Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: www.tandfonline.com/journals/tstf20

Asymptotic distributions of degenerated U-statistics

Jun Shao

To cite this article: Jun Shao (2025) Asymptotic distributions of degenerated U-statistics, Statistical Theory and Related Fields, 9:4, 434-443, DOI: 10.1080/24754269.2025.2579414

To link to this article: https://doi.org/10.1080/24754269.2025.2579414

STATISTICAL THEORY AND RELATED FIELDS
2025, VOL. 9, NO. 4, 434–443
https://doi.org/10.1080/24754269.2025.2579414

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS   Check for updates

Asymptotic distributions of degenerated U-statistics

Jun Shao

Department of Statistics, University of Wisconsin, Madison, WI, USA

ABSTRACT

A degenerated U-statistic with rank 2 is known to converge in distribution to a weighted sum of independent chi-square random variables. However, a result for the asymptotic distribution of a degenerated U-statistic with rank higher than 2 is not available in the literature. We derive explicitly the asymptotic distribution of a degenerated U-statistic with any rank, which is useful for hypothesis testing when the test statistic is a degenerated U-statistic under null hypothesis. Interestingly, the limit for a degenerated U-statistic with rank higher than 2 is a weighted sum of independent polynomials of the standard normal random variables.

ARTICLE HISTORY

Received 12 June 2025
Accepted 18 October 2025

KEYWORDS

Degenerated U-statistics; orthonormal eigenfunctions; rank of U-statistics; sample mean and sample product; V-statistics

1. Introduction

U-statistics are popular for unbiased estimation and statistical testing (Hoeffding, 1948). Let k be a fixed positive integer and $h(Y_1,\dots,Y_k)$ be a known function symmetric in Y_j's with $E\{h^2(Y_1,\dots,Y_k)\} < \infty$, where $Y_1, Y_2, \dots$ are independent and identically distributed random vectors. A U-statistic with kernel h of order k based on data $Y_1,\dots,Y_n$ is of the form

$$U_{k,n} = \frac{1}{\binom{n}{k}} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,n\}} h(Y_{i_1},\dots,Y_{i_k}), \tag{1}$$

where the summation is over all $\binom{n}{k}$ combinations of k distinct integers $i_1,\dots,i_k$ from $1,\dots,n$. By symmetry, the U-statistic $U_{k,n}$ in (1) is an unbiased estimator of $\theta = E\{h(Y_1,\dots,Y_k)\}$, which is an unknown parameter depending on the population of Y_1. In fact, when the order statistics of univariate $Y_1,\dots,Y_n$ are sufficient and complete, $U_{k,n}$ is the uniformly minimum variance unbiased estimator of θ (Serfling, 1980; Shao, 2003).

Several examples are given as follows.

Example 1.1: For univariate Y_i, the sample kth order product

$$\frac{1}{\binom{n}{k}} \sum_{\{i_1,\dots,i_k\}\subset\{1,\dots,n\}} Y_{i_1}\cdots Y_{i_k} \tag{2}$$

is a U-statistic in (1) with product kernel $h(Y_1,\dots,Y_k) = Y_1\cdots Y_k$ and is an unbiased estimator of $\theta = \{E(Y_1)\}^k$; in the special case of $k = 1$, (2) is the popular sample mean of $Y_1,\dots,Y_n$.

CONTACT Jun Shao jshao@wisc.edu Department of Statistics, University of Wisconsin, Madison, WI 53706, USA

Example 1.2: For univariate Y_i, Gini's mean difference $\frac{2}{n(n-1)} \sum_{\{i,j\}\subset\{1,\ldots,n\}} |Y_i - Y_j|$ is a U-statistic in (1) with $h(Y_1, Y_2) = |Y_1 - Y_2|$ and is an unbiased estimator of the concentration measure $\theta = E|Y_1 - Y_2|$.

Example 1.3: Let $I(B)$ denote the indicator of event B. For univariate Y_i, the one-sample Wilcoxon statistic $\frac{2}{n(n-1)} \sum_{\{i,j\}\subset\{1,\ldots,n\}} I(Y_i + Y_j \le 0)$ is a U-statistic in (1) with $h(Y_1, Y_2) = I(Y_1 + Y_2 \le 0)$ and is an unbiased estimator of $\theta = P(Y_1 + Y_2 \le 0)$.

Example 1.4: Examples of U-statistics in (1) for vector Y_i and $\theta =$ a measure of dependence of random variables are given in Section 4.

Define

$$\begin{aligned} \psi(y_1) &= E\{h(y_1, Y_2, \ldots, Y_k)\}, \\ \psi(y_1, y_2) &= E\{h(y_1, y_2, Y_3, \ldots, Y_k)\}, \\ &\vdots \\ \psi(y_1, y_2, \ldots, y_j) &= E\left\{h(y_1, y_2, \ldots, y_j, Y_{j+1}, \ldots, Y_k)\right\}, \end{aligned} \tag{3}$$

