RESEARCH ARTICLE
doi: 10.2306/scienceasia1513-1874.2014.40.182

ScienceAsia 40 (2014): **182–191**

Bounds of the normal approximation to random-sum Wilcoxon statistics

Mongkhon Tuntapthai, Nattakarn Chaidee*

Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University, Bangkok 10330 Thailand

*Corresponding author, e-mail: nattakarn.c@chula.ac.th.

Received 14 May 2013
Accepted 16 Feb 2014

ABSTRACT: Consider sequences $\{X_i\}_{i=1}^{\infty}$ and $\{Y_j\}_{j=1}^{\infty}$ of independent and identically distributed (i.i.d.) random variables, random variables K_1, K_2 ranging over of all positive integers, where the X_i's, Y_j's, K_1, and K_2 are all independent. We obtain Berry-Esseen bounds for random-sum Wilcoxon's statistics in the form $(W_{K_1,K_2} - U)/V$ and $(W_{K_1,K_2} - a)/b$ where $W_{K_1,K_2} = \sum_{i=1}^{K_1}\sum_{j=1}^{K_2} I(X_i > Y_j)$ and U, V are random variables, and a, b are constants. We also show that the rate of convergence is $O((EK_2)^{-1/2})$ provided by $EK_1/EK_2 \to \tau$ for some constant $\tau > 0$ when EK_1 and EK_2 tend to infinity.

KEYWORDS: Wilcoxon's rank-sum statistics, Mann-Whitney's statistics, random-indexed summands, Berry-Esseen bounds, non-random centring

INTRODUCTION

Let X, X_1, $\dots$, X_m and Y, Y_1, $\dots$, Y_n be two sequences of independent and identically distributed (i.i.d.) random variables. Note that the distributions of X and Y are not necessarily identical. Suppose the X_i's and Y_j's are independent and continuous random variables. To test the null hypothesis that the distributions of X and Y are equal, Wilcoxon[1] introduced rank-sum statistics by ranking the random variables between two sequences and then computing the sum of these rankings. Mann and Whitney[2] proposed the *Mann-Whitney's statistic*,

$$W_{m,n} = \sum_{i=1}^{m}\sum_{j=1}^{n} I\left(X_i > Y_j\right).$$

Mann and Whitney[2] showed that their statistics and Wilcoxon's rank-sum statistics are equivalent. They also showed that under the null hypothesis, Mann-Whitney's distribution can be approximated by the standard normal distribution denoted by Φ.

Consider a *Wilcoxon's rank-sum statistic* defined by $\mathcal{W}_m = \sum_{k=1}^{m} R_k$ where R_k is the ranking-number of X_k between two samples X_1, $\dots$, X_m and Y_1, $\dots$, Y_n. Pestman[3] obtained normal approximation theorems and Alberink[4] investigated Berry-Esseen bound for Wilcoxon's rank-sum statistics by assuming the null hypothesis. From the result of Alberink[4] and the well-known fact that $\mathcal{W}_m = W_{m,n} + m(m+1)/2$, we obtain the bound for Mann-Whitney's statistics as the next statement.

Theorem 1 (Ref. 4) *Under the null hypothesis that the distributions of X and Y are equal, we have the following result:*

$$\sup_{x\in\mathbb{R}}\left|P\left(\frac{W_{m,n} - mn/2}{\sqrt{mn(m+n)/12}} \leqslant x\right) - \Phi(x)\right| \leqslant \frac{17}{\sqrt{m+n}}\left\{1 + \frac{3\sqrt{3}(m^2+n^2)}{4\sqrt{mn}(m+n)}\right\}.$$

It is obvious that Mann-Whitney's statistics are example of U-statistics introduced by Hoeffding[5] and Lehmann[6]. So that the bounds for Mann-Whitney's statistics were established by applying the results of Grams and Serfling[7] or Chen and Shao[8]. Now, we need some notation:

$$\begin{aligned}
\theta &= P\left(X > Y\right),\\
\mu &= \theta\left(1-\theta\right),\\
\sigma_1^2 &= E\left|E\left\{I\left(X > Y\right)\middle| X\right\} - \theta\right|^2,\\
\gamma_1 &= E\left|E\left\{I\left(X > Y\right)\middle| X\right\} - \theta\right|^3,\\
\sigma_2^2 &= E\left|E\left\{I\left(X > Y\right)\middle| Y\right\} - \theta\right|^2,\\
\gamma_2 &= E\left|E\left\{I\left(X > Y\right)\middle| Y\right\} - \theta\right|^3.
\end{aligned}$$

Consider a *two-sample U-statistic* given by

$$U_{m,n} = \frac{1}{m}\frac{1}{n}\sum_{i=1}^{m}\sum_{j=1}^{m} h(X_i, Y_j),$$

where $h(X_i, Y_j) = I(X_i > Y_j) - \theta$. Observe that $Eh(X, Y) = 0$, $Eh^2(X, Y) = \mu$ and that $\sigma_\ell^2 \leqslant \mu$ for all $\ell = 1, 2$. Also, we note that

$$\frac{U_{m,n}}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}} = \frac{W_{m,n} - mn\theta}{\sqrt{mn^2\sigma_1^2 + nm^2\sigma_2^2}}.$$

Hence we can use the bounds for $U_{m,n}$ of Chen and Shao[8] to derive the bounds for Mann-Whitney's statistics as the next statement.

Theorem 2 (Ref. 8)

$$\sup_{x\in\mathbb{R}}\left|P\left(\frac{W_{m,n} - mn\theta}{\sqrt{mn\left(n\sigma_1^2 + m\sigma_2^2\right)}} \leqslant x\right) - \Phi(x)\right| \leqslant \frac{(1+\sqrt{2})\sqrt{\mu}\,(1/m + 1/n)}{\sqrt{\sigma_1^2/m + \sigma_2^2/n}} + \frac{6.6\left(\gamma_1/m^2 + \gamma_2/n^2\right)}{\left(\sigma_1^2/m + \sigma_2^2/n\right)^{3/2}},$$

and there exists a constant C (which does not depend on m and n) such that for all $x \in \mathbb{R}$,

$$\left|P\left(\frac{W_{m,n} - mn\theta}{\sqrt{mn\left(n\sigma_1^2 + m\sigma_2^2\right)}} \leqslant x\right) - \Phi(x)\right| \leqslant \frac{9\,\mu\,(1/m + 1/n)^2}{(1+x)^2\left(\sigma_1^2/m + \sigma_2^2/n\right)} + \frac{13.5\sqrt{\mu}\,(1/m + 1/n)}{\mathrm{e}^{x/3}\sqrt{\sigma_1^2/m + \sigma_2^2/n}} + \frac{C\left(\gamma_1/m^2 + \gamma_2/n^2\right)}{(1+x)^3\left(\sigma_1^2/m + \sigma_2^2/n\right)^{3/2}}.$$

In particular, if $\theta = 1/2$ and $\sigma_1^2 = \sigma_2^2 = 1/12$, then the constant 17 in Theorem 1 can be replaced by $(1+\sqrt{2})\sqrt{3}$.

Mann-Whitney's and Wilcoxon's statistics are usually used to compare two treatments in many comparative experiments. However, these statistics cannot be used when the sample sizes of observation from two treatments are not measured[9,10]. For example, the number of normal and abnormal images in the medical imaging studies and diagnostic marker studies. In 2011, under the situation that the sample sizes are random variables, Tang and Balakrishnan[10] proposed a random-sum Wilcoxon's statistic as the following:

Let $\{X_i\}_{i=1}^{\infty}$ and $\{Y_j\}_{j=1}^{\infty}$ be two sequences of i.i.d. random variables such that the X_i's and Y_j's are independent. Assume that the X_i's and Y_j's are continuous random variables. Let K_1 and K_2 be independent positive integer-valued random variables which are independent of the X_i's and Y_j's. *A random-sum Wilcoxon's statistic* is defined by

$$W_{K_1,K_2} = \sum_{i=1}^{K_1}\sum_{j=1}^{K_2} I\left(X_i > Y_j\right).$$

Tang and Balakrishnan[10] also obtained normal approximation theorems for random-sum Wilcoxon's statistics as the next statement.

