J. Japan Statist. Soc.
Vol. 36 No. 1 2006 73–89

A CONVEX COMBINATION OF TWO-SAMPLE U-STATISTICS

Koichiro Toda* and Hajime Yamato**

A convex combination of one-sample U-statistics was introduced by Toda and Yamato (2001) and its Edgeworth expansion was derived by Yamato *et al.* (2003). We introduce a convex combination of two-sample U-statistics, which includes two-sample U-statistic, V-statistic and limit of Bayes estimate. Its Edgeworth expansion is derived with remainder term $o(N^{-1/2})$, under the condition that the kernel is non-degenerate. We give some examples of the expansion for three statistics, two-sample U-statistic, V-statistic and limit of Bayes estimate, based on some distributions.

Key words and phrases: Convex combination, two-sample U-statistic, two-sample V-statistic.

1. Introduction

Let F and G be continuous distributions on the real line. Let $\theta = \theta(F, G)$ be a regular functional of F and G. That is, there exists a measurable function $h(x_1, \ldots, x_{k_1}; y_1, \ldots, y_{k_2})$ such that

$$\theta = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x_1, \ldots, x_{k_1}; y_1, \ldots, y_{k_2}) \prod_{i=1}^{k_1} dF(x_i) \prod_{j=1}^{k_2} dG(y_j). \tag{1.1}$$

We assume that $h(x_1, \ldots, x_{k_1}; y_1, \ldots, y_{k_2})$ is symmetric with respect to $x_1, \ldots, x_{k_1}$ and $y_1, \ldots, y_{k_2}$, respectively, and the integers k_1 and k_2 are the smallest integers satisfying (1.1). The function h is called the kernel of θ and (k_1, k_2) is called the degree of h and/or θ.

Let $X_1, \ldots, X_{n_1}$ and $Y_1, \ldots, Y_{n_2}$ be two independent samples of sizes n_1 and n_2 from the distributions F and G, respectively. As estimators of θ, two-sample U-statistic and V-statistic are well-known. The two-sample U-statistic U_{n_1,n_2} is given by

$$U_{n_1,n_2} = \binom{n_1}{k_1}^{-1}\binom{n_2}{k_2}^{-1} \sum_{(n_1,k_1)}\sum_{(n_2,k_2)} h(X_{i_1}, \ldots, X_{i_{k_1}}; Y_{j_1}, \ldots, Y_{j_{k_2}}), \tag{1.2}$$

where the summation $\sum_{(n_1,k_1)}$ is taken over all possible $i_1, \ldots, i_{k_1}$ satisfying $1 \le i_1 < \cdots < i_{k_1} \le n_1$. The two-sample V-statistic V_{n_1,n_2} is given by

$$V_{n_1,n_2} = n_1^{-k_1} n_2^{-k_2} \tag{1.3}$$

Received March 4, 2005. Revised June 28, 2005. Accepted November 11, 2005.

*Kagoshima Koto Preparatory School, Kagoshima 890-0051, Japan.
**Department of Mathematics and Computer Science, Kagoshima University, Kagoshima 890-0065, Japan.

$$\times \sum_{i_1=1}^{n_1} \cdots \sum_{i_{k_1}=1}^{n_1} \sum_{j_1=1}^{n_2} \cdots \sum_{j_{k_2}=1}^{n_2} h(X_{i_1}, \ldots, X_{i_{k_1}}; Y_{j_1}, \ldots, Y_{j_{k_2}}).$$

The followings are examples of $\theta(F, G)$: **(i)** The kernel $h(x, y) = x - y$ gives the parameter $\theta = E(X_1) - E(Y_1)$. **(ii)** The kernel $h(x, y) = 1$ $(x \le y)$ and $= 0$ $(x > y)$ gives the parameter $\theta = P(X_1 \le Y_1)$. The corresponding U-statistic is related to the two-sample Wilcoxon (Mann-Whitney) rank sum statistic W by the relation $W = n_1 n_2 U_{n_1,n_2} + n_2(n_2 + 1)/2$. **(iii)** The kernel $h(x_1, x_2; y_1, y_2) = 1/3$ $(x_1, x_2 < y_1, y_2$ or $y_1, y_2 < x_1, x_2)$ and $= -1/6$ (otherwise) gives the parameter $\Delta = \int_{-\infty}^{\infty} [F(x) - G(x)]^2 d([F(x) + G(x)]/2)$, which is regarded as a distance between the two distributions F and G. **(iv)** The kernel $h(x_1, x_2; y_1, y_2) = 1$ $(|y_1 - y_2| > |x_1 - x_2|)$ and $= 0$ (otherwise) gives the parameter $\theta = P(|Y_1 - Y_2| > |X_1 - X_2|)$. The associated U-statistic appears in the testing problem of two-sample scale by Lehmann (1951). **(v)** The kernel $h(x_1, x_2; y_1, y_2) = 1$ $(x_1 + x_2 < y_1 + y_2)$ and $= 0$ (otherwise) gives the parameter $\theta = P(X_1 + X_2 < Y_1 + Y_2)$, which is a measure of the difference in location considered by Hollander (1967). (See, for example, Koroljuk and Borovskich (1994), and Randles and Wolfe (1979).)

Yamato (1977) derives the Bayes estimate of θ using Dirichlet prior process of Ferguson (1973), and gives the limit of Bayes estimate which is given by

$$(1.4) \qquad B_{n_1,n_2} = \binom{n_1 + k_1 - 1}{k_1}^{-1} \binom{n_2 + k_2 - 1}{k_2}^{-1} \sum_{r_1+\cdots+r_{n_1}=k_1} \sum_{s_1+\cdots+s_{n_2}=k_2} h(\underbrace{X_1, \ldots, X_1}_{r_1}, \ldots, \underbrace{X_{n_1}, \ldots, X_{n_1}}_{r_{n_1}}; \underbrace{Y_1, \ldots, Y_1}_{s_1}, \ldots, \underbrace{Y_{n_2}, \ldots, Y_{n_2}}_{s_{n_2}}),$$

where the summation $\sum_{r_1+\cdots+r_{n_1}=k_1}$ is taken over all nonnegative integers $r_1, \ldots, r_{n_1}$ satisfying $r_1 + \cdots + r_{n_1} = k_1$. This statistic is abbreviated to LB-statistic.

All the above two-sample statistics U_{n_1,n_2}, V_{n_1,n_2} and B_{n_1,n_2} can be denoted by a convex combination of two-sample U-statistics. This convex combination Y_{n_1,n_2} is introduced in Section 2. In this paper, we put $N = n_1 + n_2$ and consider the asymptotic properties of statistics under the condition such that

$$(1.5) \qquad \frac{n_1}{N} \to p, \qquad \frac{n_2}{N} \to 1 - p \quad \text{as} \quad N \to \infty$$

where $0 < p < 1$ is a constant. In Section 3, we give an asymptotic expansion of Y_{n_1,n_2} as N tends to ∞.

The two-sample U-statistic U_{n_1,n_2} has asymptotic normality (see, for example, Koroljuk and Borovskich (1994), Lee (1990), and Randles and Wolfe (1979)). From this, it is shown that the two-sample Y-statistic Y_{n_1,n_2} has also the same asymptotic normality. To see the difference between asymptotic distributions of these two statistics, we need their Edgeworth expansion. The Edgeworth expansion of the two-sample U-statistic U_{n_1,n_2} was derived by Koroljuk and Borovskich (1994), and Maesono (1985). We shall derive the Edgeworth expansion of Y_{n_1,n_2} in Section 4.

For the same parameter θ, Edgeworth expansion of Y_{n_1,n_2} depends on the weight function w. In Section 5, for 4 kernels we give examples of the expansion for the statistics V_{n_1,n_2}, S_{n_1,n_2} and B_{n_1,n_2}, based on some special distributions.