for any integer j between 1 and k, where the expectation E is with respect to $Y_{j+1}, \ldots, Y_k$. The U-statistic $U_{k,n}$ in (1) is *degenerated* if and only if $\psi(y_1)$ in (3) is constant. It can be shown that if $\psi(y_1, y_2, \ldots, y_j)$ in (3) is constant, then so is $\psi(y_1, y_2, \ldots, y_{j'})$ with any integer $j' < j$. Thus, the smallest integer ℓ with non-constant $\psi(y_1, y_2, \ldots, y_\ell)$ is called the rank of $U_{k,n}$ in (1) and $U_{k,n}$ is degenerated if and only if its rank is higher than 1. For instance, the sample kth order product in (2) has $\psi(y_1, \ldots, y_j) = y_1 \cdots y_j E(Y_{j+1} \cdots Y_k) = y_1 \cdots y_j \{E(Y_1)\}^{k-j}$, which is 0 if and only if $E(Y_1) = 0$ and $k \ge 2$; hence, the rank of sample kth order product in (2) is $\ell = 1$ when $E(Y_1) \ne 0$ and $\ell = k$ when $E(Y_1) = 0$. The one-sample Wilcoxon statistic and Gini's mean difference are both of rank 1 and non-degenerated. More examples of degenerated U-statistic are given in Section 4.

It is well known that, for a non-degenerated U-statistic $U_{k,n}$, its convergence rate to $\theta = E(U_{k,n})$ is $n^{-1/2}$ as $n \to \infty$, and $n^{1/2}(U_{k,n} - \theta)$ converges in distribution to a normal random variable with mean 0 and variance $k^2 \mathrm{Var}\{\psi(Y_1)\} > 0$ (Hoeffding, 1948; Serfling, 1980). For a degenerated U-statistic $U_{k,n}$ with rank $\ell = 2$, its convergence rate to $\theta = E(U_{k,n})$ is n^{-1} and $n(U_{k,n} - \theta)$ converges in distribution to a random variable with mean 0 and variance $\{k^2(k-1)^2/2\}\mathrm{Var}\{\psi(Y_1, Y_2)\} > 0$, which is proved in detail in Section 5.5.2 of Serfling (1980); in fact, Serfling (1980) shows that the limit random variable is a weighted (possibly infinite) sum of independent chi-square random variables.

For a degenerated U-statistic $U_{k,n}$ with rank $\ell \ge 3$, however, a general result for the asymptotic distribution of $U_{k,n}$ is not available. For example, what is the asymptotic distribution of the sample kth order product in (2) when $k \ge 3$ and $E(Y_1) = 0$?

The purpose of this paper is to fill in this gap by establishing the asymptotic distribution of $U_{k,n}$ in (1) with rank $\ell \ge 3$. In Section 2, we derive the asymptotic distribution of the sample kth order product in (2), which is the distribution of an explicitly given kth order polynomial of a standard normal random variable. This result sets up a base stone for general degenerated U-statistics. In Section 3, we show that, for $U_{k,n}$ in (1) with rank $\ell \ge 3$, $n^{\ell/2}(U_{k,n} - \theta)$ converges in distribution to a (possibly infinite) weighted sum of independent order ℓ polynomials of the standard normal random variable. The result interestingly extends the result in Serfling (1980) for $\ell = 2$ since a chi-square random variable is an order 2 polynomial of the standard normal random variable. Our results have important applications in hypothesis testing, since a U-statistic with rank $\ell \geq 2$ is often encountered under the null hypothesis of interest (Lai et al., 2021; Serfling, 1980; Zhu et al., 2012). Section 4 provides examples of U-statistics for measuring dependence of random variables. Section 5 gives some concluding remarks.

2. Asymptotic distribution of sample *k*th order product

In this section we show that the sample kth order product in (2) converges in distribution to a kth order polynomial of a standard normal random variable.

Theorem 2.1: *Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables with $E(X_i) = 0$ and $E(X_i^2) = \sigma^2$. For any $\ell = 1, 2, \ldots$,*

$$\frac{n^{\ell/2}}{\binom{n}{\ell}} \sum_{\{i_1,\ldots,i_\ell\} \subset \{1,\ldots,n\}} X_{i_1} \cdots X_{i_\ell} \xrightarrow{d} \sigma^\ell p_\ell(Z), \tag{4}$$

where $\xrightarrow{d}$ denotes convergence in distribution as $n \to \infty$, Z denotes a standard normal random variable, $p_\ell(Z)$ is a polynomial of order ℓ given recursively by

$$p_\ell(Z) = Zp_{\ell-1}(Z) - (\ell - 1)p_{\ell-2}(Z),\ p_1(Z) = Z,\ p_0(Z) = 1, \tag{5}$$

$E\{p_\ell(Z)\} = 0$, *and* $\mathrm{Var}\{p_\ell(Z)\} = \ell!$.

It is interesting to know that $p_2(Z) = Z^2 - 1$, $p_3(Z) = Z^3 - 3Z$, and $p_4(Z) = Z^4 - 6Z^2 + 3$. The result in (4) with $\ell = 1$ or 2 is well known.