Theorem 3 (Ref. 10) *For each $\ell = 1, 2$, let K_ℓ be a positive integer-valued random variable depending on a parameter τ_ℓ. Suppose that*

(C1) *the random variable K_ℓ has its asymptotic distribution as the normal distribution with mean EK_ℓ and variance* $\mathrm{Var}(K_\ell)$, *as $\tau_\ell \to \infty$;*

(C2) *$EK_2/EK_1 \to \eta$, for some finite non-zero constant η, as $\tau_\ell \to \infty$;*

(C3) $\mathrm{Var}(K_\ell)/EK_\ell \to \delta_\ell$, *for some finite constant δ_ℓ as $\tau_\ell \to \infty$;*

(C4) *$(EK_\ell)^p/\tau_\ell \to \lambda_\ell$, for some finite non-zero constant λ_ℓ and for some $p \geqslant 1$, as $\tau_\ell \to \infty$.*

Then

$$\left|P\left(\frac{\tau_1^{-3/2p}\left(W_{K_1,K_2} - \theta EK_1EK_2\right)}{\sqrt{\mathrm{Var}\left(\tau_1^{-3/2p}W_{K_1,K_2}\right)}} \leqslant x\right) - \Phi(x)\right|$$

tends to zero, where

$$\mathrm{Var}\left(\tau_1^{-3/2p}W_{K_1,K_2}\right) \to \lambda_1^3\eta^2\sigma_1^2 + \lambda_1^3\eta\sigma_2^2 + \left(\delta_1\lambda_1^3\eta^2 + \delta_2\lambda_1^3\eta\right)\theta^2,$$

as $\tau_1 \to \infty$ and $\tau_2 \to \infty$.

MAIN RESULTS

The main results of this article are Berry-Esseen bounds for random-sum Wilcoxon's statistics. In Theorem 4, we consider a random-sum Wilcoxon's statistic with random centring and random norming as in the form $\left(W_{K_1,K_2} - U\right)/V$ where U and V are random variables. In Theorem 5, we consider a random-sum Wilcoxon's statistic with non-random centring and non-random norming as in the form $\left(W_{K_1,K_2} - a\right)/b$ where a and b are constants.

Theorem 4 *Let K_1 and K_2 be positive integer-valued random variables such that the X_i's, Y_j's, K_1, K_2 are independent. We have the following results:*

$$\sup_{x\in\mathbb{R}}\left|P\left(\frac{W_{K_1,K_2}-\theta K_1K_2}{\sqrt{\sigma_1^2K_1K_2^2+\sigma_2^2K_2K_1^2}}\leqslant x\right)-\Phi(x)\right|$$
$$\leqslant \frac{11.2}{\sqrt{EK_2}}\left\{\sqrt{\frac{\mathrm{Var}(K_1)}{EK_1}}\sqrt{\frac{EK_2}{EK_1}}+\sqrt{\frac{\mathrm{Var}(K_2)}{EK_2}}\right\}$$
$$+\frac{3.1\sqrt{\mu}\left(\sqrt{EK_2/EK_1}+\sqrt{EK_1/EK_2}\right)}{\sqrt{EK_2}\sqrt{\sigma_1^2+\sigma_2^2EK_1/EK_2}}$$
$$+\frac{11.9\left\{\gamma_1\sqrt{EK_2/EK_1}+\gamma_2\left(EK_1/EK_2\right)^{3/2}\right\}}{\sqrt{EK_2}\left(\sigma_1^2+\sigma_2^2EK_1/EK_2\right)^{3/2}},$$

and there exists a constant C such that for all $x\in\mathbb{R}$,

$$\left|P\left(\frac{W_{K_1,K_2}-\theta K_1K_2}{\sqrt{\sigma_1^2K_1K_2^2+\sigma_2^2K_2K_1^2}}\leqslant x\right)-\Phi(x)\right|$$
$$\leqslant \frac{C\,\mu}{(1+x)^2\,EK_2}\left\{\frac{\mathrm{Var}(K_1)}{\sigma_2^2EK_1}\frac{EK_2}{EK_1}+\frac{\mathrm{Var}(K_2)}{\sigma_1^2EK_2}\right\}$$
$$+\frac{C\,\mu\left(\sqrt{EK_2/EK_1}+\sqrt{EK_1/EK_2}\right)^2}{(1+x)^2\,EK_2\left(\sigma_1^2+\sigma_2^2EK_1/EK_2\right)}$$
$$+\frac{C\sqrt{\mu}\left(\sqrt{EK_2/EK_1}+\sqrt{EK_1/EK_2}\right)}{(1+x)^3\sqrt{EK_2}\sqrt{\sigma_1^2+\sigma_2^2EK_1/EK_2}}$$
$$+\frac{C\left\{\gamma_1\sqrt{EK_2/EK_1}+\gamma_2\left(EK_1/EK_2\right)^{3/2}\right\}}{(1+x)^3\sqrt{EK_2}\left(\sigma_1^2+\sigma_2^2EK_1/EK_2\right)^{3/2}}.$$

Remark 1 If K_1 and K_2 satisfy the conditions (C2), (C3) and (C4) in Theorem 3, then the rate of convergence is $O(\tau_2^{-1/2p})$ for some $p\geqslant 1$, when τ_1 and τ_2 tend to infinity.

By definition, random-sum Wilcoxon's statistics include Mann-Whitney's statistics in the special case of fixed-indices $K_1=m$ and $K_2=n$. In this case, our results are same as Theorem 2 (the bounds of Chen and Shao[8]). But the constants in a uniform bound of Theorem 2 are smaller. The next corollary is an immediately consequence of Theorem 4.

Corollary 1

$$\sup_{x\in\mathbb{R}}\left|P\left(\frac{W_{m,n}-mn\theta}{\sqrt{mn\left(n\sigma_1^2+m\sigma_2^2\right)}}\leqslant x\right)-\Phi(x)\right|$$
$$\leqslant \frac{3.1\sqrt{\mu}}{\sqrt{n}}\frac{\sqrt{n/m}+\sqrt{m/n}}{\sqrt{\sigma_1^2+\sigma_2^2m/n}}$$
$$+\frac{11.9}{\sqrt{n}}\frac{\gamma_1\sqrt{n/m}+\gamma_2\left(n/m\right)^{3/2}}{\left(\sigma_1^2+\sigma_2^2m/n\right)^{3/2}},$$

and there exists an absolute constant C (which does not depend on m and n) such that for all $x\in\mathbb{R}$,

$$\left|P\left(\frac{W_{m,n}-mn\theta}{\sqrt{mn\left(n\sigma_1^2+m\sigma_2^2\right)}}\leqslant x\right)-\Phi(x)\right|$$
$$\leqslant \frac{C\,\mu}{(1+x)^2\,n}\frac{\left(\sqrt{n/m}+\sqrt{m/n}\right)^2}{\left(\sigma_1^2+\sigma_2^2m/n\right)}$$
$$+\frac{C\sqrt{\mu}}{(1+x)^3\sqrt{n}}\frac{\sqrt{n/m}+\sqrt{m/n}}{\sqrt{\sigma_1^2+\sigma_2^2m/n}}$$
$$+\frac{C}{(1+x)^3\sqrt{n}}\frac{\gamma_1\sqrt{n/m}+\gamma_2\left(n/m\right)^{3/2}}{\left(\sigma_1^2+\sigma_2^2m/n\right)^{3/2}}.$$

It is understood that Wilcoxon's statistics can be approximated by the sum of independent random variables, introduced by Hajek[11]. Now, we denote the projection of W_{K_1,K_2} on the X_i's, Y_j's, K_1, K_2, by

$$\widehat{W}_{K_1,K_2}:=\sum_{i=1}^{K_1}E\left\{W_{K_1,K_2}-\theta EK_1EK_2|\,X_i\right\}$$
$$+\sum_{j=1}^{K_2}E\left\{W_{K_1,K_2}-\theta EK_1EK_2|\,Y_j\right\}$$
$$+E\left\{W_{K_1,K_2}-\theta EK_1EK_2|\,K_1,K_2\right\}$$
$$=(EK_2)\sum_{i=1}^{K_1}E\left\{I\left(X_i>Y_j\right)-\theta|\,X_i\right\}$$
$$+(EK_1)\sum_{j=1}^{K_2}E\left\{I\left(X_i>Y_j\right)-\theta|\,Y_j\right\}$$
$$+\theta\left(K_1K_2-EK_1EK_2\right).$$

