that

$$
\begin{aligned}
&\left[\iint h^*(y_1,y_2)dF_n(y_1)dF_n(y_2)\right]^2 \\
=\ & \frac{1}{n^4}\Bigg[\sum_{i=1}^n \{h^*(X_i,X_i)\}^2 + 4\sum_{C_{n,2}} \zeta_2(X_i,X_j) \\
&+8\sum_{C_{n,2}} \frac{1}{2}\{h^*(X_i,X_i)h^*(X_i,X_j) + h^*(X_i,X_j)h^*(X_j,X_j)\} \\
&+2\sum_{C_{n,2}} h^*(X_i,X_i)h^*(X_j,X_j) \\
&+24\sum_{C_{n,3}} \zeta_1(X_i,X_j,X_k) \\
&+12\sum_{C_{n,3}} \frac{1}{3}\{h^*(X_i,X_i)h^*(X_j,X_k) + h^*(X_j,X_j)h^*(X_i,X_k) \\
&\qquad\qquad +h^*(X_k,X_k)h^*(X_i,X_j)\} \\
&+24\sum_{C_{n,4}} \zeta_0(X_i,X_j,X_k,X_\ell)\Bigg].
\end{aligned}
$$

From the above calculations, we get

$$
\begin{aligned}
\hat{\sigma}_B^2 &= \iint \binom{n}{2}^{-2}\left[\sum_{C_{n,2}} h^*(x_i,x_j)\right]^2 \prod_{k=1}^n dF_n(x_k) - \left[\iint h^*(y_1,y_2)dF_n(y_1)dF_n(y_2)\right]^2 \\
&= \frac{6(n-1)}{n^5}\sum_{i=1}^n \{h^*(X_i,X_i)\}^2 + \frac{4(3n^2-8n+6)}{n^5(n-1)}\sum_{C_{n,2}} \zeta_2(X_i,X_j) \\
&\quad +\frac{16(n-3)}{n^5}\sum_{C_{n,2}} \frac{1}{2}\{h^*(X_i,X_i)h^*(X_i,X_j) + h^*(X_i,X_j)h^*(X_j,X_j)\} \\
&\quad -\frac{4(2n-3)}{n^5}\sum_{C_{n,2}} h^*(X_i,X_i)h^*(X_j,X_j) + \frac{24(n^2-6n+6)}{n^5(n-1)}\sum_{C_{n,3}} \zeta_1(X_i,X_j,X_k) \\
&\quad -\frac{24(2n-3)}{n^5(n-1)}\sum_{C_{n,3}} \frac{1}{3}\{h^*(X_i,X_i)h^*(X_j,X_k) + h^*(X_j,X_j)h^*(X_i,X_k) \\
&\qquad\qquad +h^*(X_k,X_k)h^*(X_i,X_j)\} \\
&\quad -\frac{48(2n-3)}{n^5(n-1)}\sum_{C_{n,4}} \zeta_0(X_i,X_j,X_k,X_\ell).
\end{aligned}
$$

Using the H-decomposition, we will obtain asymptotic representation of $V_B = n\hat{\sigma}_B^2$

where

$$\begin{aligned}
V_B &= \frac{6(n-1)}{n^4}\sum_{i=1}^{n}\{h^*(X_i,X_i)\}^2 + \frac{4(3n^2-8n+6)}{n^4(n-1)}\sum_{C_{n,2}}\zeta_2(X_i,X_j) \\
&+\frac{16(n-3)}{n^4}\sum_{C_{n,2}}\frac{1}{2}\{h^*(X_i,X_i)h^*(X_i,X_j)+h^*(X_i,X_j)h^*(X_j,X_j)\} \\
&-\frac{4(2n-3)}{n^4}\sum_{C_{n,2}}h^*(X_i,X_i)h^*(X_j,X_j)+\frac{24(n^2-6n+6)}{n^4(n-1)}\sum_{C_{n,3}}\zeta_1(X_i,X_j,X_k) \\
&-\frac{24(2n-3)}{n^4(n-1)}\sum_{C_{n,3}}\frac{1}{3}\{h^*(X_i,X_i)h^*(X_j,X_k)+h^*(X_j,X_j)h^*(X_i,X_k) \\
&\qquad +h^*(X_k,X_k)h^*(X_i,X_j)\} \\
&-\frac{48(2n-3)}{n^4(n-1)}\sum_{C_{n,4}}\zeta_0(X_i,X_j,X_k,X_\ell).
\end{aligned}$$

For $x \neq y$, we have

$$h^*(x,y) = g_2(x,y) + g_1(x) + g_1(y).$$

Let us obtain the H-decomposition for ζ_2. It is easy to see that

$$\begin{aligned}
& E[\zeta_2(x,X_2)] \\
=\ & E[g_2^2(x,X_2)+2g_2(x,X_2)g_1(X_2)] + g_1^2(x) + \xi_1^2
\end{aligned}$$

and

$$\begin{aligned}
& E[\zeta_2(x,y)] \\
=\ & g_2^2(x,y) + 2g_2(x,y)\{g_1(x)+g_1(x)\} + g_1^2(x) + g_1^2(y) + 2g_1(x)g_1(y).
\end{aligned}$$