Before presenting the proof of Theorem 1, we compare the asymptotic efficiency between the sample kth order product in (2) and the simple estimator $\overline{Y}^k$ as estimators of $\theta = \{E(Y_1)\}^k$, where $\overline{Y}$ is the sample mean of $Y_1, \ldots, Y_n$. When $E(Y_1) = 0$ and $E(Y_1^2) = \sigma^2$, an immediate consequence of Theorem 1 is that $n^{k/2}$ (the sample kth order product)/$\sigma^k \xrightarrow{d} p_k(Z)$ with $k \geq 2$. On the other hand, $n^{k/2}\overline{Y}^k/\sigma^k \xrightarrow{d} Z^k$ and, thus, the asymptotic mean squared error of the sample kth order product in (2) over that of $\overline{Y}^k$ is

$$\frac{k!}{E(Z^{2k})} = \frac{k(k-1)\cdots 2 \cdot 1}{(2k-1)(2k-3)\cdots 3 \cdot 1} = \frac{k}{2k-1}\frac{k-1}{2k-3}\cdots\frac{2}{3}\frac{1}{1} < 1,$$

which is 2/3 when $k = 2$ and 2/5 when $k = 3$. Thus, the sample kth order product in (2) is asymptotically more efficient than $\overline{Y}^k$ when $E(Y_1) = 0$ or is nearly 0.

Proof of Theorem 1: Result (4) with $\ell = 2$ is actually shown in Section 5.5.2 of Serfling (1980) with $p_2(Z) = Z^2 - 1$. Without loss of generality, we assume that $\sigma = 1$ in the proof. Let $\overline{X}$ be the sample mean of $X_1, \ldots, X_n$ and $\overline{X}_2$ be the sample mean of $X_1^2, \ldots, X_n^2$.

We first prove (4) with $\ell = 3$. From the identity

$$(n-1)\left\{\frac{2}{n(n-1)}\sum_{\{i,j\}\subset\{1,\ldots,n\}} X_iX_j\right\} = n\overline{X}^2 - \overline{X}_2, \tag{6}$$

multiplying $n^{1/2}\overline{X}$ to each side we obtain that

$$\frac{2}{n^{3/2}}\sum_{\{i,j\}\subset\{1,\ldots,n\}} X_iX_j\sum_{l=1}^{n} X_l = n^{3/2}\overline{X}^3 - n^{1/2}\overline{X}\overline{X}_2,$$

which is the same as

$$\frac{3!}{n^{3/2}}\sum_{\{i,j,l\}\subset\{1,\ldots,n\}} X_iX_jX_l = n^{3/2}\overline{X}^3 - n^{1/2}\overline{X}\overline{X}_2 - \frac{2}{n^{3/2}}\sum_{\{i,j\}\subset\{1,\ldots,n\}}(X_i^2X_j + X_iX_j^2).$$

From $n^{1/2}(\overline{X}_2 - 1) = O_p(1)$ and $n^{1/2}\overline{X} = O_p(1)$, where, throughout, $O_p(a_n)$ denotes a term bounded by a_n in probability, and from the fact that $\frac{2}{n(n-1)}\sum_{\{i,j\}\subset\{1,\ldots,n\}}(X_i^2X_j + X_iX_j^2)$ is a non-degenerated U-statistic that equals $2\overline{X} + O_p(n^{-1})$ (Serfling, 1980), we obtain that

$$\begin{aligned}
\frac{n^{3/2}}{\binom{n}{3}}\sum_{\{i,j,l\}\subset\{1,\ldots,n\}} X_iX_jX_l &= n^{3/2}\overline{X}^3 - n^{1/2}\overline{X} - 2n^{1/2}\overline{X} + O_p(n^{-1/2}) \\
&= n^{1/2}\overline{X}p_2(n^{1/2}\overline{X}) - 2p_1(n^{1/2}\overline{X}) + O_p(n^{-1/2}) \\
&= p_3(n^{1/2}\overline{X}) + O_p(n^{-1/2})
\end{aligned} \tag{7}$$

with $p_3(Z) = Zp_2(Z) - 2p_1(Z) = Z^3 - 3Z$ as given in (5). Since $n^{1/2}\overline{X} \xrightarrow{d} Z$, result (7) implies result (4) with $\ell = 3$, noting that $E\{p_3(Z)\} = E(Z^3 - 3Z) = 0$ and $\mathrm{Var}\{p_3(Z)\} = E(Z^3 - 3Z)^2 = E(Z^6 - 6Z^4 + 9Z^2) = 6 = 3!$, and using the fact that $E(Z^d) = 0$ for any odd integer d and $E(Z^d) = (d-1)(d-3)\cdots 3\cdot 1$ for any even integer d.

We next prove (4) with $\ell = 4$. From (7), multiplying $n^{1/2}\overline{X}$ to each side we obtain that

$$\begin{aligned}
&\frac{n^2}{\binom{n}{4}}\sum_{\{i,j,l,r\}\subset\{1,\ldots,n\}} X_iX_jX_lX_r \\
&\quad = n^{1/2}\overline{X}p_3(n^{1/2}\overline{X}) \\
&\qquad - \frac{3!}{n^2}\sum_{\{i,j,l\}\subset\{1,\ldots,n\}}(X_iX_jX_l^2 + X_iX_j^2X_l + X_i^2X_jX_l) + O_p(n^{-1/2}) \\
&\quad = n^{1/2}\overline{X}p_3(n^{1/2}\overline{X}) - 3\left(\frac{2}{n}\sum_{i<j} X_iX_j\right) + O_p(n^{-1/2}) \\
&\quad = n^{1/2}\overline{X}p_3(n^{1/2}\overline{X}) - 3(n\overline{X}^2 - \overline{X}_2) + O_p(n^{-1/2}) \\
&\quad = n^{1/2}\overline{X}p_3(n^{1/2}\overline{X}) - 3p_2(n^{1/2}\overline{X}) + O_p(n^{-1/2}) \\
&\quad = p_4(n^{1/2}\overline{X}) + O_p(n^{-1/2}),
\end{aligned} \tag{8}$$