By taking the conditional expectation given by K_1 and K_2 (see p. 22 of Ref. 12), we observe that

$$\mathrm{Var}(\widehat{W}_{K_1,K_2})$$
$$=E\left\{\mathrm{Var}\left(\widehat{W}_{K_1,K_2}\,|K_1,K_2\right)\right\}$$
$$+\mathrm{Var}\left\{E\left(\widehat{W}_{K_1,K_2}\,|K_1,K_2\right)\right\}$$
$$=\sigma_1^2EK_1\left(EK_2\right)^2+\sigma_2^2EK_2\left(EK_1\right)^2$$
$$+\theta^2\left\{\left(EK_2\right)^2\mathrm{Var}(K_1)+\left(EK_1\right)^2\mathrm{Var}(K_2)\right\}$$
$$+\theta^2\mathrm{Var}(K_1)\mathrm{Var}(K_2).$$

Set

$$\mathrm{Var}^*(\widehat{W}_{K_1,K_2}):=\mathrm{Var}(\widehat{W}_{K_1,K_2})-\theta^2\mathrm{Var}(K_1)\mathrm{Var}(K_2).$$

Similarly, we can see that

$$\begin{aligned}
&\mathrm{Var}(W_{K_1,K_2}) \\
&= E\left\{\mathrm{Var}(W_{K_1,K_2}|\,K_1,K_2)\right\} \\
&\quad + \mathrm{Var}\left\{E\left(W_{K_1,K_2}|\,K_1,K_2\right)\right\} \\
&= E\left\{\mu K_1K_2\right\} \\
&\quad + E\left\{\sigma_1^2K_1K_2\left(K_2-1\right)\right\} \\
&\quad + E\left\{\sigma_2^2K_1\left(K_1-1\right)K_2\right\} + \mathrm{Var}(\theta K_1K_2) \\
&= \mu EK_1EK_2 \\
&\quad + \sigma_1^2EK_1\left\{(EK_2)^2+\mathrm{Var}(K_2)-EK_2\right\} \\
&\quad + \sigma_2^2EK_2\left\{(EK_1)^2+\mathrm{Var}(K_1)-EK_1\right\} \\
&\quad + \theta^2\left\{(EK_2)^2\mathrm{Var}(K_1)+(EK_1)^2\mathrm{Var}(K_2)\right\} \\
&\quad + \theta^2\mathrm{Var}(K_1)\mathrm{Var}(K_2).
\end{aligned}$$

This implies that $\mathrm{Var}(W_{K_1,K_2})/\mathrm{Var}^*(\widehat{W}_{K_1,K_2}) \to 1$, when EK_1 and EK_2 tend to infinity. Hence, the quantity $\mathrm{Var}(W_{K_1,K_2})$ for normalization in the normal approximation theorem, can be replaced by $\mathrm{Var}^*(\widehat{W}_{K_1,K_2})$.

In the next statement, we show a uniform bound for random-sum Wilcoxon's statistics with non-random centring θEK_1EK_2 and non-random norming $\mathrm{Var}^*(\widehat{W}_{K_1,K_2})$.

Theorem 5 *Let $\xi_1,\dots,\xi_m$ be i.i.d. positive integer-valued random variables with $E(\xi_1)=a_1$, $\mathrm{Var}(\xi_1)=b_1^2$ and $E|\xi_1|^3<\infty$. Let $\zeta_1,\dots,\zeta_n$ be i.i.d. positive integer-valued random variables satisfying that $E(\zeta_1)=a_2$, $\mathrm{Var}(\zeta_1)=b_2^2$ and $E|\zeta_1|^3<\infty$. Suppose the X_i's, Y_j's, ξ_k's, ζ_ℓ's are independent. Put $K_1=\sum_{j=1}^m\xi_j$, $K_2=\sum_{\ell=1}^n\zeta_\ell$. Then*

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\left|P\left(\frac{W_{K_1,K_2}-\theta EK_1EK_2}{\sqrt{\mathrm{Var}^*(\widehat{W}_{K_1,K_2})}}\leqslant x\right)-\Phi(x)\right| \\
&\leqslant \frac{11.2}{\sqrt{n}}\left\{\frac{a_2b_1\tau}{a_1}+\frac{b_2}{a_2}\right\} \\
&+ \frac{3.1}{\sqrt{n}}\frac{\sqrt{\mu}\left(\sqrt{\tau/a_1}+\sqrt{a_1/a_2^2\tau}\right)}{\sqrt{\sigma_1^2+\sigma_2^2a_1\tau/a_2}} \\
&+ \frac{11.9}{\sqrt{n}}\frac{\gamma_1/\sqrt{a_1\tau}+\gamma_2\left(a_1\tau\right)^{3/2}/a_2^2}{\left(\sigma_1^2+\sigma_2^2a_1\tau/a_2\right)^{3/2}} \\
&+ \frac{1}{\sqrt{n}}\left\{\frac{2b_2}{a_2}+\frac{6.1E|\xi_1|^3}{\sqrt{\tau}\left(E\xi_1^2\right)^{3/2}}\right\} \\
&+ \frac{\sigma_1}{\theta\sqrt{n}}\left(\frac{1+\sqrt{2}}{\sqrt{a_1\tau}}+\frac{3b_2\sqrt{a_1}}{a_2b_1}\right. \\
&\qquad \left. +\frac{2b_2^2\sqrt{a_1}}{a_2^2b_1\sqrt{n}}+\frac{2b_1}{a_1\sqrt{\tau}}\right) \\
&+ \frac{\sigma_2}{\theta\sqrt{n}}\left\{\frac{16E|\xi_1|^3}{a_1b_1^2\sqrt{a_2m}}+\frac{2}{\sqrt{a_2}}+\frac{2a_1b_2}{a_2b_1\sqrt{a_2\tau}}\right\} \\
&+ \frac{\sigma_2}{\theta\sqrt{n}}\left(\frac{2b_1}{a_1\sqrt{a_2n}}+\frac{2\sqrt{2}}{b_1\sqrt{a_2}}\right. \\
&\qquad \left. +\frac{4\sqrt{2}b_1^2}{ma_1^2\sqrt{a_2}}+\frac{2\sqrt{2}}{ma_1\sqrt{a_2}}\right) \\
&+ \frac{1}{\sqrt{n}}\left\{\frac{6.1E|\zeta_1|^3}{b_2^3}+\frac{2b_1}{a_1\sqrt{\tau}}\right\},
\end{aligned}$$

where $\tau=m/n$.

Remark 2 The random variables $K_1=\sum_{j=1}^m\xi_j$ and $K_2=\sum_{\ell=1}^n\zeta_\ell$ depend on the parameters m and n, respectively. It is easy to see that K_1, K_2 satisfy the conditions (C1), (C3) and (C4) in Theorem 3. Moreover, if m/n converges to some constant when m and n tend to infinity, then K_1, K_2 satisfy the condition (C2). Hence the rate of convergence is $O\left(n^{-1/2}\right)$.

APPLICATION IN LROC ANALYSIS

A location receiver operating characteristic (LROC) curve, introduced by Starr, Metz, Lusted and Goodenoughet[13], provides a useful method in radiology based on a location of data. And the area under the LROC curve has been widely used to measure the diagnostic accuracy of imaging study because it represents the probability that positive and negative case are correctly classified. By using an empirical estimation model, Tang and Balakrishnan[10] showed that the area under the LROC curve is equivalent to the random-sum Wilcoxon's statistic as the following:

Let K be a binomial distributed random variables with parameters m and p. Suppose that the X_i's, Y_j's and K are all independent. An *estimator for the area under the LROC curve* is given by

$$A_{\mathrm{LROC}}=\frac{1}{m}\frac{1}{n}\sum_{i=1}^{K}\sum_{j=1}^{n}I\left(X_i\geqslant Y_j\right).$$

Tang and Balakrishnan[10] also obtained the asymptotic distribution of A_{LROC} as the following statement.