Since $E[\zeta_2(X_1,X_2)] = 2\xi_1^2 + \xi_2^2$, we have the decompositions $\tilde{g}_1$ and $\tilde{g}_2$ of ζ_2 where

$$\begin{aligned}
\tilde{g}_1(x) &= E[\zeta_2(x,X_2)] - 2\xi_1^2 - \xi_2^2 \\
&= g_1^2(x) - 2\xi_1^2 - \xi_2^2 + E[g_2^2(x,X_2)+2g_2(x,X_2)g_1(X_2)]
\end{aligned}$$

and

$$\begin{aligned}
\tilde{g}_2(x,y) &= E[\zeta_2(x,y)] - 2\xi_1^2 - \xi_2^2 - \tilde{g}_1(x) - \tilde{g}_1(y) \\
&= g_2^2(x,y) + 2g_2(x,y)\{g_1(x)+g_1(x)\} + 2g_1(x)g_1(y) + \xi_2^2 \\
&\quad -E[g_2^2(x,X_2)+g_2^2(y,X_2)+\{2g_2(x,X_2)+2g_2(y,X_2)\}g_1(X_2)].
\end{aligned}$$

Thus we can obtain the H-decomposition

$$\begin{aligned}
& \binom{n}{2}^{-1}\sum_{C_{n,2}}\zeta_2(X_i,X_j) \\
=\ & E\left[\binom{n}{2}^{-1}\sum_{k=1}^{2}\zeta_2(X_i,X_j)\right] + \binom{n}{2}^{-1}\sum_{C_{n,2}}\binom{n-k}{2-k}^2 A_k \\
=\ & 2\xi_1^2 + \xi_2^2 + \frac{2}{n}\sum_{i=1}^{n}\tilde{g}_1(X_i) + \frac{2}{n(n-1)}\sum_{C_{n,2}}\tilde{g}_2(X_i,X_j) \qquad (13)
\end{aligned}$$

Similarly we can obtain H-decomposition of ζ_1, ζ_0 as follows. It is easy to see that

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

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

and

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

Since $E[\zeta_1(X_1, X_2, X_3)] = \xi_1^2$, we have the decompositions $\hat{g}_1, \hat{g}_2$ and $\hat{g}_3$ of ζ_1 where

$$\begin{aligned} \hat{g}_1(x) &= E[\zeta_1(x, X_2, X_3)] - \xi_1^2 \\ &= \frac{1}{3}\{g_1^2(x) - \xi_1^2\} + \frac{2}{3}E[g_2(x, X_2)g_1(X_2)], \end{aligned}$$

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

and

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

Thus we get the H-decomposition

$$\begin{aligned} \binom{n}{3}^{-1} \sum_{C_{n,3}} \zeta_1(X_i, X_j, X_k) &= \xi_1^2 + \frac{3}{n}\sum_{i=1}^{n} \hat{g}_1(X_i) + \frac{6}{n(n-1)} \sum_{C_{n,2}} \hat{g}_2(X_i, X_j) \\ &\quad + \frac{6}{n(n-1)(n-2)} \sum_{C_{n,3}} \hat{g}_3(X_i, X_j, X_k). \end{aligned} \tag{14}$$

Furthermore, we can get the decompositions $\breve{g}_1, \breve{g}_2, \breve{g}_3$ and $\breve{g}_4$ of ζ_0 as follows.

$$\begin{aligned}
E[\zeta_0(X_1, X_2, X_3, X_4)] &= E[\zeta_0(x, X_2, X_3, X_4)] = \breve{g}_1(x) = 0, \\
E[\zeta_0(x, y, X_3, X_4)] &= \frac{2}{3} g_1(x) g_1(y), \\
E[\zeta_0(x, y, z, X_4)] &= \frac{1}{3}[g_2(x, y)g_1(z) + g_2(x, z)g_1(y) + g_2(y, z)g_1(x) \\
&\quad +2\{g_1(x)g_1(y) + g_1(x)g_1(z)g_1(y)g_1(z)\}], \\
\breve{g}_2(x, y) &= \frac{2}{3} g_1(x) g_1(y), \\
\breve{g}_3(x, y, z) &= \frac{1}{3}\{g_2(x, y)g_1(z) + g_2(x, z)g_1(y) + g_2(y, z)g_1(x)\}
\end{aligned}$$

and

$$\breve{g}_4(x, y, z, w) = \frac{1}{3}\{g_2(x, y)g_2(z, w) + g_2(x, z)g_2(y, w) + g_2(x, w)g_2(y, z)\}.$$

Thus we have the H-decomposition for ζ_0

$$\begin{aligned}
&\binom{n}{4}^{-1} \sum_{C_{n,4}} \zeta_0(X_i, X_j, X_k, X_l) \\
= &\ \frac{12}{n(n-1)} \sum_{C_{n,2}} \breve{g}_2(X_i, X_j) + \frac{24}{n(n-1)(n-2)} \sum_{C_{n,3}} \breve{g}_3(X_i, X_j, X_k) \\
&+ \frac{24}{n(n-1)(n-2)(n-3)} \sum_{C_{n,4}} \breve{g}_4(X_i, X_j, X_k, X_l). \qquad (15)
\end{aligned}$$

Combining the equations (13), (14), (15) and moment evaluations, we get the desired result (8).

Acknowledgement

This research was supported by JSPS Grant-in-Aid for Scientific Research (B) No.21340026 and Exploratory Research No.21650065.

References

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). he jackknife estimate of variance. *Ann. Statist.*, **9**, 586-596.

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

Hoeffding, W. (1961). he strong law of large numbers for U-statistics, *Univ. of North Carolina Institute of statistics. Mimeo Series.* No.**302**.

Maesono, Y. (1998). Mean square errors of variance estimators and their Edgeworth expansions. *Journal of the Japan Statistical Society*, **28**, 1–19.

Petrov, V.V. (1975). *Sums of Independent Random Variables*, Springer Berlin.

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). n some convergence properties of U-statistics. *Calcutta Statist. Assoc. Bull.*, **10**, 1-18.

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

von Bahr, B and Esséen C.G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.*, **36**, 299303.

Received August 25, 2011
Revised November 20, 2011