where the second equality holds since $\frac{3!}{n(n-1)(n-2)}\sum_{\{i,j,l\}\subset\{1,\ldots,n\}}(X_iX_jX_l^2+X_iX_j^2X_l+X_i^2X_jX_l)$ is a U-statistic (of order 3 and rank 2) that equals $\frac{2}{n(n-1)}\sum_{i<j}X_iX_j+O_p(n^{-3/2})$ (Serfling, 1980), the third equality follows from (6), the fourth equality follows from $n^{1/2}(\overline{X}_2-1)=O_p(1)$ and $p_2(Z)=Z^2-1$, and the last equality follows from $Zp_3(Z)-3p_2(Z)=p_4(Z)$ by (5). Therefore, (4) holds for $\ell=4$ with $p_4(Z)=Z^4-6Z^2+3$, noting that $E\{p_4(Z)\}=E(Z^4)-6E(Z^2)+3=0$ and $\mathrm{Var}\{p_4(Z)\}=E(Z^4-6Z^2+3)^2=E(Z^8+36Z^4+9-12Z^6+6Z^4-36Z^2)=24=4!$.

Following the argument in the proof for $\ell=3,4$, for a general $\ell\geq 5$, we multiply $n^{1/2}\overline{X}$ to the quantities in the identity for the case of $\ell-1$, e.g., (6)–(8) for $\ell=2,3,4$, to obtain

$$\begin{aligned}&\frac{n^{\ell/2}}{\binom{n}{\ell}}\sum_{\{i_1,\ldots,i_\ell\}\subset\{1,\ldots,n\}}X_{i_1}\cdots X_{i_\ell}\\&\quad=n^{1/2}\overline{X}p_{\ell-1}(n^{1/2}\overline{X})-(\ell-1)p_{\ell-2}(n^{1/2}\overline{X})+O_p(n^{-1/2})\\&\quad=p_\ell(n^{1/2}\overline{X})+O_p(n^{-1/2}),\end{aligned}\tag{9}$$

where the last equality follows from (5). Thus, the convergence $\xrightarrow{d}$ in (4) holds.

To finish the proof, it remains to show that $E\{p_\ell(Z)\}=0$ and $\mathrm{Var}\{p_\ell(Z)\}=\ell!$ for $\ell=5,6,\ldots$. Since the mean and variance of the left side of (9) are 0 and $\ell!\{n(n-1)\cdots(n-\ell+1)\}^{-1}$, respectively, by (9), we just need to show that $\{p_\ell^2(n^{1/2}\overline{X}),n=1,2,\ldots\}$ is uniformly integrable (Serfling, 1980; Shao, 2003), which is true since $\{(n^{1/2}\overline{X})^2,n=1,2,\ldots\}$ is uniformly integrable (Serfling, 1980; Shao, 2003) and $p_\ell^2(n^{1/2}\overline{X})$ is a polynomial of $n^{1/2}\overline{X}$ with order 2ℓ. This completes the proof. ■

3. Asymptotic distributions of general degenerated U-statistics

Based on Theorem 1 in Section 2, we establish the following general result.

Theorem 3.1: *Assume that $U_{k,n}$ in (1) is degenerated with rank ℓ between 2 and k. There are nonrandom $\lambda_{\ell,t}$, $t=1,2,\ldots$, such that $\sum_{t=1}^\infty\lambda_{\ell,t}^2=\mathrm{Var}\{\psi(Y_1,\ldots,Y_\ell)\}>0$, where $\psi(y_1,\ldots,y_\ell)$ is defined in (3), and*

$$n^{\ell/2}(U_{k,n}-\theta)\xrightarrow{d}\binom{k}{\ell}\sum_{t=1}^{\infty}\lambda_{\ell,t}p_\ell(Z_t),\tag{10}$$

where $p_\ell(\cdot)$ is the polynomial of order ℓ given in (5) and $Z_1,Z_2,\ldots$ are independent standard normal random variables. Furthermore, the limit in the right side of (10) has mean 0 and variance equal to $\lim_{n\to\infty}\mathrm{Var}(n^{\ell/2}U_{k,n})=\ell!\binom{k}{\ell}^2\mathrm{Var}\{\psi(Y_1,\ldots,Y_\ell)\}$.

Before presenting the proof of Theorem 2, we discuss about the relative efficiency between the U-statistic (1) and its corresponding V-statistic

$$V_{k,n}=\frac{1}{n^k}\sum_{i_1=1}^{n}\cdots\sum_{i_k=1}^{n}h(Y_{i_1},\ldots,Y_{i_k}).$$

For non-degenerated U-statistics (rank$=1$), it is known that $n^{1/2}(U_{k,n}-\theta)$ and $n^{1/2}(V_{k,n}-\theta)$ have the same asymptotic distribution, provided that $E\{h(Y_{i_1},\ldots,Y_{i_k})\}<\infty$ for all $1\leq$

$i_1 \leq \cdots \leq i_k \leq k$; see, e.g., Section 3.5.3 of Shao (2003). For degenerated U-statistics (rank $\ell \geq 2$), the asymptotic distribution for $V_{k,n}$ can be derived using the same technique in the proof of Theorem 2. However, $U_{k,n}$ is asymptotically more efficient than $V_{k,n}$ in terms of the asymptotic mean squared error. The special case of $\ell = 2$ is discussed in detail in Section 3.5.3 of Shao (2003). For $\ell \geq 3$, a similar result can be obtained.