Theorem 6 (Ref. 10) *Let n be any fixed-index and K be a binomial distributed random variable with parameters m and p. Suppose that $m/n\to\tau$ for some constant $\tau>0$ as $m\to\infty$ and $n\to\infty$. Under* *the null hypothesis that the distributions of X and Y are equal, we have*

$$\left|P\left(\frac{\sqrt{n}\,(A_{\mathrm{LROC}}-p/2)}{m\sqrt{p\,(4n-3np+mp)\,/12}}\leqslant x\right)-\Phi(x)\right|$$

tends to zero as $m\to\infty$ and $n\to\infty$.

In the next theorem, we investigate Berry-Esseen bounds for A_{LROC}.

Theorem 7 *Let n be any fixed-index and K be a binomial distributed random variable with parameters m and p. Set $\tau=m/n$. Suppose that the X_i's, Y_j's and K are independent. Under the null hypothesis that the distributions of X and Y are equal, we have the following results:*

$$\sup_{x\in\mathbb{R}}\left|P\left(\frac{A_{\mathrm{LROC}}-K/2m}{mn\sqrt{nK\,(n+K)\,/12}}\leqslant x\right)-\Phi(x)\right|
\leqslant\frac{11.2}{\sqrt{n}}\frac{\sqrt{1-p}}{\sqrt{p\tau}}+\frac{5.4}{\sqrt{n}}\frac{\sqrt{1/p\tau}+\sqrt{p\tau}}{\sqrt{1+p\tau}}
+\frac{15.5}{\sqrt{n}}\frac{\sqrt{1/p\tau}+(1/p\tau)^{3/2}}{(1+p\tau)^{3/2}},$$

and there exists an absolute constant C (which does not depend on m, n and p) such that for all $x\in\mathbb{R}$,

$$\left|P\left(\frac{A_{\mathrm{LROC}}-K/2m}{mn\sqrt{nK\,(n+K)\,/12}}\leqslant x\right)-\Phi(x)\right|
\leqslant\frac{C}{(1+x)^2\,n}\left\{\frac{1-p}{p\tau}+\frac{\left(\sqrt{1/p\tau}+\sqrt{p\tau}\right)^2}{1+p\tau}\right\}
+\frac{C}{(1+x)^3\sqrt{n}}\left(\frac{\sqrt{1/p\tau}+\sqrt{p\tau}}{\sqrt{1+p\tau}}
+\frac{\sqrt{1/p\tau}+(1/p\tau)^{3/2}}{(1+p\tau)^{3/2}}\right).$$

Proof: Under the assumption that the distributions of X and Y are equal, Alberink[4] showed that $\theta=1/2$, $\sigma_1^2=\sigma_2^2=1/12$, and $\gamma_1=\gamma_2=1/32$. From these facts, the bounds in Theorem 4 can be simplified to Theorem 7. □

PROOF OF MAIN RESULTS

In the proof of Theorem 4, we apply Theorem 2 (Berry-Esseen theorem for Mann-Whitney's statistics) to investigate the bounds for the randomly-indexed summands. In the proof of Theorem 5, we obtain the bounds for random summands with non-random centring and non-random norming by improving the arguments of Chen, Goldstein and Shao (see p. 271 of Ref. 14) for the case of two random indices.

Proof of Theorem 4: (Uniform bound.) By using the conditional probability given by K_1, K_2 and applying the uniform bound of Theorem 2, we can see that

$$\begin{aligned}
&\left|P\left(\frac{W_{K_1,K_2}-\theta K_1K_2}{\sqrt{\sigma_1^2K_1K_2^2+\sigma_2^2K_2K_1^2}}\leqslant x\right)-\Phi(x)\right|\\
&\leqslant P\left(|K_1-EK_1|>0.09EK_1\right)\\
&\quad+\sum_{\{0.91EK_1\leqslant k_1\leqslant 1.01EK_1\}}P\left(K_1=k_1\right)\\
&\qquad\times\left|P\left(\frac{W_{k_1,K_2}-k_1\theta K_2}{\sqrt{k_1\sigma_1^2K_2^2+k_1^2\sigma_2^2K_2}}\leqslant x\right)-\Phi(x)\right|\\
&\leqslant P\left(|K_1-EK_1|>0.09EK_1\right)\\
&\quad+P\left(|K_2-EK_2|>0.09EK_2\right)\\
&\quad+\sum_{\substack{\{0.91EK_1\leqslant k_1\leqslant 1.09EK_1\}\\ \{0.91EK_2\leqslant k_2\leqslant 1.09EK_2\}}}P\left(K_1=k_1\right)P\left(K_2=k_2\right)\\
&\qquad\times\left|P\left(\frac{W_{k_1,k_2}-k_1k_2\theta}{\sqrt{k_1k_2^2\sigma_1^2+k_2k_1^2\sigma_2^2}}\leqslant x\right)-\Phi(x)\right|\\
&\leqslant 11.2\left\{\frac{\sqrt{\mathrm{Var}(K_1)}}{EK_1}+\frac{\sqrt{\mathrm{Var}(K_2)}}{EK_2}\right\}\\
&\quad+\sum_{\substack{\{0.91EK_1\leqslant k_1\leqslant 1.09EK_1\}\\ \{0.91EK_2\leqslant k_2\leqslant 1.09EK_2\}}}P\left(K_1=k_1\right)P\left(K_2=k_2\right)\\
&\quad\times\frac{1}{\sqrt{k_2}}\left\{\frac{(1+\sqrt{2})\sqrt{\mu}\left(\sqrt{k_2/k_1}+\sqrt{k_1/k_2}\right)}{\sqrt{\sigma_1^2+\sigma_2^2k_1/k_2}}\right.\\
&\qquad\left.+\frac{6.6\left(\gamma_1\sqrt{k_2/k_1}+\gamma_2\left(k_1/k_2\right)^{3/2}\right)}{\left(\sigma_1^2+k_1\sigma_2^2/k_2\right)^{3/2}}\right\}\\
&\leqslant\frac{11.2}{\sqrt{EK_2}}\left\{\sqrt{\frac{\mathrm{Var}(K_1)}{EK_1}}\sqrt{\frac{EK_2}{EK_1}}+\sqrt{\frac{\mathrm{Var}(K_2)}{EK_2}}\right\}\\
&\quad+\frac{3.1\sqrt{\mu}\left(\sqrt{EK_2/EK_1}+\sqrt{EK_1/EK_2}\right)}{\sqrt{EK_2}\sqrt{\sigma_1^2+\sigma_2^2EK_1/EK_2}}\\
&\quad+\frac{11.9\left\{\gamma_1\sqrt{EK_2/EK_1}+\gamma_2\left(EK_1/EK_2\right)^{3/2}\right\}}{\sqrt{EK_2}\left(\sigma_1^2+\sigma_2^2EK_1/EK_2\right)^{3/2}},
\end{aligned}\tag{1}$$

where we used the facts that for $\ell=1,2$,

$$0.91EK_\ell\leqslant k_\ell\leqslant 1.09EK_\ell$$

and that

$$\left(\sigma_1^2+\sigma_2^2 EK_1/EK_2\right) \leqslant \tfrac{109}{91}\left(\sigma_1^2+k_1\sigma_2^2/k_2\right)$$

in the last inequality.

(Non-uniform bound.) For all $k_1, k_2 \in \mathbb{N}$, put $p_{\overline{k}} = P(K_1=k_1)\,P(K_2=k_2)$ where $\overline{k}=(k_1,k_2)$ and set

$$\overline{\mathsf{A}} := \left\{(k_1,k_2)\in\mathbb{N}\times\mathbb{N}\,\middle|\,\begin{matrix}0.5EK_1\leqslant k_1\leqslant 1.5EK_1\\ 0.5EK_2\leqslant k_2\leqslant 1.5EK_2\end{matrix}\right\}.$$

From (1), it suffices to investigate the non-uniform bound for $|x|>1$. Without loss of generality, we may assume that $x>1$.