2. A convex combination of two-sample U-statistics

Let $w(\alpha_1,\dots,\alpha_j;k)$ be a nonnegative and symmetric function of positive integers $\alpha_1,\dots,\alpha_j$ such that $j=1,\dots,k$ and $\alpha_1+\cdots+\alpha_j=k$ for a given integer k. We assume that at least one of $w(\alpha_1,\dots,\alpha_j;k)$'s is positive. We put

$$(2.1)\qquad d(k,j)=\sum\nolimits^{+}_{\alpha_1+\cdots+\alpha_j=k} w(\alpha_1,\dots,\alpha_j;k),\qquad j=1,2,\dots,k,$$

where the summation $\sum^{+}_{\alpha_1+\cdots+\alpha_j=k}$ is taken over all positive integers $\alpha_1,\dots,\alpha_j$ satisfying $\alpha_1+\cdots+\alpha_j=k$ for j and k given.

For $j_1=1,\dots,k_1$ and $j_2=1,\dots,k_2$, let $h_{(j_1,j_2)}(x_1,\dots,x_{j_1};y_1,\dots,y_{j_2})$ be the kernel given by

$$\begin{aligned}(2.2)\qquad & h_{(j_1,j_2)}(x_1,\dots,x_{j_1};y_1,\dots,y_{j_2})\\ &=\frac{1}{d(k_1,j_1)d(k_2,j_2)}\sum\nolimits^{+}_{r_1+\cdots+r_{j_1}=k_1}\sum\nolimits^{+}_{s_1+\cdots+s_{j_2}=k_2}\\ &\quad w(r_1,\dots,r_{j_1};k_1)w(s_1,\dots,s_{j_2};k_2)\\ &\quad\times h(\underbrace{x_1,\dots,x_1}_{r_1},\dots,\underbrace{x_{j_1},\dots,x_{j_1}}_{r_{j_1}};\underbrace{y_1,\dots,y_1}_{s_1},\dots,\underbrace{y_{j_2},\dots,y_{j_2}}_{s_{j_2}}).\end{aligned}$$

Let $U^{(j_1,j_2)}_{n_1,n_2}$ be the two-sample U-statistic associated with this kernel $h_{(j_1,j_2)}$ for $j_1=1,\dots,k_1$ and $j_2=1,\dots,k_2$.

The kernel $h_{(j_1,j_2)}(x_1,\dots,x_{j_1};y_1,\dots,y_{j_2})$ is symmetric with respect to $x_1,\dots,x_{j_1}$ and $y_1,\dots,y_{j_2}$, respectively, because of the symmetry of $w(\alpha_1,\dots,\alpha_j;k)$. If $d(k_1,j_1)$ or $d(k_2,j_2)$ equal zero for some j_1 or j_2, respectively, then the associated $w(\alpha_1,\dots,\alpha_{j_1};k_1)$'s or $w(\alpha_1,\dots,\alpha_{j_2};k_2)$'s are equal to zero. In this case, we let the corresponding statistic $U^{(j_1,j_2)}_{n_1,n_2}$ to be zero. Especially, if $w(1,\dots,1;k)>0$ then we have

$$h_{(k_1,k_2)}=h\quad\text{and}\quad U^{(k_1,k_2)}_{n_1,n_2}=U_{n_1,n_2},$$

because of $d(k_1,k_1)=w(1,\dots,1;k_1)$ and $d(k_2,k_2)=w(1,\dots,1;k_2)$. If $w(1,\dots,1;k)>0$ and $w(1,\dots,1,2;k)>0$, then

$$\begin{aligned}(2.3)\qquad & h_{(k_1-1,k_2)}(x_1,x_2,\dots,x_{k_1-1};y_1,\dots,y_{k_2})\\ &=\frac{1}{k_1-1}[h(x_1,x_1,x_2,\dots,x_{k_1-1};y_1,\dots,y_{k_2})\\ &\qquad+h(x_1,x_2,x_2,x_3,\dots,x_{k_1-1};y_1,\dots,y_{k_2})\\ &\qquad+\cdots+h(x_1,x_2,\dots,x_{k_1-2},x_{k_1-1},x_{k_1-1};y_1,\dots,y_{k_2})]\end{aligned}$$

and

$$\begin{aligned}(2.4)\qquad & h_{(k_1,k_2-1)}(x_1,\dots,x_{k_1};y_1,\dots,y_{k_2-1}) \\ &= \frac{1}{k_2-1}[h(x_1,\dots,x_{k_1};y_1,y_1,y_2,\dots,y_{k_2-1}) \\ &\quad + h(x_1,\dots,x_{k_1};y_1,y_2,y_2,y_3,\dots,y_{k_2-1}) \\ &\quad + \cdots + h(x_1,\dots,x_{k_1};y_1,\dots,y_{k_2-2},y_{k_2-1},y_{k_2-1})].\end{aligned}$$

Definition 2.1. *As an estimator of θ, a convex combination of two-sample U-statistics is defined by*

$$(2.5)\quad Y_{n_1,n_2} = \frac{1}{D(n_1,k_1)D(n_2,k_2)}\sum_{j_1=1}^{k_1}\sum_{j_2=1}^{k_2} d(k_1,j_1)d(k_2,j_2)\binom{n_1}{j_1}\binom{n_2}{j_2}U_{n_1,n_2}^{(j_1,j_2)},$$

where $D(n,k) = \sum_{j=1}^{k} d(k,j)\binom{n}{j}$.

Since w's are nonnegative and at least one of them is positive, $D(n,k)$ is positive.

If we choose the weight function w, the statistic Y_{n_1,n_2} is determined as an estimator of θ. For example, let w be the function given by

$$w(1,1,\dots,1;k) = 1 \quad \text{and} \quad w(\alpha_1,\dots,\alpha_j;k) = 0 \quad \text{for} \quad j = 1,\dots,k-1.$$

Then we have $d(k,k) = 1$, $d(k,j) = 0$ $(j = 1,\dots,k-1)$ and $D(n,k) = \binom{n}{k}$. Thus Y_{n_1,n_2} is equal to the two-sample U-statistic U_{n_1,n_2}.

Let w be the function given by

$$w(\alpha_1,\dots,\alpha_j;k) = \frac{k!}{\alpha_1!\cdots\alpha_j!}$$

for positive integers $\alpha_1,\dots,\alpha_j$ such that $j = 1,\dots,k$ and $\alpha_1+\cdots+\alpha_j = k$. Then it holds that $\sum_{\alpha_1+\cdots+\alpha_j=k}^{+} w(\alpha_1,\dots,\alpha_j;k) = j!\mathcal{S}(k,j)$, where $\mathcal{S}(k,j)$ is the Stirling number of the second kind. (For the Stirling numbers, see for example, Charalambides and Singh (1988).) Hence, we have $d(k,j) = j!\mathcal{S}(k,j)$ for $j = 1,\dots,k$ and $D(n,k) = \sum_{j=1}^{k}\mathcal{S}(k,j)(n)_j = n^k$, where $(n)_j = n(n-1)\cdots(n-j+1)$. Thus the kernel $h_{(j_1,j_2)}(x_1,\dots,x_{j_1};y_1,\dots,y_{j_2})$ is equal to

$$\begin{aligned}&\frac{1}{j_1!\mathcal{S}(k_1,j_1)j_2!\mathcal{S}(k_2,j_2)}\sum\nolimits_{r_1+\cdots+r_{j_1}=k_1}^{+}\sum\nolimits_{s_1+\cdots+s_{j_2}=k_2}^{+}\frac{k_1!}{r_1!\cdots r_{j_1}!}\frac{k_2!}{s_1!\cdots s_{j_2}!} \\ &\qquad\times h(\underbrace{x_1,\dots,x_1}_{r_1},\dots,\underbrace{x_{j_1},\dots,x_{j_1}}_{r_{j_1}};\underbrace{y_1,\dots,y_1}_{s_1},\dots,\underbrace{y_{j_2},\dots,x_{j_2}}_{s_{j_2}}).\end{aligned}$$

By the U-statistics $U_{n_1,n_2}^{(j_1,j_2)}$ associated with these kernels, the statistic Y_{n_1,n_2} is written as

$$V_{n_1,n_2} = \frac{1}{n_1^{k_1}n_2^{k_2}}\sum_{j_1=1}^{k_1}\sum_{j_2=1}^{k_2}\mathcal{S}(k_1,j_1)\mathcal{S}(k_2,j_2)(n_1)_{j_1}(n_2)_{j_2}U_{n_1,n_2}^{(j_1,j_2)},$$

which is equal to the two-sample V-statistic V_{n_1,n_2} given by (1.3) (see, Toda and Yamato (2001), p. 227–228).

