Asymptotic normality for U-statistics of associated random variables

I. Dewan *, B.L.S. Prakasa Rao

Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi 110016, India

Abstract

Let $\{X_n,\ n \geqslant 1\}$ be a stationary sequence of associated random variables and U_n be a U-statistic based on this sample. We establish a central limit theorem for U_n when the U-statistic is degenerate or non-degenerate using an orthogonal expansion for the kernel associated with U_n. We extend the results to U-statistics of kernels of degree 3 and to V-statistics of arbitrary degree. We also establish a central limit theorem for the two sample U-statistic based on observations of two independent stationary associated sequences.

MSC: 60F05; 62M99

Keywords: U-statistics; Degenerate; Non-degenerate; Central limit theorem; Associated random variables

1. Introduction

Let $\{X_n,\ n \geqslant 1\}$ be a sequence of random variables. A finite family $\{X_1, \ldots, X_n\}$ of random variables is said to be associated if

$$\text{Cov}(h(X_1, \ldots, X_n), g(X_1, \ldots, X_n)) \geqslant 0$$

for any coordinatewise non-decreasing functions h, g on $\mathbb{R}^n$ such that the covariance exists. An infinite family of random variables is said to be associated if every finite subfamily is associated. (cf. Esary et al., 1967). Further suppose that $\{X_n,\ n \geqslant 1\}$ is a stationary sequence of associated random variables. Let F be the distribution function and f the density function for X_1.

Associated random variables are of considerable interest in reliability studies, percolation theory and statistical mechanics. For a review of several probabilistic and statistical results for associated sequences, see Prakasa Rao and Dewan (2000).

Let $\psi(x, y)$ be a real-valued function symmetric in its arguments. Define a U-statistic by

$$U_n = \frac{\sum_{1 \leqslant i < j \leqslant n} \psi(X_i, X_j)}{\binom{n}{2}}. \tag{1.1}$$

Suppose that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^2(x, y)\, \mathrm{d}F(x)\, \mathrm{d}F(y) < \infty. \tag{1.2}$$

Let us consider an orthonormal basis $\{e_k(x),\ k \geqslant 0\}$, with $e_0(x) = 1$, such that

$$\psi(x, y) = \sum_{k=0}^{\infty} \lambda_k e_k(x) e_k(y). \tag{1.3}$$

Then

$$\int_{-\infty}^{\infty} e_k(x)\psi(x, y)\, \mathrm{d}F(x) = \lambda_k e_k(y) \tag{1.4}$$

and

$$\sum_{k=0}^{\infty} \lambda_k^2 < \infty. \tag{1.5}$$

Definition 1.1 (*Gregory*, 1977). The U-statistic U_n and its kernel ψ are called *degenerate* if

$$\int_{-\infty}^{\infty} \psi(x, y)\, \mathrm{d}F(y) = 0 \tag{1.6}$$

for all x.

Gregory (1977) discussed the central limit theorem for degenerate U-statistics based on i.i.d. sequences and its applications to Cramer–Von Mises test. Hall (1979) made a unified study of the invariance principle of degenerate and nondegenerate U-statistics in the i.i.d. case. Eagleson (1979) extended the above methods to derive the limiting distribution of U-statistics based on stationary mixing samples. Our aim in this paper is to obtain similar results for stationary associated real random sequences.

We now discuss few definitions and results from Newman (1984) which will be used in proving our main results.

Definition 1.2 (*Newman*, 1984). Let f and f_1 be two complex-valued functions on $\mathbb{R}^n$. Then we say that $f \ll f_1$ if $f_1 - \mathrm{Re}(\mathrm{e}^{\mathrm{i}\alpha} f)$ is componentwise non-decreasing for all real α.

Remark 1.3 (*Newman*, 1984). Let f and f_1 be two real-valued functions. Then $f \ll f_1$ if and only if $f_1 + f$ and $f_1 - f$ are both non-decreasing. In particular, if $f \ll f_1$ and f, f_1 are functions of a single variable, then f_1 will be non-decreasing.