By using the non-uniform bound of Theorem 2 and taking the conditional probability given by K_1 and K_2, we have

$$\begin{aligned}
&\left|P\left(\frac{W_{K_1,K_2}-\theta K_1K_2}{\sqrt{\sigma_1^2K_1K_2^2+\sigma_2^2K_2K_1^2}}\leqslant x\right)-\Phi(x)\right|\\
&\leqslant \sum_{\overline{k}\notin\overline{\mathsf{A}}} p_{\overline{k}}\left|P\left(\frac{W_{k_1,k_2}-k_1k_2\theta}{\sqrt{k_1k_2^2\sigma_1^2+k_2k_1^2\sigma_2^2}}\leqslant x\right)-\Phi(x)\right|\\
&\quad+\frac{C\,\mu}{(1+x)^2}\sum_{\overline{k}\in\overline{\mathsf{A}}} p_{\overline{k}}\frac{\left(\sqrt{k_2/k_1}+\sqrt{k_1/k_2}\right)^2}{k_2\left(\sigma_1^2+\sigma_2^2k_1/k_2\right)}\\
&\quad+\frac{C\sqrt{\mu}}{(1+x)^3}\sum_{\overline{k}\in\overline{\mathsf{A}}} p_{\overline{k}}\frac{\sqrt{k_2/k_1}+\sqrt{k_1/k_2}}{\sqrt{k_2}\sqrt{\sigma_1^2+\sigma_2^2k_1/k_2}}\\
&\quad+\frac{C}{(1+x)^3}\sum_{\overline{k}\in\overline{\mathsf{A}}} p_{\overline{k}}\frac{\gamma_1\sqrt{k_2/k_1}+\gamma_2\left(k_1/k_2\right)^{3/2}}{\sqrt{k_2}\left(\sigma_1^2+\sigma_2^2k_2/k_1\right)^{3/2}}\\
&\leqslant \sum_{\overline{k}\notin\overline{\mathsf{A}}} p_{\overline{k}}\left|P\left(\frac{W_{k_1,k_2}-k_1k_2\theta}{\sqrt{k_1k_2^2\sigma_1^2+k_2k_1^2\sigma_2^2}}\leqslant x\right)-\Phi(x)\right|\\
&\quad+\frac{C\,\mu\left(\sqrt{EK_2/EK_1}+\sqrt{EK_1/EK_2}\right)^2}{(1+x)^2\,EK_2\left(\sigma_1^2+\sigma_2^2EK_1/EK_2\right)}\\
&\quad+\frac{C\sqrt{\mu}\left(\sqrt{EK_2/EK_1}+\sqrt{EK_1/EK_2}\right)}{(1+x)^3\sqrt{EK_2}\sqrt{\sigma_1^2+\sigma_2^2EK_1/EK_2}}\\
&\quad+\frac{C\left\{\gamma_1\sqrt{EK_2/EK_1}+\gamma_2\left(EK_1/EK_2\right)^{3/2}\right\}}{(1+x)^3\sqrt{EK_2}\left(\sigma_1^2+\sigma_2^2EK_1/EK_2\right)^{3/2}}.
\end{aligned}\tag{2}$$

Set

$$\widehat{W}_{\overline{k}} := \sum_{i=1}^{k_1} \mathrm{g}_1(X_i) + \sum_{j=1}^{k_2}\mathrm{g}_2(Y_j),$$

where

$$\begin{aligned}\mathrm{g}_1(X_i) &:= E\left(W_{k_1,k_2}-k_1k_2\theta\,|\,X_i\right)\\ &= k_2\left\{E\left\{I\left(X>Y\right)|\,X=X_i\right\}-\theta\right\},\end{aligned}$$

and

$$\begin{aligned}\mathrm{g}_2(Y_j) &:= E\left(W_{k_1,k_2}-k_1k_2\theta\,|\,Y_j\right)\\ &= k_1\left\{E\left\{I\left(X>Y\right)|\,Y=Y_j\right\}-\theta\right\}.\end{aligned}$$

Note that $E(\widehat{W}_{\overline{k}})=0$ and

$$\mathrm{Var}(\widehat{W}_{\overline{k}}) = k_1k_2^2\sigma_1^2+k_2k_1^2\sigma_2^2.$$

Set

$$\Delta_{\overline{k}} := W_{k_1,k_2}-k_1k_2\theta-\widehat{W}_{\overline{k}}.$$

Hence we can check that

$$\Delta_{\overline{k}} = \sum_{i=1}^{k_1}\sum_{j=1}^{k_2}\left\{I\left(X_j,Y_j\right)-\theta-\frac{\mathrm{g}_1(X_i)}{k_2}-\frac{\mathrm{g}_2(Y_j)}{k_1}\right\},$$

and that

$$E\left|\Delta_{\overline{k}}\right|^2 = k_1k_2\left\{\mu-\sigma_1^2-\sigma_2^2\right\}.$$

Observe that

$$\begin{aligned}
&\left|P\left(\frac{W_{k_1,k_2}-k_1k_2\theta}{\sqrt{k_1k_2^2\sigma_1^2+k_2k_1^2\sigma_2^2}}\leqslant x\right)-\Phi(x)\right|\\
&\leqslant P\left(\frac{\widehat{W}_{\overline{k}}}{\sqrt{k_1k_2^2\sigma_1^2+k_2k_1^2\sigma_2^2}}>\frac{x-2}{3}\right)\\
&\quad+P\left(\frac{\left|\Delta_{\overline{k}}\right|}{\sqrt{k_1k_2^2\sigma_1^2+k_2k_1^2\sigma_2^2}}>\frac{1+x}{3}\right)\\
&\quad+|1-\Phi(x)|\\
&\leqslant \frac{C}{(1+x)^2}\left\{1+\frac{E\left|\widehat{W}_{\overline{k}}\right|^2+E\left|\Delta_{\overline{k}}\right|^2}{k_1k_2\left(k_2\sigma_1^2+k_1\sigma_2^2\right)}\right\}\\
&\leqslant \frac{C}{(1+x)^2}\left\{1+\frac{\mu}{k_2\sigma_1^2+k_1\sigma_2^2}\right\}.
\end{aligned}\tag{3}$$

From (3) and the fact that $\sigma_\ell^2\leqslant\mu$ for $\ell=1,2$, we can see that

$$\begin{aligned}
&\sum_{\overline{k}\notin\overline{\mathsf{A}}} p_{\overline{k}}\left|P\left(\frac{W_{k_1,k_2}-k_1k_2\theta}{\sqrt{k_1k_2^2\sigma_1^2+k_2k_1^2\sigma_2^2}}\leqslant x\right)-\Phi(x)\right|\\
&\leqslant \frac{C}{(1+x)^2}\sum_{\{|k_1-EK_1|>0.5EK_1\}}P\left(K_1=k_1\right)\left(1+\frac{\mu}{k_1\sigma_2^2}\right)\\
&\quad+\frac{C}{(1+x)^2}\sum_{\{|k_2-EK_2|>0.5EK_2\}}P\left(K_2=k_2\right)\left(1+\frac{\mu}{k_2\sigma_1^2}\right)\\
&\leqslant \frac{C\mu}{(1+x)^2\sigma_2^2}P\left(|K_1-EK_1|>0.5EK_1\right)\\
&\quad+\frac{C\mu}{(1+x)^2\sigma_1^2}P\left(|K_2-EK_2|>0.5EK_2\right)
\end{aligned}$$

$$\leqslant \frac{C\,\mu}{(1+x)^2\,EK_2}\left\{\frac{\mathrm{Var}(K_1)}{\sigma_2^2 EK_1}\frac{EK_2}{EK_1}+\frac{\mathrm{Var}(K_2)}{\sigma_1^2 EK_2}\right\}. \tag{4}$$

Hence, combining (2) and (4), one can prove the non-uniform bound of Theorem 4. □