Let w be the function given by

$$w(\alpha_1, \dots, \alpha_j; k) = 1$$

for positive integers $\alpha_1, \dots, \alpha_j$ such that $j = 1, \dots, k$ and $\alpha_1 + \cdots + \alpha_j = k$. Then, we have $d(k,j) = \binom{k-1}{j-1}$ for $j = 1, \dots, k$ and $D(n,k) = \sum_{j=1}^{k} \binom{k-1}{j-1}\binom{n}{j} = \binom{n+k-1}{k}$. Thus the kernel $h_{(j_1,j_2)}(x_1, \dots, x_{j_1}; y_1, \dots, y_{j_2})$ is equal to

$$\binom{k_1-1}{j_1-1}^{-1}\binom{k_2-1}{j_2-1}^{-1} {\sum}^{+}_{r_1+\cdots+r_{j_1}=k_1} {\sum}^{+}_{s_1+\cdots+s_{j_2}=k_2} h(\underbrace{x_1, \dots, x_1}_{r_1}, \dots, \underbrace{x_{j_1}, \dots, x_{j_1}}_{r_{j_1}}; \underbrace{y_1, \dots, y_1}_{s_1}, \dots, \underbrace{y_{j_2}, \dots, x_{j_2}}_{s_{j_2}}).$$

By the U-statistics $U_{n_1,n_2}^{(j_1,j_2)}$ associated with these kernels, the statistic Y_{n_1,n_2} is written as

$$B_{n_1,n_2} = \binom{n_1+k_1-1}{k_1}^{-1}\binom{n_2+k_2-1}{k_2}^{-1} \times \sum_{j_1=1}^{k_1}\sum_{j_2=1}^{k_2}\binom{k_1-1}{j_1-1}\binom{k_2-1}{j_2-1}\binom{n_1}{j_1}\binom{n_2}{j_2} U_{n_1,n_2}^{(j_1,j_2)},$$

which is equal to B_{n_1,n_2} given by (1.4) (see, Toda and Yamato (2001), p. 227).

Let w be the function given by

$$w(\alpha_1, \dots, \alpha_j; k) = \frac{k!}{\alpha_1 \cdots \alpha_j}$$

for positive integers $\alpha_1, \dots, \alpha_j$ such that $j = 1, \dots, k$ and $\alpha_1 + \cdots + \alpha_j = k$. Then it holds that ${\sum}^{+}_{\alpha_1+\cdots+\alpha_j=k} w(\alpha_1, \dots, \alpha_j; k) = j!|\boldsymbol{s}(k,j)|$, where $\boldsymbol{s}(k,j)$ is the Stirling number of the first kind. Hence, we have $d(k,j) = j!|\boldsymbol{s}(k,j)|$ for $j = 1, \dots, k$ and

$$D(n,k) = \sum_{j=1}^{k} |\boldsymbol{s}(k,j)|(n)_j.$$

Thus the kernel $h_{(j_1,j_2)}(x_1, \dots, x_{j_1}; y_1, \dots, y_{j_2})$ is equal to

$$\frac{1}{j_1!j_2!|\boldsymbol{s}(k_1,j_1)\boldsymbol{s}(k_2,j_2)|} {\sum}^{+}_{r_1+\cdots+r_{j_1}=k_1} {\sum}^{+}_{s_1+\cdots+s_{j_2}=k_2} \frac{k_1!}{r_1 \cdots r_{j_1}} \frac{k_2!}{s_1 \cdots s_{j_2}} \times h(\underbrace{x_1, \dots, x_1}_{r_1}, \dots, \underbrace{x_{j_1}, \dots, x_{j_1}}_{r_{j_1}}; \underbrace{y_1, \dots, y_1}_{s_1}, \dots, \underbrace{y_{j_2}, \dots, x_{j_2}}_{s_{j_2}}).$$

By the U-statistics $U_{n_1,n_2}^{(j_1,j_2)}$ associated with these kernels, the statistic Y_{n_1,n_2} is written as

$$S_{n_1,n_2} = \frac{1}{D(n_1,k_1)D(n_2,k_2)} \sum_{j_1=1}^{k_1}\sum_{j_2=1}^{k_2} |\boldsymbol{s}(k_1,j_1)\boldsymbol{s}(k_2,j_2)|(n_1)_{j_1}(n_2)_{j_2} U_{n_1,n_2}^{(j_1,j_2)},$$

(compare with Nomachi *et al.* (2002), p. 97–98).

3. Asymptotic expansion of Y-statistic

If $k_1 = 1$ and $k_2 = 1$, then the statistic Y_{n_1,n_2} is equal to U_{n_1,n_2}. Therefore hereafter we assume that $k_1k_2 \geq 2$. We suppose $d(k,k) > 0$, which is equivalent to $w(1,\dots,1;k) > 0$. Then, with $\delta_k = kd(k,k-1)/d(k,k)$ it holds that

$$\frac{d(k,k)}{D(n,k)}\binom{n}{k} = 1 - \frac{\delta_k}{n} + O\left(\frac{1}{n^2}\right), \tag{3.1}$$

$$\frac{d(k,k-1)}{D(n,k)}\binom{n}{k-1} = \frac{\delta_k}{n} + O\left(\frac{1}{n^2}\right). \tag{3.2}$$

and

$$\frac{d(k,j)}{D(n,k)}\binom{n}{j} = O\left(\frac{1}{n^2}\right), \qquad j = 1,\dots,k-2. \tag{3.3}$$

For the U-statistic U_{n_1,n_2}, $d(k,k)n^{(k)}/[D(n,k)k!] = 1$ and $\delta_k = 0$. For the V-statistic V_{n_1,n_2} and the statistic S_{n_1,n_2}, $\delta_k = k(k-1)/2$. For the statistic B_{n_1,n_2}, $\delta_k = k(k-1)$ (see Nomachi *et al.* (2002)).

We assume that

$$E\Bigg\{h(\underbrace{X_1,\dots,X_1}_{r_1},\dots,\underbrace{X_{j_1},\dots,X_{j_1}}_{r_{j_1}};\underbrace{Y_1,\dots,Y_1}_{s_1},\dots,\underbrace{Y_{j_2},\dots,Y_{j_2}}_{s_{j_2}})\Bigg\}^2 < \infty \tag{3.4}$$

for positive integers $r_1,\dots,r_{j_1}$ and $s_1,\dots,s_{j_2}$ satisfying $r_1 + \cdots + r_{j_1} = k_1$ ($j_1 = 1,\dots,k_1$) and $s_1 + \cdots + s_{j_2} = k_2$ ($j_2 = 1,\dots,k_2$), respectively. Then there exist $E\{U^{(j_1,j_2)}_{n_1,n_2}\}^2$ for $j_1 = 1,\dots,k_1$ and $j_2 = 1,\dots,k_2$. Thus we have for $j_1 = 1,\dots,k_1$ and $j_2 = 1,\dots,k_2$

$$E\{U^{(j_1,j_2)}_{n_1,n_2}\}^2 < \infty, \qquad \text{and} \qquad \mathrm{Var}(U^{(j_1,j_2)}_{n_1,n_2}) = O(N^{-1}). \tag{3.5}$$

Then we have the following asymptotic expansion of the statistic Y_{n_1,n_2}.