Definition 1.4 (*Newman*, 1984). $f \overset{A}{\ll} f_1$ if $f \ll f_1$ and both f_1 and f depend only on $x_j' s$ with $j \in A$.

Remark 1.5. Suppose that f is real-valued function. Then $f \ll f_1$ for f_1 real iff for $x < y$,

$$f(y) - f(x) \leqslant f_1(y) - f_1(x) \tag{1.7}$$

and

$$f(x) - f(y) \leqslant f_1(y) - f_1(x). \tag{1.8}$$

These relations hold iff, for $x < y$,

$$|f(y) - f(x)| \leqslant f_1(y) - f_1(x). \tag{1.9}$$

Theorem 1.6 (Newman, 1984). *Let $\{X_n, n \geqslant 1\}$ be a stationary sequence of associated random variables. For each j, let $Y_j = f_j(X_1, X_2, \ldots)$, and $\tilde{Y}_j = \tilde{f}_j(X_1, X_2, \ldots)$. Suppose that $f_j \overset{A}{\ll} \tilde{f}_j$ for each j where $A = \{k\colon k \geqslant 1\}$. Then*

$$\left| \phi - \prod_{j=1}^{n} \phi_j \right| \leqslant 2 \sum_{1 \leqslant k < \ell \leqslant n} |r_k r_\ell \operatorname{Cov}(\tilde{Y}_k, \tilde{Y}_\ell)|, \tag{1.10}$$

where ϕ and ϕ_j are given by

$$\phi = E\left(\exp\left[i \sum_{j=1}^{n} r_j Y_j\right]\right) \text{ and } \phi_j = E(\exp[i r_j Y_j]).$$

Theorem 1.7 (Newman, 1984). *Let $\{X_n, n \geqslant 1\}$ be a stationary sequence of associated random variables. For each j, let $Y_j = f(X_j)$ and $\tilde{Y}_j = \tilde{f}(X_j)$. Suppose that $f \ll \tilde{f}$. Let*

$$\sum_{j=2}^{\infty} \operatorname{Cov}(\tilde{Y}_1, \tilde{Y}_j) < \infty. \tag{1.11}$$

Then

$$n^{-1/2} \sum_{j=1}^{n} (Y_j - EY_j) \xrightarrow{\mathscr{L}} \sigma Z \quad \text{as } n \to \infty \tag{1.12}$$

where Z is a standard normal random variable and

$$\sigma^2 = \operatorname{Var}(Y_1) + 2 \sum_{j=2}^{\infty} \operatorname{Cov}(Y_1, Y_j). \tag{1.13}$$

Remark 1.8. Note that if $\tilde{f}$ is differentiable and $\sup_x |\tilde{f}'(x)|$ is finite, then, using Newman's (1980), inequality (1.11) holds provided

$$\sum_{j=2}^{\infty} \operatorname{Cov}(X_1, X_j) < \infty. \tag{1.14}$$

Remark 1.9. Suppose $f \ll \tilde{f}$. Following the Remark 1.5, we must have, for $x < y$,

$$|f(y) - f(x)| \leqslant \tilde{f}(y) - \tilde{f}(x). \tag{1.15}$$

Let $\tilde{f}(x) = cx$ for some constant $c > 0$. Then $f \ll \tilde{f}$ iff, for $x < y$,

$$|f(y) - f(x)| \leqslant c(y - x). \tag{1.16}$$

which indicates that f is Lipschitzian. A sufficient condition for (1.16) is that

$$\sup_x |f'(x)| \leqslant C. \tag{1.17}$$

2. Limit theorem for U-statistics for kernels of degree 2

Let C denote a generic positive constant in the following discussion.