Proof of Theorem 5: Let Z_1, Z_2 and Z_3 be independent standard normal random variables which are independent of the X_i's, Y_j's, K_1, K_2. Set

$$T := \frac{W_{K_1,K_2}-\theta EK_1 EK_2}{\sqrt{\mathrm{Var}^*(\widehat{W}_{K_1,K_2})}},$$

$$T_1 := \frac{\theta\,(K_1K_2 - EK_1EK_2) + Z_1 Q}{\sqrt{\mathrm{Var}^*(\widehat{W}_{K_1,K_2})}},$$

$$T_2 := \frac{\theta(K_2-EK_2)EK_1 + Z_2\theta K_2\sqrt{\mathrm{Var}(K_1)} + Z_1 q}{\sqrt{\mathrm{Var}^*(\widehat{W}_{K_1,K_2})}},$$

$$T_3 := \frac{Z_3\theta EK_1\sqrt{\mathrm{Var}(K_2)} + Z_2\theta EK_2\sqrt{\mathrm{Var}(K_1)} + Z_1 q}{\sqrt{\mathrm{Var}^*(\widehat{W}_{K_1,K_2})}},$$

where

$$Q := \sqrt{\sigma_1^2 K_1 K_2^2 + \sigma_2^2 K_2 K_1^2},$$

$$q := \sqrt{\sigma_1^2 EK_1 EK_2^2 + \sigma_2^2 EK_2 EK_1^2}.$$

It is clear that T_3 has the standard normal distribution and easy to see that

$$\begin{aligned}
&\left|P\left(\frac{W_{K_1,K_2}-\theta EK_1EK_2}{\sqrt{\mathrm{Var}^*(\widehat{W}_{K_1,K_2})}}\leqslant x\right)-\Phi(x)\right| \\
&\leqslant |P(T\leqslant x) - P(T_1 \leqslant x)| \\
&\quad + |P(T_1\leqslant x) - P(T_2 \leqslant x)| \\
&\quad + |P(T_2\leqslant x) - P(T_3 \leqslant x)| \\
&= \left|P\left(\frac{W_{K_1,K_2}-\theta K_1K_2}{Q}\leqslant y_1\right)-\Phi(y_1)\right| \\
&\quad + \left|P\left(\frac{K_1-EK_1}{\sqrt{\mathrm{Var}(K_1)}} + \frac{Z_1(Q-q)}{\theta K_2\sqrt{\mathrm{Var}(K_1)}}\leqslant y_2\right)-\Phi(y_2)\right| \\
&\quad + \left|P\left(\frac{K_2-EK_2}{\sqrt{\mathrm{Var}(K_2)}} + \frac{Z_2(K_2-EK_2)\sqrt{\mathrm{Var}(K_1)}}{EK_1\sqrt{\mathrm{Var}(K_2)}}\leqslant y_3\right)-\Phi(y_3)\right| \\
&=: R_1 + R_2 + R_3,
\end{aligned} \tag{5}$$

where y_1, y_2 and y_3 are given by

$$y_1 := \frac{x\sqrt{\mathrm{Var}^*(\widehat{W}_{K_1,K_2})} - \theta\,(K_1K_2 - EK_1EK_2)}{Q},$$

$$y_2 := \frac{x\sqrt{\mathrm{Var}^*(\widehat{W}_{K_1,K_2})} - \theta\,(K_2 - EK_2)\,EK_1 - Z_1 q}{\theta K_2\sqrt{\mathrm{Var}(K_1)}},$$

$$y_3 := \frac{x\sqrt{\mathrm{Var}(\widehat{W}_{\mathrm{r}})} - Z_2\theta EK_2\sqrt{\mathrm{Var}(K_1)} - Z_1 q}{\theta EK_1\sqrt{\mathrm{Var}(K_2)}}.$$

Firstly, following the arguments of (1) (in the proof of Theorem 4), we can see that

$$\begin{aligned}
R_1 &\leqslant \frac{11.2}{\sqrt{EK_2}}\left\{\sqrt{\frac{\mathrm{Var}(K_1)}{EK_1}}\sqrt{\frac{EK_2}{EK_1}}+\sqrt{\frac{\mathrm{Var}(K_2)}{EK_2}}\right\} \\
&\quad + \frac{3.1\sqrt{\mu}\left(\sqrt{EK_2/EK_1}+\sqrt{EK_1/EK_2}\right)}{\sqrt{EK_2}\sqrt{\sigma_1^2+\sigma_2^2 EK_1/EK_2}} \\
&\quad + \frac{11.9\left\{\gamma_1\sqrt{EK_2/EK_1}+\gamma_2\,(EK_1/EK_2)^{3/2}\right\}}{\sqrt{EK_2}\,(\sigma_1^2+\sigma_2^2EK_1/EK_2)^{3/2}} \\
&= \frac{11.2}{\sqrt{n}}\left\{\frac{a_2b_1\tau}{a_1}+\frac{b_2}{a_2}\right\} \\
&\quad + \frac{3.1}{\sqrt{n}}\frac{\sqrt{\mu}\left(\sqrt{\tau/a_1}+\sqrt{a_1/a_2^2\tau}\right)}{\sqrt{\sigma_1^2+\sigma_2^2a_1\tau/a_2}} \\
&\quad + \frac{11.9}{\sqrt{n}}\frac{\gamma_1/\sqrt{a_1\tau}+\gamma_2\,(a_1\tau)^{3/2}/a_2^2}{(\sigma_1^2+\sigma_2^2a_1\tau/a_2)^{3/2}}.
\end{aligned} \tag{6}$$

Secondly, using the conditional probability given by K_2 and Z_1,

$$\begin{aligned}
R_2 &\leqslant P\left(|K_2-EK_2|>0.5EK_2\right) \\
&\quad + \sum_{\{0.5EK_2\leqslant k_2\leqslant 1.5EK_2\}} P(K_2=k_2) \\
&\qquad \times |P(S_1+\Delta\leqslant y_2)-\Phi(y_2)|,
\end{aligned} \tag{7}$$

where

$$S_1 = \frac{K_1-EK_1}{\sqrt{\mathrm{Var}(K_1)}},$$

and

$$\Delta = \frac{Z_1\left(\sqrt{k_2^2\sigma_1^2K_1+k_2\sigma_2^2K_1^2}-q\right)}{k_2\theta\sqrt{\mathrm{Var}(K_1)}}.$$

For each $i=1,\ldots,m$, let

$$\Delta_i = \frac{Z_1\left(\sqrt{k_2^2\sigma_1^2\epsilon_i+k_2\sigma_2^2\epsilon_i^2}-q\right)}{k_2\theta\sqrt{\mathrm{Var}(K_1)}},$$