PROPOSITION 3.1. *Under the conditions* $w(1,\dots,1;k) > 0$ *and* (3.4), *we have*

$$Y_{n_1,n_2} - \theta = (U_{n_1,n_2} - \theta) + \frac{1}{N}b^{(0)} + R_{n_1,n_2} \tag{3.6}$$

where $E[|R_{n_1,n_2}|^2] = o(N^{-2})$,

$$b^{(0)} = \frac{\delta_{k_1}}{p}(\theta^{(k_1-1,k_2)} - \theta) + \frac{\delta_{k_2}}{1-p}(\theta^{(k_1,k_2-1)} - \theta) \tag{3.7}$$

and

$$\theta^{(k_1-1,k_2)} = E[h(X_1, X_1, X_2, X_3, \dots, X_{k_1-1}; Y_1, Y_2, \dots, Y_{k_2})],$$
$$\theta^{(k_1,k_2-1)} = E\big[h(X_1, X_2, \dots, X_{k_1}; Y_1, Y_1, Y_2, Y_3, \dots, Y_{k_2-1})\big].$$

PROOF. We note

$$E[U^{(k_1,k_2-1)}_{n_1,n_2}] = \theta^{(k_1,k_2-1)} \quad \text{and} \quad E[U^{(k_1-1,k_2)}_{n_1,n_2}] = \theta^{(k_1-1,k_2)}.$$

From (2.5), we can write

$$Y_{n_1,n_2} - \theta = I^{(1)}_{n_1,n_2} + I^{(2)}_{n_1,n_2} + I^{(3)}_{n_1,n_2} + b_{n_1,n_2} + R^*_{n_1,n_2}, \tag{3.8}$$

where

$$\begin{aligned}
I^{(1)}_{n_1,n_2} &= \frac{d(k_1,k_1)d(k_2,k_2)}{D(n_1,k_1)D(n_2,k_2)}\binom{n_1}{k_1}\binom{n_2}{k_2}(U_{n_1,n_2} - \theta),\\
I^{(2)}_{n_1,n_2} &= \frac{d(k_1,k_1)d(k_2,k_2-1)}{D(n_1,k_1)D(n_2,k_2)}\binom{n_1}{k_1}\binom{n_2}{k_2-1}(U^{(k_1,k_2-1)}_{n_1,n_2} - \theta^{(k_1,k_2-1)}),\\
I^{(3)}_{n_1,n_2} &= \frac{d(k_1,k_1-1)d(k_2,k_2)}{D(n_1,k_1)D(n_2,k_2)}\binom{n_1}{k_1-1}\binom{n_2}{k_2}(U^{(k_1-1,k_2)}_{n_1,n_2} - \theta^{(k_1-1,k_2)}),\\
b_{n_1,n_2} &= \frac{d(k_1,k_1)d(k_2,k_2-1)}{D(n_1,k_1)D(n_2,k_2)}\binom{n_1}{k_1}\binom{n_2}{k_2-1}(\theta^{(k_1,k_2-1)} - \theta)\\
&\quad + \frac{d(k_1,k_1-1)d(k_2,k_2)}{D(n_1,k_1)D(n_2,k_2)}\binom{n_1}{k_1-1}\binom{n_2}{k_2}(\theta^{(k_1-1,k_2)} - \theta),
\end{aligned}$$

and

$$\begin{aligned}
R^*_{n_1,n_2} = \frac{1}{D(n_1,k_1)D(n_2,k_2)}\Bigg[&d(k_1,k_1)\binom{n_1}{k_1}\sum_{j_2=1}^{k_2-2} d(k_2,j_2)\binom{n_2}{j_2}(U^{(k_1,j_2)} - \theta)\\
&+ d(k_2,k_2)\binom{n_2}{k_2}\sum_{j_1=1}^{k_1-2} d(k_1,j_1)\binom{n_1}{j_1}(U^{(j_1,k_2)}_{n_1,n_2} - \theta)\\
&+ \sum_{j_1=1}^{k_1-1}\sum_{j_2=1}^{k_2-1} d(k_1,j_1)d(k_2,j_2)\\
&\times \binom{n_1}{j_1}\binom{n_2}{j_2}(U^{(j_1,j_2)}_{n_1,n_2} - \theta)\Bigg].
\end{aligned}$$

We evaluate $I^{(1)}_{n_1,n_2}$, $I^{(2)}_{n_1,n_2}$, $I^{(3)}_{n_1,n_2}$, b_{n_1,n_2} and $R^*_{n_1,n_2}$ as the followings (i), (ii), (iii) and (iv).

(i) From (3.1), we have

$$I^{(1)}_{n_1,n_2} = \frac{d(k_1,k_1)}{D(n_1,k_1)}\binom{n_1}{k_1}\frac{d(k_2,k_2)}{D(n_2,k_2)}\binom{n_2}{k_2}(U_{n_1,n_2} - \theta) \tag{3.9}$$

$$= \left(1 - \frac{\delta_{k_1}}{n_1} - \frac{\delta_{k_2}}{n_2} + O\left(\frac{1}{N^2}\right)\right)(U_{n_1,n_2} - \theta)$$
$$= (U_{n_1,n_2} - \theta) + R^{**}_{n_1,n_2}$$

where $E(R^{**}_{n_1,n_2})^2 = O(N^{-3})$ because of $\mathrm{Var}[U_{n_1,n_2}] = O(N^{-1})$.

(ii) From (3.1) and (3.2), we have

$$I^{(2)}_{n_1,n_2} = \frac{d(k_1,k_1)}{D(n_1,k_1)}\binom{n_1}{k_1}\frac{d(k_2,k_2-1)}{D(n_2,k_2)}\binom{n_2}{k_2-1}(U^{(k_1,k_2-1)}_{n_1,n_2} - \theta^{(k_1,k_2-1)})$$
$$= \left(1 - \frac{\delta_{k_1}}{n_1} + O\left(\frac{1}{N}\right)\right)\frac{\delta_{k_2}}{n_2}(U^{(k_1,k_2-1)}_{n_1,n_2} - \theta^{(k_1,k_2-1)}).$$

By (3.5), we have

$$E[I^{(2)}_{n_1,n_2}]^2 = O(N^{-3}). \tag{3.10}$$

Similarly,

$$E[I^{(3)}_{n_1,n_2}]^2 = O(N^{-3}). \tag{3.11}$$

(iii) From (3.1) and (3.2), we have

$$b_{n_1,n_2} = \left(1 - \frac{\delta_{k_1}}{n_1} + O\left(\frac{1}{N}\right)\right)\frac{\delta_{k_2}}{n_2}(\theta^{(k_1,k_2-1)} - \theta)$$
$$+ \left(1 - \frac{\delta_{k_2}}{n_2} + O\left(\frac{1}{N}\right)\right)\frac{\delta_{k_1}}{n_1}(\theta^{(k_1-1,k_2)} - \theta)$$
$$= \frac{\delta_{k_1}}{n_1}(\theta^{(k_1-1,k_2)} - \theta) + \frac{\delta_{k_2}}{n_2}(\theta^{(k_1,k_2-1)} - \theta) + O(N^{-2}).$$