Proof of Theorem 2: Because $\psi(y_1, \ldots, y_j)$ is constant for $j = 1, \ldots, \ell - 1$, from formula (2) on Page 190 of Serfling (1980),

$$n^{\ell/2}(U_{k,n} - \theta) = n^{\ell/2}\binom{k}{\ell}\widetilde{U}_{\ell,n} + O_p(n^{-1/2}),$$

where $\widetilde{U}_{\ell,n}$ is of the form of a U-statistic (1) with order ℓ and kernel function $\psi(Y_1, \ldots, Y_\ell) - \theta$. Let $\{\phi_t(\cdot), t = 1, 2, \ldots\}$ denote orthonormal eigenfunctions corresponding to the eigenvalues $\{\lambda_{\ell,t}, t = 1, 2, \ldots\}$ defined in connection with $\psi(y_1, \ldots, y_\ell) - \theta$ such that $\phi_1(Y_i), \phi_2(Y_i), \ldots$ are independent and identically distributed, $E\{\phi_t(Y_i)\} = 0$, $E\{\phi_t^2(Y_i)\} = 1$, and

$$\begin{aligned} &E\left\{\psi(Y_1, \ldots, Y_\ell) - \theta - \sum_{t=1}^{T} \lambda_{\ell,t}\phi_t(Y_1)\cdots\phi_t(Y_\ell)\right\}^2 \\ &\quad = E\{\psi(Y_1, \ldots, Y_\ell) - \theta\}^2 - \sum_{t=1}^{T} \lambda_{\ell,t}^2 \to 0 \end{aligned} \tag{11}$$

as $T \to \infty$ (Dunford & Schwartz, 1963; Serfling, 1980), which implies that

$$E\{\psi(Y_1, \ldots, Y_\ell) - \theta\}^2 = \mathrm{Var}\{\psi(Y_1, \ldots, Y_\ell)\} = \sum_{t=1}^{\infty} \lambda_{\ell,t}^2 < \infty. \tag{12}$$

Then, result (10) follows from

$$n^{\ell/2}\widetilde{U}_{\ell,n} \xrightarrow{d} W_\ell, \quad W_\ell = \sum_{t=1}^{\infty} \lambda_{\ell,t}\, p_\ell(Z_t). \tag{13}$$

Define

$$\widetilde{U}_{\ell,n}^{(T)} = \sum_{t=1}^{T} \frac{\lambda_{\ell,t}}{\binom{k}{\ell}} \sum_{\{i_1,\ldots,i_\ell\}\subset\{1,\ldots,n\}} \phi_t(Y_{i_1})\cdots\phi_t(Y_{i_\ell}).$$

Then, $\widetilde{U}_{\ell,n} - \widetilde{U}_{\ell,n}^{(T)}$ is of the form of a U-statistic with order ℓ and kernel

$$g(Y_1, \ldots, Y_\ell) = \psi(Y_1, \ldots, Y_\ell) - \theta - \sum_{t=1}^{T} \lambda_{\ell,t}\phi_t(Y_1)\cdots\phi_t(Y_\ell).$$

From the theory of U-statistic, e.g., Section 3.2 of Shao (2003),

$$E\left\{n^\ell\left(\widetilde{U}_{\ell n} - \widetilde{U}_{\ell,n}^{(T)}\right)^2\right\} = \ell!\mathrm{Var}\{g(Y_1, \ldots, Y_\ell)\} + O(n^{-1})$$

$$= \ell! \sum_{t=T+1}^{\infty} \lambda_{\ell,t}^2 + O(n^{-1}),$$

where $O(a_n)$ denotes a term bounded by a_n and the last equality follows from (11)–(12). Let $\mathrm{ch}(s, R)$ denote the characteristic function of random variable R. Then,

$$\begin{aligned} \left|\mathrm{ch}\left(s, n^{\ell/2}\widetilde{U}_{\ell,n}\right) - \mathrm{ch}\left(s, n^{\ell/2}\widetilde{U}_{\ell,n}^{(T)}\right)\right| &\le |s| \left[E\left\{n^{\ell}\left(\widetilde{U}_{\ell,n} - \widetilde{U}_{\ell,n}^{(T)}\right)^2\right\}\right]^{1/2} \\ &= |s| \left(\ell! \sum_{t=T+1}^{\infty} \lambda_{\ell,t}^2\right)^{1/2}. \end{aligned}$$