Theorem 2.1. *Let $\{X_n, n \geqslant 1\}$ be a stationary sequence of associated random variables. Let U_n be a degenerate* U-*statistic where the kernel $\psi(.,.)$ satisfies* (1.2). *Assume that the eigenfunctions $e_k(x)$ given by* (1.4) *are differentiable and*

$$\sup_j \sup_x |e_j'(x)| < \infty. \tag{2.1}$$

Furthermore, assume that

$$\sum_{j=1}^{\infty} \operatorname{Cov}(X_1, X_j) < \infty \tag{2.2}$$

and

$$\sum_{k=1}^{\infty} |\lambda_k| < \infty. \tag{2.3}$$

Then

$$nU_n \xrightarrow{\mathscr{L}} \sum_{k=1}^{\infty} \lambda_k (\tilde{Z}_k^2 - 1) \quad \text{as } n \to \infty. \tag{2.4}$$

where $\{\tilde{Z}_k\}$ is a sequence of correlated jointly normal random variables with mean zero and

$$\operatorname{Cov}(\tilde{Z}_k, \tilde{Z}_j) = \operatorname{Cov}(e_k(X_1), e_j(X_1)) + 2\sum_{i=1}^{\infty} \operatorname{Cov}(e_k(X_1), e_j(X_{(1+i)})).$$

for $k \neq j$.

Proof. Since the kernel ψ satisfies (1.2) and it is degenerate, we have

$$\psi(x, y) = \sum_{k=1}^{\infty} \lambda_k e_k(x) e_k(y). \tag{2.5}$$

Given $\varepsilon > 0$, there exists a positive integer N such that if

$$\psi_N(x, y) = \sum_{k=1}^{N} \lambda_k e_k(x) e_k(y), \tag{2.6}$$

then

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi(x, y) - \psi_N(x, y)|^2 \, dF(x) \, dF(y) = \sum_{k=N+1}^{\infty} \lambda_k^2 < \varepsilon. \tag{2.7}$$

Let $U_{n,N}$ be the U-statistic based on the kernel $\psi_N(x,y)$. Then

$$U_{n,N} = \frac{1}{\binom{n}{2}} \sum_{1\leqslant i<j\leqslant n} \psi_N(X_i,X_j)$$

$$= \frac{1}{\binom{n}{2}} \sum_{k=1}^{N} \left(\sum_{1\leqslant i<j\leqslant n} \lambda_k e_k(X_i) e_k(X_j) \right)$$

$$= \frac{1}{n(n-1)} \sum_{k=1}^{N} \lambda_k \left(\sum_{i=1}^{n}\sum_{j=1}^{n} e_k(X_i)e_k(X_j) - \sum_{i=1}^{n} e_k^2(X_i) \right)$$

$$= \frac{1}{n(n-1)} \sum_{k=1}^{N} \lambda_k \left(\left(\sum_{i=1}^{n} e_k(X_i) \right)^2 - \sum_{i=1}^{n} e_k^2(X_i) \right). \tag{2.8}$$

Then

$$nU_{n,N} = \frac{n}{n-1} \sum_{k=1}^{N} \lambda_k \left(\left(\frac{\sum_{i=1}^{n} e_k(X_i)}{\sqrt{n}} \right)^2 - \frac{1}{n} \sum_{i=1}^{n} e_k^2(X_i) \right). \tag{2.9}$$

Because of (2.1), (2.2) and using the Strong law of large numbers for differentiable functions of associated random variables (Bagai and Prakasa Rao, 1995), we get that

$$\frac{1}{n} \sum_{i=1}^{n} e_k^2(X_i) \to 1 \text{ a.s.} \quad \text{as } n \to \infty. \tag{2.10}$$

Next, we consider the joint distribution of

$$\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} e_1(X_i), \ldots, \frac{1}{\sqrt{n}} \sum_{i=1}^{n} e_N(X_i) \right). \tag{2.11}$$