Thus we get

$$b_{n_1,n_2} = \frac{1}{N}b^{(0)} + o(N^{-1}). \tag{3.12}$$

(iv) Since $\mathrm{Var}[U^{(j_1,j_2)}_{n_1,n_2}] = O(N^{-1})$ for $j_1, j_2 \geq 1$ and $[d(k,j)/D(n,k)]\binom{n}{j} = O(n^{-2})$ $(j = 1, \dots, k-2)$, we have

$$E(R^{*}_{n_1,n_2})^2 = O(N^{-4}). \tag{3.13}$$

Applying (3.9), (3.10), (3.11), (3.12) and (3.13) to (3.8), we get (3.6). □

4. Edgeworth expansion

For the two-sample U-statistic U_{n_1,n_2}, $\sqrt{N}(U_{n_1,n_2} - \theta)$ converges to Normal distribution $N(0,\sigma^2)$ as N tends to ∞ (see, for example, Lee (1990), p. 141, and Randles and Wolfe (1979), p. 92). Therefore by (3.6), $\sqrt{N}(Y_{n_1,n_2} - \theta)$ converges to the same Normal distribution. To see the difference between asymptotic distributions of these two statistics, we shall derive the Edgeworth expansion of the statistic Y_{n_1,n_2}. We put as follows.

$$\psi_{1,0}(x_1) = E[h(x_1, X_2, \dots, X_{k_1}; Y_1, \dots, Y_{k_2})],$$
$$\psi_{0,1}(y_1) = E[h(X_1, X_2, \dots, X_{k_1}; y_1, Y_2, \dots, Y_{k_2})],$$

$$\psi_{2,0}(x_1, x_2) = E[h(x_1, x_2, X_3, \dots, X_{k_1}; Y_1, \dots, Y_{k_2})],$$
$$\psi_{0,2}(y_1, y_2) = E[h(X_1, X_2, \dots, X_{k_1}; y_1, y_2, Y_3, \dots, Y_{k_2})],$$
$$\psi_{1,1}(x_1; y_1) = E[h(x_1, X_2, \dots, X_{k_1}; y_1, Y_2, \dots, Y_{k_2})],$$
$$h^{(1,0)}(x_1) = \psi_{1,0}(x_1) - \theta, \qquad h^{(0,1)}(y_1) = \psi_{0,1}(y_1) - \theta,$$
$$h^{(2,0)}(x_1, x_2) = \psi_{2,0}(x_1, x_2) - \psi_{1,0}(x_1) - \psi_{1,0}(x_2) + \theta,$$
$$h^{(0,2)}(y_1, y_2) = \psi_{2,0}(x_1, x_2) - \psi_{0,1}(y_1) - \psi_{0,1}(y_2) + \theta,$$
$$h^{(1,1)}(x_1; y_1) = \psi_{1,1}(x_1; y_1) - \psi_{1,0}(x_1) - \psi_{0,1}(y_1) + \theta,$$
$$\delta_{1,0}^2 = \mathrm{Var}(h^{(1,0)}(X_1)) = E[\psi_{1,0}(X_1) - \theta]^2,$$
$$\delta_{0,1}^2 = \mathrm{Var}(h^{(0,1)}(Y_1)) = E[\psi_{0,1}(Y_1) - \theta]^2.$$

In this paper, we assume that

$$\delta_{1,0}^2 > 0 \qquad \text{and} \qquad \delta_{0,1}^2 > 0.$$

That is, we assume that the kernel h is non-degenerate. Next, we put

$$\sigma_N^2 = \mathrm{Var}[U_{n_1,n_2}] = E[(U_{n_1,n_2} - \theta)^2]$$

and

$$\sigma_N^{*2} = \frac{k_1^2}{n_1}\delta_{1,0}^2 + \frac{k_2^2}{n_2}\delta_{0,1}^2.$$

Then, we have the relation

$$\sigma_N^2 = \sigma_N^{*2} + O(N^{-2}).$$

Furthermore, we put

$$\sigma^2 = \frac{k_1^2}{p}\delta_{1,0}^2 + \frac{k_2^2}{1-p}\delta_{0,1}^2$$

and

$$\begin{aligned}\eta_{2,N} = \frac{1}{\sigma_N^{*3}}\bigg\{ & \frac{k_1^3}{n_1^2}E[(h^{(1,0)}(X_1))^3] + \frac{k_2^3}{n_2^2}E[(h^{(0,1)}(Y_1))^3] \\ & + \frac{6k_1^2k_2^2}{n_1n_2}E[h^{(1,0)}(X_1)h^{(0,1)}(Y_1)h^{(1,1)}(X_1;Y_1)] \\ & + \frac{3k_1^3(k_1-1)}{n_1^2}E[h^{(1,0)}(X_1)h^{(1,0)}(X_2)h^{(2,0)}(X_1,X_2)] \\ & + \frac{3k_2^3(k_2-1)}{n_2^2}E[h^{(0,1)}(Y_1)h^{(0,1)}(Y_2)h^{(0,2)}(Y_1,Y_2)]\bigg\}.\end{aligned}$$

The right-hand side of $\eta_{2,N}$ is due to Maesono (1985). The last three expectations on the right-hand side are rewritten as follows:

$$\begin{aligned}& E[h^{(1,0)}(X_1)h^{(0,1)}(Y_1)h^{(1,1)}(X_1;Y_1)] \\ & \qquad = E[\psi_{1,0}(X_1)\psi_{0,1}(Y_1)\psi_{1,1}(X_1;Y_1)] - \theta(\delta_{1,0}^2 + \delta_{0,1}^2) - \theta^3,\end{aligned}$$

$$E[h^{(1,0)}(X_1)h^{(1,0)}(X_2)h^{(2,0)}(X_1,X_2)]$$
$$= E[\psi_{1,0}(X_1)\psi_{1,0}(X_2)\psi_{2,0}(X_1,X_2)] - \theta\delta_{1,0}^2 - \theta^3,$$
$$E[h^{(0,1)}(Y_1)h^{(0,1)}(Y_2)h^{(0,2)}(Y_1,Y_2)]$$
$$= E[\psi_{0,1}(Y_1)\psi_{0,1}(Y_2)\psi_{0,2}(Y_1,Y_2)] - \theta\delta_{0,1}^2 - \theta^3.$$

We put

$$Q(x) = \Phi(x) + \eta_{2,N}(1-x^2)\phi(x).$$

Let φ_1 and φ_2 be the characteristic functions of the random variables $h^{(1,0)}(X_1)$ and $h^{(0,1)}(Y_1)$, respectively. The following is the Edgeworth expansion of the two-sample U-statistic U_{n_1,n_2} by Koroljuk and Borovskich (1994).

LEMMA 4.1. (Koroljuk and Borovskich (1994), Theorem 6.3.2) *Suppose that the Cramer condition*

$$\lim_{|t|\to\infty} \sup |\varphi_j(t)| < 1, \qquad j = 1,2 \tag{4.1}$$

and the moment condition

$$E[|h(X_1,\ldots,X_{k_1};Y_1,\ldots,Y_{k_2})|^3] < \infty \tag{4.2}$$

is satisfied. Then

$$\sup_x |P(\sigma_N^{-1}(U_{n_1,n_2}-\theta) \le x) - Q(x)| = O(N^{-3/5}) \tag{4.3}$$

as $N \to \infty$.

Before giving the Edgeworth expansion of the statistic Y_{n_1,n_2}, we show the relation between the two statistics U_{n_1,n_2} and Y_{n_1,n_2}.

PROPOSITION 4.2. *Suppose that $w(1,\ldots,1;k) > 0$ and (3.4) is satisfied. Then*

$$\sup_x \left| P(\sigma_N^{*-1}(Y_{n_1,n_2}-\theta) \le x) - P\left(\sigma_N^{-1}(U_{n_1,n_2}-\theta) + \frac{1}{\sqrt{N}\sigma} b^{(0)} \le x\right) \right| = o(N^{-1/2}). \tag{4.4}$$

PROOF. From (3.6), we have

$$\sigma_N^{*-1}(Y_{n_1,n_2}-\theta) = \sigma_N^{*-1}(U_{n_1,n_2}-\theta) + \frac{1}{N\sigma_N^*} b^{(0)} + R_{n_1,n_2}^{***}$$

where $E[|R_{n_1,n_2}^{***}|] = o(N^{-1/2})$.