This and result (12) imply that, for any fixed s and $\epsilon > 0$,

$$\sup_n \left|\mathrm{ch}\left(s, n^{\ell/2}\widetilde{U}_{\ell,n}\right) - \mathrm{ch}\left(s, n^{\ell/2}\widetilde{U}_{\ell,n}^{(T)}\right)\right| < \epsilon$$

for all sufficiently large T. Applying Theorem 1 with X_t and $p_\ell(Z)$ replaced by $\phi_t(Y_i)$ and $p_\ell(Z_t)$, respectively, and using the fact that $\phi_t(Y_i)$, $t = 1, 2, \ldots$, $i = 1, \ldots, n$, are independent, we obtain that, for every $T = 1, 2, \ldots$,

$$\begin{aligned} n^{\ell/2}\widetilde{U}_{\ell,n}^{(T)} &= \frac{n^{\ell/2}}{\binom{n}{\ell}} \sum_{t=1}^{T} \lambda_{\ell,t} \sum_{\{i_1,\ldots,i_\ell\}\subset\{1,\ldots,n\}} \phi_t(Y_{i_1})\cdots\phi_t(Y_{i_\ell}) \xrightarrow{d} W_\ell^{(T)} \\ &= \sum_{t=1}^{T} \lambda_{\ell,t} p_\ell(Z_t). \end{aligned} \tag{14}$$

Hence, for any $s, \epsilon > 0$ and T,

$$\left|\mathrm{ch}\left(s, n^{\ell/2}\widetilde{U}_{\ell,n}^{(T)}\right) - \mathrm{ch}\left(s, W_\ell^{(T)}\right)\right| < \epsilon$$

for all sufficiently large n. For any s and $\epsilon > 0$,

$$\begin{aligned} \left|\mathrm{ch}\left(s, W_\ell^{(T)}\right) - \mathrm{ch}(s, W_\ell)\right| &\le |s| \left\{E\left(W_\ell - W_\ell^{(T)}\right)^2\right\}^{1/2} \\ &= |s| \left(\ell! \sum_{t=T+1}^{\infty} \lambda_{\ell,t}^2\right)^{1/2} < \epsilon \end{aligned} \tag{15}$$

for all sufficiently large T. Therefore, for any s and $\epsilon > 0$,

$$\left|\mathrm{ch}\left(s, n^{\ell/2}\widetilde{U}_{\ell,n}\right) - \mathrm{ch}(s, W_\ell)\right| < 3\epsilon$$

for all sufficiently large n. This shows that the characteristic function of $n^{\ell/2}\widetilde{U}_{\ell,n}$ converges to the characteristic function of W_ℓ and, thus, (13) holds and the proof of (10) is completed.

It remains to show that the mean of W_ℓ is 0 and the variance of W_ℓ is $\ell!\mathrm{Var}\{\psi(Y_1, \ldots, Y_\ell)\}$, for W_ℓ given by (13). Let $W_\ell^{(T)}$ be given by (14). Then, result (15) actually shows that $W_\ell^{(T)} \xrightarrow{d} W_\ell$ as $T \to \infty$. Since $\{(W_\ell^{(T)})^2, T = 1, 2, \ldots\}$ is uniformly integrable (e.g., Page 86 of Shao, 2003), $E(W_\ell) = \lim_{T\to\infty} E(W_\ell^{(T)}) = 0$ as $E\{p_\ell(Z_t)\} = 0$ and $\mathrm{Var}(W_\ell) = \lim_{T\to\infty} \mathrm{Var}(W_\ell^{(T)}) = \lim_{T\to\infty} \sum_{t=1}^{T} \lambda_{\ell,t}^2 \mathrm{Var}\{p_\ell(Z_t)\} = \ell! \lim_{T\to\infty} \sum_{t=1}^{T} \lambda_{\ell,t}^2 = \ell! \sum_{t=1}^{\infty} \lambda_{\ell,t}^2 = \ell! \mathrm{Var}\{\psi(Y_1,\ldots,Y_\ell)\}$ by (12). By Hoeffding (1948), $\mathrm{Var}(n^{\ell/2} U_{k,n}) \to \ell! \binom{k}{\ell}^2 \mathrm{Var}\{\psi(Y_1,\ldots,Y_\ell)\}$ as $n\to\infty$. This completes the proof. ■

4. U-statistics for measures of dependence

In this section, we provide details of Example 4 in Section 1 for U-statistics in (1) with multivariate Y_i's in estimating measures of dependence of random variables.

Consider two random variables R and S with joint cumulative distribution function $F_{R,S}(r,s)$ and marginal cumulative distribution functions $F_R(r)$ and $F_S(s)$ for R and S, respectively. A measure of dependence of R and S is

$$\theta = \iint \left\{F_{R,S}(r,s) - F_R(r)F_S(s)\right\}^2 \mathrm{d}F_{R,S}(r,s), \tag{16}$$

which is 0 if and only if $F_{R,S}(r,s) = F_R(r)F_S(s)$ for any r and s, i.e., R and S are independent.

Let $Y_i = (R_i, S_i)$, $i = 1,\ldots,n$, be independent and identically distributed with $F_{R,S}(r,s)$. As an unbiased estimator of θ in (16), the U-statistic $U_{5,n}$ of order $k = 5$ is given by (1) with

$$\begin{aligned} h(Y_1,\ldots,Y_5) = \frac{1}{5!}\sum_P \big\{ & I(R_{i_2} \le R_{i_1})I(S_{i_2} \le S_{i_1})I(R_{i_3} \le R_{i_1})I(S_{i_3} \le S_{i_1}) \\ & - 2I(R_{i_2} \le R_{i_1})I(S_{i_2} \le S_{i_1})I(R_{i_4} \le R_{i_1})I(S_{i_5} \le S_{i_1}) \\ & + I(R_{i_2} \le R_{i_1})I(S_{i_3} \le S_{i_1})I(R_{i_4} \le R_{i_1})I(S_{i_5} \le S_{i_1})\big\}, \end{aligned} \tag{17}$$