where $\epsilon_i = K_1 - \xi_i + a_1$. Note that

$$
\begin{aligned}
&E\left|S_1\Delta\right| \\
&\leqslant \frac{E\left|S_1\left\{k_2^2\sigma_1^2K_1 + k_2\sigma_2^2K_1^2 - q^2\right\}\right|}{qk_2\theta\sqrt{\mathrm{Var}(K_1)}} \\
&\leqslant \frac{\sigma_1^2E\left|S_1\left\{k_2^2(K_1-EK_1)+k_2(k_2-EK_2)EK_1\right\}\right|}{qk_2\theta\sqrt{\mathrm{Var}(K_1)}} \\
&\quad + \frac{\sigma_1^2E\left|S_1\left\{(k_2-EK_2)EK_1EK_2+\left(E^2K_2-EK_2^2\right)EK_1\right\}\right|}{qk_2\theta\sqrt{\mathrm{Var}(K_1)}} \\
&\quad + \frac{\sigma_2^2E\left|S_1\left\{k_2(K_1-EK_1)^2+2k_2(K_1-EK_1)EK_1\right\}\right|}{qk_2\theta\sqrt{\mathrm{Var}(K_1)}} \\
&\quad + \frac{\sigma_2^2E\left|S_1\left\{(k_2-EK_2)E^2K_1+\left(E^2K_1-EK_1^2\right)EK_2\right\}\right|}{qk_2\theta\sqrt{\mathrm{Var}(K_1)}} \\
&\leqslant \frac{\sigma_1\left\{k_2^2\sqrt{\mathrm{Var}(K_1)} + k_2\left|k_2 - EK_2\right|EK_1\right\}}{k_2\theta EK_2\sqrt{EK_1}\sqrt{\mathrm{Var}(K_1)}} \\
&\quad + \frac{\sigma_1\left\{\left|k_2 - EK_2\right|EK_1EK_2 + EK_1\mathrm{Var}(K_2)\right\}}{k_2\theta EK_2\sqrt{EK_1}\sqrt{\mathrm{Var}(K_1)}} \\
&\quad + \frac{\sigma_2\left\{k_2E\left|K_1 - EK_1\right|^3 + 2k_2EK_1\sqrt{\mathrm{Var}(K_1)}\right\}}{k_2\theta EK_1\sqrt{EK_2}\sqrt{\mathrm{Var}(K_1)}} \\
&\quad + \frac{\sigma_2\left\{\left|k_2 - EK_2\right|E^2K_1 + EK_2\mathrm{Var}(K_1)\right\}}{k_2\theta EK_1\sqrt{EK_2}\sqrt{\mathrm{Var}(K_1)}} \\
&\leqslant \frac{\sigma_1}{\theta}\left\{\frac{k_2}{a_2n\sqrt{a_1m}} + \frac{\left|k_2 - EK_2\right|\sqrt{a_1}}{a_2nb_1}\right\} \\
&\quad + \frac{\sigma_1}{\theta}\left\{\frac{\left|k_2 - EK_2\right|\sqrt{a_1}}{k_2b_1} + \frac{b_2^2\sqrt{a_1}}{k_2a_2b_1}\right\} \\
&\quad + \frac{\sigma_2}{\theta}\left\{\frac{16E\left|\xi_1\right|^3}{a_1b_1^2\sqrt{a_2mn}} + \frac{2}{\sqrt{a_2n}}\right\} \\
&\quad + \frac{\sigma_2}{\theta}\left\{\frac{\left|k_2 - EK_2\right|a_1\sqrt{m}}{k_2b_1\sqrt{a_2n}} + \frac{b_1\sqrt{a_2}}{k_2a_1}\right\}, \qquad (8)
\end{aligned}
$$

where we used the fact that (Rosenthal's inequality)

$$
\begin{aligned}
&E\left|K_1 - EK_1\right|^3 \\
&\qquad \leqslant 8\left\{\textstyle\sum_{i=1}^m E\left|\xi_i\right|^3 + \left(\sum_{i=1}^m E\xi_i^2\right)^{3/2}\right\}
\end{aligned}
$$

in the last inequality.

Also, we can see that

$$
\begin{aligned}
&E\left|(\xi_i - a_1)(\Delta - \Delta_i)\right| \\
&\leqslant E\left\{\frac{\left|\xi_i - a_1\right|\left|k_2^2\sigma_1^2(\xi_i - a_1) + k_2\sigma_2^2\left\{K_1^2 - \epsilon_i^2\right\}\right|}{k_2\theta\sqrt{\mathrm{Var}(K_1)}\sqrt{k_2^2\sigma_1^2\epsilon_i + k_2\sigma_2^2\epsilon_i^2}}\right\} \\
&\leqslant E\left\{\frac{\sigma_1(\xi_i - a_1)^2}{\theta\sqrt{\epsilon_i}\sqrt{\mathrm{Var}(K_1)}}\right\} \\
&\quad + E\left\{\frac{\sigma_2\left|\xi_i - a_1\right|\left|2\epsilon_i + (\xi_i - a)\right|}{\theta\sqrt{k_2}\left|\epsilon_i\right|\sqrt{\mathrm{Var}(K_1)}}\right\} \\
&\leqslant E\left\{\frac{\sigma_1(\xi_i - a_1)^2}{\theta\sqrt{\epsilon_i}\sqrt{\mathrm{Var}(K_1)}}\right\} \\
&\quad + E\left\{\frac{2\sigma_2\left|\xi_i - a_1\right|}{\theta\sqrt{k_2}\sqrt{\mathrm{Var}(K_1)}} + \frac{\sigma_2(\xi_i - a)^2}{\theta\sqrt{k_2}\left|\epsilon_i\right|\sqrt{\mathrm{Var}(K_1)}}\right\} \\
&\leqslant \frac{\sigma_1b_1}{\theta\sqrt{m}}\left(\frac{2b_1}{a_1\sqrt{m}} + \frac{\sqrt{2}}{\sqrt{a_1m}}\right) \\
&\quad + \frac{2\sigma_2}{\theta\sqrt{k_2m}} + \frac{\sigma_2b_1}{\theta\sqrt{k_2m}}\left(\frac{4b_1^2}{a_1^2m} + \frac{2}{a_1m}\right), \qquad (9)
\end{aligned}
$$

where we use the fact that

$$
\begin{aligned}
&E\epsilon_i^{-1} \\
&\leqslant P\left(\left|\epsilon_i - EK_1\right| > 0.5EK_1\right) \\
&\quad + E\left\{\frac{I\left(\left|\epsilon_i - EK_1\right| \leqslant 0.5EK_1\right)}{\epsilon_i}\right\} \\
&\leqslant \frac{4\mathrm{Var}(K_1)}{(EK_1)^2} + \frac{2}{EK_1} \\
&= \frac{4b_1^2}{a_1^2m} + \frac{2}{a_1m}
\end{aligned}
$$

in the last inequality.

By applying the result of p. 258 of Ref. 14 (the uniform bound for nonlinear statistics), we have that

$$
\begin{aligned}
&\sup_{z\in\mathbb{R}}\left|P\left(S_1 + \Delta \leqslant z\right) - \Phi(z)\right| \\
&\qquad \leqslant \frac{6.1E\left|\xi\right|^3}{\sqrt{m}\left(E\xi^2\right)^{3/2}} + E\left|S_1\Delta\right| \\
&\qquad\quad + \sum_{i=1}^m \frac{E\left|(\xi_i - a_1)(\Delta - \Delta_i)\right|}{b_1\sqrt{m}}. \qquad (10)
\end{aligned}
$$