Thus,

$$\sup_x \left| P(\sigma_N^{*-1}(Y_{n_1,n_2}-\theta) \le x) - P\left(\sigma_N^{*-1}(U_{n_1,n_2}-\theta) + \frac{1}{N\sigma_N^*} b^{(0)} \le x\right) \right| = o(N^{-1/2}), \tag{4.5}$$

where we use the relation

$$\sup_x |P(W+\Delta \le x) - P(W \le x)| \le 4(E|W\Delta| + E|\Delta|) \tag{4.6}$$

for any random variables W and Δ (see Shorack (2000), p. 261).

Since

$$E[|\sigma_N^{*-1} - \sigma_N^{-1}| \cdot |U_{n_1,n_2} - \theta|] = O(N^{-1}),$$

using the relation (4.6), we have

$$\begin{aligned}\sup_x \Bigg| P\left(\sigma_N^{*-1}(U_{n_1,n_2} - \theta) + \frac{1}{N\sigma_N^*} b^{(0)} \le x\right) \\ - P\left(\sigma_N^{-1}(U_{n_1,n_2} - \theta) + \frac{1}{N\sigma_N^*} b^{(0)} \le x\right)\Bigg| \\ = O(N^{-1}).\end{aligned} \tag{4.7}$$

Since

$$\frac{1}{N\sigma_N^*} - \frac{1}{\sqrt{N}\sigma} = o(N^{-1/2}),$$

we have

$$\begin{aligned}\sup_x \Bigg| P\left(\sigma_N^{-1}(U_{n_1,n_2} - \theta) + \frac{1}{N\sigma_N^*} b^{(0)} \le x\right) \\ - P\left(\sigma_N^{-1}(U_{n_1,n_2} - \theta) + \frac{1}{\sqrt{N}\sigma} b^{(0)} \le x\right)\Bigg| \\ = o(N^{-1/2}).\end{aligned} \tag{4.8}$$

Thus by (4.5), (4.7) and (4.8), we get (4.4). □

We put

$$\begin{aligned} Q^*(x) &= Q(x) + \frac{b^{(0)}}{\sqrt{N}\sigma}\left[-\frac{1}{2}\phi(x) + x(1-x^2)\phi(x)\right] \\ &= \Phi(x) - \frac{b^{(0)}}{2\sqrt{N}\sigma}\phi(x) + \eta_{2,N}(1-x^2)\phi(x) + \frac{b^{(0)}}{\sqrt{N}\sigma}x(1-x^2)\phi(x). \end{aligned} \tag{4.9}$$

The Edgeworth expansion (4.3) of U_{n_1,n_2} is derived by standardising with its standard deviation σ_N. Since σ_N is not the standard deviation of the statistic Y_{n_1,n_2}, we shall standardise Y_{n_1,n_2} by using its asymptotic standard deviation σ_N^* and obtain the Edgeworth expansion of Y_{n_1,n_2} with the remainder term $o(N^{-1/2})$ as follows.

THEOREM 4.3. *Suppose that* $w(1,\dots,1;k) > 0$, *and* (3.4), (4.1) *and* (4.2) *are satisfied. Then*

$$\sup_x |P(\sigma_N^{*-1}(Y_{n_1,n_2} - \theta) \le x) - Q^*(x)| = o(N^{-1/2}). \tag{4.10}$$

PROOF. From (4.3), we have

$$\sup_x \left| P\left(\sigma_N^{-1}(U_{n_1,n_2}-\theta) \le x - \frac{1}{\sqrt{N}\sigma} b^{(0)}\right) - Q\left(x - \frac{1}{\sqrt{N}\sigma} b^{(0)}\right)\right| = O(N^{-3/5}).$$

Thus by this relation and (4.4), we get

$$\sup_x \left| P(\sigma_N^{*-1}(Y_{n_1,n_2}-\theta) \le x) - Q\left(x - \frac{1}{\sqrt{N}\sigma} b^{(0)}\right)\right| = o(N^{-1/2}), \tag{4.11}$$

where

$$\begin{aligned} Q\left(x - \frac{b^{(0)}}{\sqrt{N}\sigma}\right) &= \Phi\left(x - \frac{1}{\sqrt{N}\sigma} b^{(0)}\right) \\ &\quad + \eta_{2,N}\left[1 - \left(x - \frac{1}{\sqrt{N}\sigma} b^{(0)}\right)^2\right] \phi\left(x - \frac{1}{\sqrt{N}\sigma} b^{(0)}\right) \\ &= Q^*(x) + O(N^{-1}). \end{aligned} \tag{4.12}$$

Thus by (4.11) and (4.12), we get (4.10). □

COROLLARY 4.4. *The difference between the Edgeworth expansions of the two-sample U-statistic U_{n_1,n_2} and the statistic Y_{n_1,n_2} is given by*

$$\frac{b^{(0)}}{\sqrt{N}\sigma}\left[-\frac{1}{2}\phi(x) + x(1-x^2)\phi(x)\right].$$

5. Examples

The asymptotic expansion of $(Y_{n_1,n_2}-\theta)/\sigma_N$, $Q^*(x)$, depends on $b^{(0)}$ and $\eta_{2,N}$. For 4 kernels we shall give the values of $b^{(0)}$ and $\eta_{2,N}$ about V_{n_1,n_2}, S_{n_1,n_2} and B_{n_1,n_2}, based on some special distributions.

(i) We consider the kernel

$$h(x_1,\dots,x_r;y_1,\dots,y_r) = x_1\cdots x_r - y_1\cdots y_r, \qquad r = 2,3,\dots$$

which gives the parameter $\theta = \mu^r - \nu^r$, where $\mu = E(X_1)$ and $\nu = E(Y_1)$. We assume that X and Y are symmetric about μ and ν, respectively. Then we have

$$\theta^{(r-1,r)} = E(X_1^2)\cdot\mu^{r-2} - \nu^r, \qquad \theta^{(r,r-1)} = \mu^r - E(Y_1^2)\cdot\nu^{r-2},$$

and

$$b^{(0)} = \delta_r\left[\frac{1}{p}\mu^{r-2}\operatorname{Var}(X_1) - \frac{1}{1-p}\nu^{r-2}\operatorname{Var}(Y_1)\right],$$

where $\delta_r = r(r-1)/2$ for V_{n_1,n_2} and S_{n_1,n_2}, and $\delta_r = r(r-1)$ for B_{n_1,n_2}.