where the summation is over the 5! permutations $\{i_1,\ldots,i_5\}$ of $\{1,\ldots,5\}$ and $I(B)$ is the indicator of B. To see why $E\{h(Y_1,\ldots,Y_5)\} = \theta$ in (16), note that

$$\theta = \iint \left\{F_{R,S}^2(r,s) - 2\iint F_{R,S}(r,s)F_R(r)F_S(s) + \iint F_R^2(r)F_S^2(s)\right\} \mathrm{d}F_{R,S}(r,s)$$

and

$$\begin{aligned} & E\big\{I(R_{i_2} \le R_{i_1})I(S_{i_2} \le S_{i_1})I(R_{i_3} \le R_{i_1})I(S_{i_3} \le S_{i_1})\big\} \\ & \quad = E\big[E\big\{I(R_{i_2} \le R_{i_1})I(S_{i_2} \le S_{i_1})I(R_{i_3} \le R_{i_1})I(S_{i_3} \le S_{i_1}) \mid Y_{i_1}\big\}\big] \\ & \quad = \iint E\big\{I(R_{i_2} \le r)I(S_{i_2} \le s)I(R_{i_3} \le r)I(S_{i_3} \le s) \mid Y_{i_1} = (r,s)\big\}\mathrm{d}F_{R,S}(r,s) \\ & \quad = \iint E\big\{I(R_{i_2} \le r)I(S_{i_2} \le s)I(R_{i_3} \le r)I(S_{i_3} \le s)\big\}\mathrm{d}F_{R,S}(r,s) \\ & \quad = \iint E\big\{I(R_{i_2} \le r)I(S_{i_2} \le s)\big\}E\big\{I(R_{i_3} \le r)I(S_{i_3} \le s)\big\}\mathrm{d}F_{R,S}(r,s) \\ & \quad = \iint F_{R,S}^2(r,s)\mathrm{d}F_{R,S}(r,s), \end{aligned}$$

$$\begin{aligned} & E\big\{I(R_{i_2} \le R_{i_1})I(S_{i_2} \le S_{i_1})I(R_{i_4} \le R_{i_1})I(S_{i_5} \le S_{i_1})\big\} \\ & \quad = E\big[E\big\{I(R_{i_2} \le R_{i_1})I(S_{i_2} \le S_{i_1})I(R_{i_4} \le R_{i_1})I(S_{i_5} \le S_{i_1}) \mid Y_{i_1}\big\}\big] \end{aligned}$$

$$
\begin{aligned}
&= \iint E\big\{I(R_{i_2} \le r)I(S_{i_2} \le s)I(R_{i_4} \le r)I(S_{i_5} \le s) \mid Y_{i_1} = (r,s)\big\}\mathrm{d}F_{R,S}(r,s) \\
&= \iint E\big\{I(R_{i_2} \le r)I(S_{i_2} \le s)I(R_{i_4} \le r)I(S_{i_5} \le s)\big\}\mathrm{d}F_{R,S}(r,s) \\
&= \iint E\big\{I(R_{i_2} \le r)I(S_{i_2} \le s)\big\}E\big\{I(R_{i_4} \le r)\big\}E\big\{I(S_{i_5} \le s)\big\}\mathrm{d}F_{R,S}(r,s) \\
&= \iint F_{R,S}(r,s)F_R(r)F_S(s)\mathrm{d}F_{R,S}(r,s),
\end{aligned}
$$

$$
\begin{aligned}
&E\big\{I(R_{i_2} \le R_{i_1})I(S_{i_3} \le S_{i_1})I(R_{i_4} \le R_{i_1})I(S_{i_5} \le S_{i_1})\big\} \\
&\quad = E\big[E\big\{I(R_{i_2} \le R_{i_1})I(S_{i_3} \le S_{i_1})I(R_{i_4} \le R_{i_1})I(S_{i_5} \le S_{i_1}) \mid Y_{i_1}\big\}\big] \\
&\quad = \iint E\big\{I(R_{i_2} \le r)I(S_{i_3} \le s)I(R_{i_4} \le r)I(S_{i_5} \le s) \mid Y_{i_1} = (r,s)\big\}\mathrm{d}F_{R,S}(r,s) \\
&\quad = \iint E\big\{I(R_{i_2} \le r)I(S_{i_3} \le s)I(R_{i_4} \le r)I(S_{i_5} \le s)\big\}\mathrm{d}F_{R,S}(r,s) \\
&\quad = \iint E\big\{I(R_{i_2} \le r)\big\}E\big\{I(S_{i_3} \le s)\big\}E\big\{I(R_{i_4} \le r)\big\}E\big\{I(S_{i_5} \le s)\big\}\mathrm{d}F_{R,S}(r,s) \\
&\quad = \iint F_R^2(r)F_S^2(s)\mathrm{d}F_{R,S}(r,s),
\end{aligned}
$$

where we use the fact that Y_i and Y_j are independent whenever $i \neq j$.