Hence combining (7)–(10), we obtain

$$
\begin{aligned}
R_2 &\leqslant P\left(\left|K_2 - EK_2\right| \geqslant 0.5EK_2\right) \\
&\quad + \sum_{\{0.5EK_2\leqslant k_2\leqslant 1.5EK_2\}} P\left(K_2 = k_2\right) \\
&\quad \times \left|P\left(S_1 + \Delta \leqslant y_2\right) - \Phi(y_2)\right| \\
&\leqslant \frac{2\sqrt{\mathrm{Var}(K_2)}}{EK_2} + \frac{6.1E\left|\xi\right|^3}{\sqrt{m}\left(E\xi^2\right)^{3/2}} \\
&\quad + \sum_{\{0.5EK_2\leqslant k_2\leqslant 1.5EK_2\}} P\left(K_2 = k_2\right) \\
&\quad \times \left\{\frac{\sigma_1}{\theta}\left(\frac{k_2}{a_2n\sqrt{a_1m}} + \frac{\left|k_2 - EK_2\right|\sqrt{a_1}}{a_2nb_1}\right)\right. \\
&\qquad + \frac{\sigma_1}{\theta}\left(\frac{\left|k_2 - EK_2\right|\sqrt{a_1}}{k_2b_1} + \frac{b_2^2\sqrt{a_1}}{k_2a_2b_1}\right) \\
&\qquad + \frac{\sigma_2}{\theta}\left(\frac{16E\left|\xi_1\right|^3}{a_1b_1^2\sqrt{a_2mn}} + \frac{2}{\sqrt{a_2n}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{\sigma_2}{\theta}\left(\frac{|k_2-EK_2|\,a_1\sqrt{m}}{k_2b_1\sqrt{a_2n}}+\frac{b_1\sqrt{a_2}}{k_2a_1}\right)\\
&+\frac{\sigma_1}{\theta}\left(\frac{2b_1}{a_1\sqrt{m}}+\frac{\sqrt{2}}{\sqrt{a_1m}}\right)\\
&+\frac{2\sigma_2}{\theta b_1\sqrt{k_2}}+\frac{\sigma_2}{\theta\sqrt{k_2}}\left(\frac{4b_1^2}{a_1^2m}+\frac{2}{a_1m}\right)\Bigg\}\\
\leqslant{}& \frac{2\sqrt{\mathrm{Var}(K_2)}}{EK_2}+\frac{6.1E\,|\xi_1|^3}{\sqrt{m}\,(E\xi_1^2)^{3/2}}\\
&+\frac{\sigma_1}{\theta}\left\{\frac{EK_2}{a_2n\sqrt{a_1m}}+\frac{\sqrt{\mathrm{Var}(K_2)}\sqrt{a_1}}{a_2nb_1}\right\}\\
&+\frac{\sigma_1}{\theta}\left\{\frac{2\sqrt{\mathrm{Var}(K_2)}\sqrt{a_1}}{b_1EK_2}+\frac{2b_2^2\sqrt{a_1}}{a_2b_1EK_2}\right\}\\
&+\frac{\sigma_2}{\theta}\left\{\frac{16E\,|\xi_1|^3}{a_1b_1^2\sqrt{a_2nm}}+\frac{2}{\sqrt{a_2n}}\right\}\\
&+\frac{\sigma_2}{\theta}\left\{\frac{2a_1\sqrt{m}\sqrt{\mathrm{Var}(K_2)}}{b_1\sqrt{a_2n}EK_2}+\frac{2b_1\sqrt{a_2}}{a_1EK_2}\right\}\\
&+\frac{\sigma_1}{\theta}\left\{\frac{2b_1}{a_1\sqrt{m}}+\frac{\sqrt{2}}{\sqrt{a_1m}}\right\}\\
&+\frac{2\sqrt{2}\sigma_2}{\theta b_1\sqrt{EK_2}}+\frac{\sqrt{2}\sigma_2}{\theta\sqrt{EK_2}}\left\{\frac{4b_1^2}{a_1^2m}+\frac{2}{a_1m}\right\}\\
={}&\frac{1}{\sqrt{n}}\left\{\frac{2b_2}{a_2}+\frac{6.1E\,|\xi_1|^3}{\sqrt{\tau}\,(E\xi_1^2)^{3/2}}\right\}\\
&+\frac{\sigma_1}{\theta\sqrt{n}}\left\{\frac{1+\sqrt{2}}{\sqrt{a_1\tau}}+\frac{3b_2\sqrt{a_1}}{a_2b_1}\right\}\\
&+\frac{\sigma_1}{\theta\sqrt{n}}\left\{\frac{2b_2^2\sqrt{a_1}}{a_2^2b_1\sqrt{n}}+\frac{2b_1}{a_1\sqrt{\tau}}\right\}\\
&+\frac{\sigma_2}{\theta\sqrt{n}}\left\{\frac{16E\,|\xi_1|^3}{a_1b_1^2\sqrt{a_2m}}+\frac{2}{\sqrt{a_2}}\right\}\\
&+\frac{\sigma_2}{\theta\sqrt{n}}\left\{\frac{2a_1b_2}{a_2b_1\sqrt{a_2\tau}}+\frac{2b_1}{a_1\sqrt{a_2n}}\right\}\\
&+\frac{\sigma_2}{\theta\sqrt{n}}\left\{\frac{2\sqrt{2}}{b_1\sqrt{a_2}}+\frac{4\sqrt{2}b_1^2}{ma_1^2\sqrt{a_2}}+\frac{2\sqrt{2}}{ma_1\sqrt{a_2}}\right\}. \qquad (11)
\end{aligned}
$$

Finally, we denote

$$S_2=\frac{K_2-EK_2}{\sqrt{\mathrm{Var}(K_2)}},$$

and set

$$\Lambda=\frac{(K_2-EK_2)\sqrt{\mathrm{Var}(K_1)}}{EK_1\sqrt{\mathrm{Var}(K_2)}}Z_2.$$

For each $j=1,\dots,n$, let

$$\Lambda_j=\frac{\{(K_2-\zeta_j+a_2)-EK_2\}\sqrt{\mathrm{Var}(K_1)}}{EK_1\sqrt{\mathrm{Var}(K_2)}}Z_2.$$

Note that

$$
\begin{aligned}
E\,|S_2\Lambda| &\leqslant \frac{\sqrt{ES_2^2}\sqrt{E\,|K_2-EK_2|^2}\sqrt{\mathrm{Var}(K_1)}}{EK_1\sqrt{\mathrm{Var}(K_2)}}E\,|Z_2|\\
&\leqslant \frac{b_1}{a_1\sqrt{m}},
\end{aligned}
$$

and that

$$
\begin{aligned}
&E\,|(\zeta_j-a_2)(\Lambda-\Lambda_j)|\\
&\leqslant \frac{E(\zeta_j-a_2)^2\sqrt{\mathrm{Var}(K_1)}}{EK_1\sqrt{\mathrm{Var}(K_2)}}E\,|Z_2|\\
&\leqslant \frac{b_1b_2}{a_1\sqrt{mn}}.
\end{aligned}
$$

By using the conditional probability given by Z_2, Z_1 and applying the result of p. 258 of Ref. 14, we have

$$
\begin{aligned}
R_3 &= |P(S_2+\Lambda\leqslant y_3)-\Phi(y_3)|\\
&\leqslant \frac{6.1E|\zeta_1|^3}{\sqrt{n}(E\zeta_1^2)^{3/2}}+E\,|S_2\Lambda|+\sum_{j=1}^{n}\frac{E|(\zeta_j-a_2)(\Lambda-\Lambda_j)|}{b_2\sqrt{n}}\\
&\leqslant \frac{6.1E\,|\zeta_1|^3}{\sqrt{n}\,(E\zeta_1^2)^{3/2}}+\frac{2b_1}{a_1\sqrt{m}}\\
&= \frac{1}{\sqrt{n}}\left\{\frac{6.1E\,|\zeta_1|^3}{(E\zeta_1^2)^{3/2}}+\frac{2b_1}{a_1\sqrt{\tau}}\right\}. \qquad (12)
\end{aligned}
$$

We combine (5), (6), (11) and (12) to obtain the main result. □

Acknowledgements: The authors thank Professor Kritsana Neammanee for helpful discussions and the Centre of Excellence in Mathematics CHE for financial support.

REFERENCES

1. Wilcoxon F (1945) Individual comparisons by ranking methods. *Biometrics* **1**, 80–3.
2. Mann HB, Whitney DR (1947) On a test of whether one of two random variables is stochastically larger than the other. *Ann Math Stat* **18**, 50–60.
3. Pestman WR (1998) *Mathematical Statistics: An Introduction*, De Gruyter.
4. Alberink IB (2000) A Berry-Esseen bound for U-statistics in the non-i.i.d. case. *J Theor Probab* **13**, 519–33.
5. Hoeffding W (1948) A class of statistics with asymptotically normal distribution. *Ann Math Stat* **19**, 293–325.

6. Lehmann EL (1951) Consistency and unbiasedness of certain nonparametric tests. *Ann Math Stat* **22**, 165–79.
7. Grams WF, Serfling RJ (1973) Convergence rates for U-statistics and related statistics. *Ann Probab* **1**, 153–60.
8. Chen LHY, Shao QM (2007) Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli* **13**, 581–99.
9. Chakraborty DP (2006) A search model and figure of merit for observer data acquired according to the free-response paradigm. *Phys Med Biol* **51**, 3449–62.
10. Tang LL, Balakrishnan N (2011) A random-sum Wilcoxon statistic and its applications to analysis of ROC and LROC data. *J Stat Plann Infer* **141**, 335–44.
11. Hajek J (1968) Asymptotic normality of simple linear rank statistics under alternatives. *Ann Math Stat* **39**, 325–46.
12. Ross S (1995) *Stochastic Processes*, 2nd edn, Wiley, New York.
13. Starr SJ, Metz CE, Lust edn LB, Goodenough DJ (1975) Visual detection and localization of radiographic images. *Radiology* **116**, 533–8.
14. Chen LHY, Goldstein L, Shao QM (2010) *Normal Approximation by Stein's Method*, Springer, New York.