Next we evaluate $\eta_{2,N}$. We have

$$\begin{aligned} &h^{(1,0)}(x_1) = \mu^{r-1}(x_1-\mu), \qquad h^{(0,1)}(y_1) = -\nu^{r-1}(y_1-\nu), \\ &h^{(2,0)}(x_1,x_2) = \mu^{r-2}(x_1-\mu)(x_2-\mu), \\ &h^{(0,2)}(y_1,y_2) = -\nu^{r-2}(y_1-\mu)(y_2-\mu), \qquad \text{and} \qquad h^{(1,1)}(x_1;y_1) = 0. \end{aligned}$$

Thus we get

$$\delta_{1,0}^2 = \mu^{2r-2}\,\mathrm{Var}(X_1), \qquad \delta_{0,1}^2 = \nu^{2r-2}\,\mathrm{Var}(Y_1)$$

and

$$\sigma_N^{*2} = r^2\left(\frac{\mu^{2r-2}\,\mathrm{Var}(X_1)}{n_1} + \frac{\nu^{2r-2}\,\mathrm{Var}(Y_1)}{n_2}\right).$$

Therefore, we have

$$\eta_{2,N} = 3(r-1)\left(\frac{\mu^{3r-4}[\mathrm{Var}(X_1)]^2}{n_1^2} - \frac{\nu^{3r-4}[\mathrm{Var}(Y_1)]^2}{n_2^2}\right) \times \left(\frac{\mu^{2r-2}\,\mathrm{Var}(X_1)}{n_1} + \frac{\nu^{2r-2}\,\mathrm{Var}(Y_1)}{n_2}\right)^{-3/2}.$$

(ii) We consider the kernel given by

$$h(x_1, x_2; y_1, y_2) = \begin{cases} 1 & (x_1, x_2 < y_1, y_2 \text{ or } x_1, x_2 > y_1, y_2) \\ 0 & \text{otherwise.} \end{cases}$$

The corresponding parameter $\theta = Eh(X_1, X_2; Y_1, Y_2)$ is equal to $2\Delta + 1/3$, where Δ is given by **(iii)** of Section 1. We consider the uniform distributions $U(0,1)$ and $U(1/2, 3/2)$ as F and G, respectively. Then we have

$$\theta = \frac{2}{3}, \qquad \theta^{(1,2)} = \frac{5}{6} \qquad \text{and} \qquad \theta^{(2,1)} = \frac{5}{6}.$$

Thus we get

$$b^{(0)} = \frac{\delta_2}{6p(1-p)},$$

where $\delta_2 = 1$ for V_{n_1,n_2} and S_{n_1,n_2}, and $\delta_2 = 2$ for B_{n_1,n_2}.

For $0 < x_1, x_2 < 1$ and $1/2 < y_1, y_2 < 3/2$, we have

$$\psi_{1,0}(x_1) = \begin{cases} \dfrac{19}{24} & (0 < x_1 < 1/2) \\ -x_1 + \dfrac{31}{24} & (1/2 \le x_1 < 1), \end{cases}$$

$$\psi_{0,1}(y_1) = \begin{cases} y_1 - \dfrac{5}{24} & (1/2 < y_1 < 1) \\ \dfrac{19}{24} & (1 \le y_1 < 3/2), \end{cases}$$

$$\psi_{2,0}(x_1, x_2) = \left[\frac{3}{2} - \max(x_1, x_2)\right]^2 I_{(0.5,1)}(\max(x_1, x_2)) + I_{(0,0.5)}(\max(x_1, x_2)) + \left[\min(x_1, x_2) - \frac{1}{2}\right]^2 I_{(0.5,1)}(\min(x_1, x_2)),$$

$$\psi_{0,2}(y_1, y_2) = [\min(y_1, y_2)]^2 I_{(0.5,1)}(\min(y_1, y_2)) + I_{(1,1.5)}(\min(y_1, y_2)) + [1 - \max(y_1, y_2)]^2 I_{(0.5,1)}(\max(y_1, y_2)),$$

and

$$\psi_{1,1}(x_1;y_1)=\begin{cases}\dfrac{7}{8} & (0<x_1<1/2, y_1>1)\\ -\dfrac{1}{8}+\dfrac{3}{2}y_1-\dfrac{1}{2}y_1^2 & (0<x_1<1/2, y_1\le 1)\\ 1-\dfrac{1}{2}x_1^2 & (1/2<x_1<1, 1\le y_1<3/2)\\ \dfrac{3}{2}y_1-\dfrac{1}{2}(x_1^2+y_1^2) & (1/2\le y_1<1, 1/2\le x_1<1, x_1<y_1)\\ -\dfrac{1}{2}+x_1+\dfrac{1}{2}y_1-\dfrac{1}{2}(x_1^2+y_1^2) & \\ & (1/2\le y_1<1, 1/2\le x_1<1, x_1>y_1).\end{cases}$$

Thus by using Mathematica to compute the integrals we get

$$\delta_{1,0}^2=E[\psi_{1,0}(X_1)-\theta]^2=\frac{5}{192},\qquad \delta_{0,1}^2=E[\psi_{0,1}(Y_1)-\theta]^2=\frac{5}{192},$$
$$\sigma_N^{*2}=\frac{10}{81}\left(\frac{1}{n_1}+\frac{1}{n_2}\right),$$

and

$$E[(h^{(1,0)}(X_1))^3]=-\frac{1}{256},\qquad E[(h^{(0,1)}(Y_1))^3]=-\frac{1}{256},$$
$$E[h^{(1,0)}(X_1)h^{(0,1)}(Y_1)h^{(1,1)}(X_1;Y_1)]=-\frac{1}{960},$$
$$E[h^{(1,0)}(X_1)h^{(1,0)}(X_2)h^{(2,0)}(X_1,X_2)]=\frac{13}{2560},$$
$$E[h^{(0,1)}(Y_1)h^{(0,1)}(Y_2)h^{(0,2)}(Y_1,Y_2)]=\frac{107}{20480}.$$

Therefore we have

$$\eta_{2,N}=\frac{729\sqrt{10}}{256000}\left(\frac{232}{n_1^2}+\frac{241}{n_2^2}-\frac{256}{n_1n_2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^{-3/2}.$$

(iii) The kernel given by **(iv)** of Section 1 is considered. We have $\theta^{(1,2)}=1$ and $\theta^{(2,1)}=0$. We consider the uniform ditributions $U(0,1)$ and $U(1/4,3/4)$ as F and G, respectively. Then we have $\theta=7/24$ and

$$b^{(0)}=\frac{\delta_2}{p(1-p)}\left(\frac{17}{24}-p\right).$$

Next we evaluate $\eta_{2,N}$. For $0<x_1,x_2<1$ and $1/4<y_1,y_2<3/4$ we have

$$\psi_{1,0}(x_1)=\frac{1}{3}-\frac{4}{3}\left(x_1-\frac{1}{2}\right)^3,$$
$$\psi_{0,1}(y_1)=2\left\{\left(y_1-\frac{1}{4}\right)^2+\left(\frac{3}{4}-y_1\right)^2\right\}+\frac{2}{3}\left\{\left(y_1-\frac{1}{4}\right)^3-\left(\frac{3}{4}-y_1\right)^3\right\},$$
$$\psi_{2,0}(x_1,x_2)=\begin{cases}(1-2|x_1-x_2|)^2 & (2|x_1-x_2|<1)\\ 0 & (2|x_1-x_2|\ge 1),\end{cases}$$

and

$$\psi_{0,2}(y_1,y_2) = 2|y_1-y_2| - (y_1-y_2)^2.$$

Futhermore, for $0 < x_1 < 1/2$ and $y_1 + x_1 < 3/4$

$$\psi_{1,1}(x_1;y_1) = \begin{cases} \left(x_1+y_1-\frac{1}{4}\right)^2 + \left(x_1-y_1+\frac{3}{4}\right)^2 - 4x_1^2 & (y_1-x_1>1/4) \\ \left(x_1-y_1+\frac{3}{4}\right)^2 + 2\left(y_1-\frac{1}{4}\right)^2 - 2x_1^2 & (y_1-x_1\le 1/4). \end{cases}$$

For $0 < x_1 < 1/2$ and $y_1 + x_1 > 3/4$

$$\psi_{1,1}(x_1;y_1) = \begin{cases} \left(x_1+y_1-\frac{1}{4}\right)^2 - 2x_1^2 + 2\left(y_1-\frac{3}{4}\right)^2 & (y_1-x_1>1/4) \\ 2\left\{\left(y_1-\frac{1}{4}\right)^2 + \left(y_1-\frac{3}{4}\right)^2\right\} & (y_1-x_1\le 1/4). \end{cases}$$