From the form of kernel h in (17), we obtain that, with $y = (r_1, s_1)$,

$$
\begin{aligned}
\psi(y_1) &= E\{I(R_2 \le r)I(S_2 \le s)I(R_3 \le r)I(S_3 \le s)\} \\
&\quad - 2E\{I(R_2 \le r_1)I(S_2 \le s_1)I(R_4 \le r_1)I(S_5 \le s_1)\} \\
&\quad + E\{I(R_2 \le r_1)I(S_3 \le s_1)I(R_4 \le r_1)I(S_5 \le s_1)\} \\
&= F_{R,S}^2(r_1,s_1) - 2F_{R,S}(r_1,s_1)F_R(r_1)F_S(s_1) + F_R^2(r_1)F_S^2(s_1) \\
&= \{F_{R,S}(r_1,s_1) - F_R(r_1)F_S(s_1)\}^2,
\end{aligned}
$$

which is non-constant (the corresponding U-statistic has rank 1) if and only if $\theta \neq 0$ (R and S are not independent). When R and S are independent, with $y_1 = (r_1, s_1)$ and $y_2 = (r_2, s_2)$,

$$
\begin{aligned}
&\psi(y_1, y_2) \\
&\quad = \frac{1}{4}\int \{I(r_1 \le r)I(r_2 \le r) - I(r_1 \le r)F_R(r) - I(r_2 \le r)F_R(r) + F_R^2(r)\}\mathrm{d}F_R(r) \\
&\quad\quad \times \int \{I(s_1 \le s)I(s_2 \le s) - I(s_1 \le s)F_S(s) - I(s_2 \le s)F_S(s) + F_S^2(s)\}\mathrm{d}F_S(s),
\end{aligned}
$$

which is non-constant and, thus, the corresponding U-statistic has rank 2.

The previous discussion can be extended to the dependence of any fixed number of random variables. We end this paper by constructing a U-statistic for the dependence of three random variables, Q, R, and S. Let $Y = (Q, R, S)$, $F_Y(q, r, s)$ be the joint cumulative distribution of Y, and $F_Q(q)$, $F_R(r)$ and $F_S(s)$ be the marginal cumulative distributions of Q, R and S, respectively. A measure of dependence of Q, R and S is

$$\theta = \iiint \{F_Y(q,r,s) - F_Q(q)F_R(r)F_S(s)\}^2 \mathrm{d}F_Y(q,r,s),$$

which is 0 if and only if $F_Y(q,r,s) = F_Q(q)F_R(r)F_S(s)$ for any q, r and s, i.e., Q, R and S are independent. Using the same argument, we can construct a U-statistic (1) of order 7 and kernel $h(Y_1,\ldots,Y_7)$ equating

$$\begin{aligned}\frac{1}{7!}\sum_P \{ & I(Q_{i_2} \le Q_{i_1})I(R_{i_2} \le R_{i_1})I(S_{i_2} \le S_{i_1})I(Q_{i_3} \le Q_{i_1})I(R_{i_3} \le R_{i_1})I(S_{i_3} \le S_{i_1}) \\ & - 2I(Q_{i_2} \le Q_{i_1})I(R_{i_2} \le R_{i_1})I(S_{i_2} \le S_{i_1})I(Q_{i_5} \le Q_{i_1})I(R_{i_6} \le R_{i_1})I(S_{i_7} \le S_{i_1}) \\ & + I(Q_{i_2} \le Q_{i_1})I(R_{i_3} \le R_{i_1})I(S_{i_4} \le S_{i_1})I(Q_{i_5} \le Q_{i_1})I(R_{i_6} \le R_{i_1})I(S_{i_7} \le S_{i_1})\},\end{aligned}$$

where $Y_i = (Q_i, R_i, S_i)$ and the summation is over the 7! permutations $\{i_1,\ldots,i_7\}$ of $\{1,\ldots,7\}$.

5. Concluding remarks

This paper presents a rigorous and technically sophisticated derivation of the asymptotic distributions of degenerated U-statistics with any given rank ℓ, extending the existing result for rank $\ell = 2$ and offering new insights into the limiting behaviour of such statistics. This topic is of practical importance in statistical theory, since a U-statistic with rank $\ell \ge 2$ is often encountered under the null hypothesis of interest. Possible extensions include the derivations of asymptotic joint distributions of several U-statistics, some of which have high ranks, and multi-sample generalized U-statistics with high ranks (Serfling, 1980, p. 175).

Acknowledgments

The author would like to thank two anonymous referees for helpful comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Jun Shao http://orcid.org/0000-0002-0943-1179

References

Dunford, N., & Schwartz, J. (1963). *Linear operators*. Wiley.
Hoeffding, W. (1948). A class of statistics with asymptotic normal distribution. *Annals of Mathematical Statistics*, *19*(3), 293–325. https://doi.org/10.1214/aoms/1177730196
Lai, T., Zhang, Z., & Wang, Y. (2021). A kernel-based measure for conditional mean dependence. *Computational Statistics and Data Analysis*, *160*, Article 107246. https://doi.org/10.1016/j.csda.2021.107246
Serfling, R. (1980). *Approximation theorems of mathematical statistics*. John Wiley & Sons.
Shao, J. (2003). *Mathematical statistics* (2nd ed.). Springer.
Zhu, W., Jiang, Y., & Zhang, H. (2012). Nonparametric covariate-adjusted association tests based on the generalized Kendall's tau. *Journal of the American Statistical Association*, *107*(497), 1–11. https://doi.org/10.1080/01621459.2011.643707