For $1/2 < x_1 < 1$ and $y_1 + x_1 - 1 < 1/4$,

$$\psi_{1,1}(x_1;y_1) = \begin{cases} 2\left\{\left(y_1-\frac{1}{4}\right)^2 + \left(y_1-\frac{3}{4}\right)^2\right\} & (y_1-x_1+1>3/4) \\ 2\left(y_1-\frac{1}{4}\right)^2 + \left(\frac{5}{4}-x_1-y_1\right)^2 - 2(1-x_1)^2 & \\ & (y_1-x_1+1\le 3/4). \end{cases}$$

For $1/2 < x_1 < 1$ and $y_1 + x_1 - 1 > 1/4$,

$$\psi_{1,1}(x_1;y_1) = \begin{cases} \left(\frac{3}{4}-x_1+y_1\right)^2 + \left(\frac{7}{4}-x_1-y_1\right)^2 - 4(1-x_1)^2 \\ \qquad (y_1-x_1+1<3/4) \\ \left(\frac{3}{4}-x_1+y_1\right)^2 - 2(1-x_1)^2 + 2\left(\frac{3}{4}-y_1\right)^2 \\ \qquad (y_1-x_1+1>3/4). \end{cases}$$

Thus by using Mathematica to compute the integrals we get

$$\delta_{1,0}^2 = E[\psi_{1,0}(X_1)-\theta]^2 = \frac{23}{4032},$$
$$\delta_{0,1}^2 = E[\psi_{0,1}(Y_1)-\theta]^2 = \frac{37}{4032},$$
$$\sigma_N^{*2} = \frac{1}{1008}\left(\frac{23}{n_1}+\frac{37}{n_2}\right)$$

and

$$E[(h^{(1,0)}(X_1))^3] = \frac{55}{96768}, \qquad E[(h^{(0,1)}(Y_1))^3] = \frac{821}{483840},$$

$$E[h^{(1,0)}(X_1)h^{(0,1)}(Y_1)h^{(1,1)}(X_1;Y_1)] = \frac{50531}{92897280},$$
$$E[h^{(1,0)}(X_1)h^{(1,0)}(X_2)h^{(2,0)}(X_1,X_2)] = \frac{3569}{2419200},$$
$$E[h^{(0,1)}(Y_1)h^{(0,1)}(Y_2)h^{(0,2)}(Y_1,Y_2)] = -\frac{1009}{1036800}.$$

Therefore we have

$$\eta_{2,N} = \frac{\sqrt{7}}{400}\left(\frac{193312}{n_1^2} - \frac{47328}{n_2^2} + \frac{252655}{n_1 n_2}\right)\left(\frac{23}{n_1} + \frac{37}{n_2}\right)^{-3/2}.$$

(iv) We consider the kernel given by **(v)** of Section 1. We consider the uniform ditributions $U(0,1)$ and $U(1/2,3/2)$ as F and G, respectively. Then we have

$$\theta^{(1,2)} = \frac{5}{6}, \qquad \theta^{(2,1)} = \frac{11}{12}, \qquad \theta = \frac{23}{24},$$

and

$$b^{(0)} = -\frac{\delta_2(3-2p)}{24p(1-p)}.$$

For $0 < x_1, x_2 < 1$ and $1/2 < y_1, y_2 < 3/2$,

$$\psi_{1,0}(x_1) = 1 - \frac{1}{6}x_1^3, \qquad \psi_{0,1}(y_1) = 1 - \frac{1}{6}\left(\frac{3}{2} - y_1\right)^3,$$

$$\psi_{2,0}(x_1,x_2) = \begin{cases} 1 & (0 < x_1 + x_2 \le 1) \\ 1 - \frac{1}{2}(x_1 + x_2 - 1)^2 & (1 < x_1 + x_2 < 2), \end{cases}$$

$$\psi_{0,2}(y_1,y_2) = \begin{cases} 1 - \frac{1}{2}[2 - (y_1 + y_2)]^2 & (1 < y_1 + y_2 < 2) \\ 1 & (2 \le y_1 + y_2 < 3), \end{cases}$$

and

$$\psi_{1,1}(x_1;y_1) = \begin{cases} \frac{7}{8} + \frac{1}{2}(y_1 - x_1) - \frac{1}{2}(y_1 - x_1)^2 & \left(-\frac{1}{2} < y_1 - x_1 < \frac{1}{2}\right) \\ 1 & \left(\frac{1}{2} \le y_1 - x_1 < \frac{3}{2}\right). \end{cases}$$

Thus by using Mathematica to compute the integrals we get

$$\delta_{1,0}^2 = \frac{1}{448}, \qquad \delta_{0,1}^2 = \frac{1}{448}, \qquad \sigma_N^{*2} = \frac{1}{112}\left(\frac{1}{n_1} + \frac{1}{n_2}\right),$$
$$E[(h^{(1,0)}(X_1))^3] = -\frac{1}{8960}, \qquad E[(h^{(0,1)}(Y_1))^3] = -\frac{1}{8960},$$

$$E[h^{(1,0)}(X_1)h^{(0,1)}(Y_1)h^{(1,1)}(X_1;Y_1)] = -\frac{9727}{7257600},$$
$$E[h^{(1,0)}(X_1)h^{(1,0)}(X_2)h^{(2,0)}(X_1,X_2)] = -\frac{9727}{7257600},$$
$$E[h^{(0,1)}(Y_1)h^{(0,1)}(Y_2)h^{(0,2)}(Y_1,Y_2)] = -\frac{32531}{7257600}.$$

Therefore we have

$$\eta_{2,N} = -\frac{\sqrt{7}}{675}\left(\frac{9997}{n_1^2} + \frac{32801}{n_2^2} + \frac{38908}{n_1 n_2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-3/2}.$$

Acknowledgements

We would like to express our thanks to Prof. Maesono for the discussions on the related matters, and the two referees for their kind suggestions.

References

Charalambides, Ch. A. and Singh, J. (1988). A review of the Stirling numbers, their generalizations and statistical applications, *Commun. Statist.-Theory Meth.*, **17**, 2533–2595.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems, *Ann. Statist.*, **1**, 209–230.

Hollander, M. (1967). Asymptotic efficiency of two nonparametric competitors of Wilcoxon's two sample test, *J. Amer. Statist. Assoc.*, **62**, 939–949.

Koroljuk, V. S. and Borovskich, Yu. V. (1994). *Theory of U-statistics*, Kluwer Academic Publishers, Dordrecht.

Lee, A. J. (1990). *U-statistics*, Marcel Dekker, New York.

Lehmann, E. L. (1951). Consistency and unbiasedness of certain nonparametric tests, *Ann. Math. Statist.*, **22**, 165–179.

Maesono, Y. (1985). Edgeworth expansion for two-sample U-statistics, *Rep. Fac. Sci. Kagoshima Univ.*, (Math., Phys. & Chem.), No. 18, 35–43.

Nomachi, T., Kondo, M. and Yamato, H. (2002). Higher Order efficiency of linear combinations of U-statistics as estimators of estimable parameters, *Scientiae Mathematicae Japonicae*, **56**, 95–106.

Randles, R. H. and Wolfe, D. A. (1979). *Introduction to the Theory of Nonparametric Statistics*, John Wiley, New York.

Shorack, G. R. (2000). *Probability for Statisticians*, Springer, New York.

Toda, K. and Yamato, H. (2001). Berry-Esseen bounds for some statistics including LB-statistic and V-statistic, *J. Japan Statist. Soc.*, **31**, 225–237.

Yamato, H. (1977). Relations between limiting Bayes estimates and U-statistics for estimable parameters, *J. Japan Statist. Soc.*, **7**, 57–66.

Yamato, H., Nomachi, T. and Toda, K. (2003). Edgeworth expansions of some statistics including the LB-statistic and V-statistic, *J. Japan Statist. Soc.*, **33**, 77–94.