Consider a linear combination

$$T = \sum_{k=1}^{N} \frac{a_k}{\sqrt{n}} \sum_{i=1}^{n} e_k(X_i)$$

$$= \sum_{i=1}^{n} \frac{1}{\sqrt{n}} \sum_{k=1}^{N} a_k e_k(X_i)$$

$$= \sum_{i=1}^{n} \frac{1}{\sqrt{n}} B_N(X_i), \tag{2.12}$$

where

$$B_N(X_i) = \sum_{k=1}^{N} a_k e_k(X_i). \tag{2.13}$$

Then, under condition (2.1), B_N satisfies the conditions of Theorem 1.7 for every vector $(a_1,\ldots,a_N) \in R^N$. Note that

$$ET = 0,$$

$$\operatorname{Var} T = \sigma_N^2 = \sum_{k=1}^{N} a_k^2 + \frac{2}{n} \sum_{j=1}^{n-1} (n-j)\operatorname{Cov}(B_N(X_1), B_N(X_{1+j})). \tag{2.14}$$

Hence, by Theorem 1.7,

$$n^{-1/2}\sum_{i=1}^{n} B_N(X_i) \xrightarrow{\mathscr{L}} \mathrm{N}(0,\sigma_N^2) \quad \text{as } n\to\infty. \tag{2.15}$$

Therefore, using (2.9), (2.10) and (2.15), we get that

$$nU_{n,N} \xrightarrow{\mathscr{L}} \sum_{k=1}^{N} \lambda_k(\tilde{Z}_k^2-1) \quad \text{as } n\to\infty \tag{2.16}$$

where $\{\tilde{Z}_k, 1\leqslant k\leqslant N\}$ are jointly normal random variables with mean zero and

$$\operatorname{Cov}(\tilde{Z}_k,\tilde{Z}_j) = \operatorname{Cov}(e_k(X_1),e_j(X_1)) + 2\sum_{i=1}^{\infty}\operatorname{Cov}(e_k(X_1),e_j(X_{(1+i)}) \tag{2.17}$$

for $k\neq j$.

Note that

$$\begin{aligned} E(\tilde{Z}_k^2) &= \operatorname{Var}(e_k(X_1)) + 2\sum_{i=1}^{\infty}\operatorname{Cov}(e_k(X_1),e_k(X_{1+i})) \\ &\leqslant 1 + C\sum_{i=1}^{\infty}\operatorname{Cov}(X_1,X_{1+i}) \quad \text{(by using (2.1) and the Newman's Inequality)} \\ &\leqslant C \quad \text{(by using (2.2))} \end{aligned} \tag{2.18}$$

uniformly in $k\geqslant 1$.

Let $\eta_n(t)$, $\phi(t)$, $\phi_N(t)$ and $\eta_{n,N}(t)$ be the characteristic functions of $nU_n, \sum_{k=1}^{\infty}\lambda_k(\tilde{Z}_k^2-1), \sum_{k=1}^{N}\lambda_k(\tilde{Z}_k^2-1)$ and $nU_{n,N}$, respectively.

Then

$$\begin{aligned} |\eta_n(t)-\phi(t)| \leqslant{}& |\eta_n(t)-\eta_{n,N}(t)| \\ &+ |\eta_{n,N}(t)-\phi_N(t)| + |\phi_N(t)-\phi(t)|. \end{aligned} \tag{2.19}$$

Relation (2.16) implies that, given $\varepsilon>0$, for large n depending on N, ε and t,

$$|\eta_{n,N}(t)-\phi_N(t)| \leqslant \frac{\varepsilon}{3}. \tag{2.20}$$

Since

$$\begin{aligned} E\left|\sum_{k=N+1}^{\infty}\lambda_k(\tilde{Z}_k^2-1)\right| &\leqslant E\left\{\sum_{k=N+1}^{\infty}|\lambda_k||\tilde{Z}_k^2-1|\right\} \\ &\leqslant (C+1)\sum_{k=N+1}^{\infty}|\lambda_k| < \varepsilon, \end{aligned} \tag{2.21}$$

it follows that, for large N, depending on t

$$\begin{aligned} |\phi_N(t)-\phi(t)| &= |E(\mathrm{e}^{\mathrm{i}t\sum_{k=1}^{N}\lambda_k(\tilde{Z}_k^2-1)} - \mathrm{e}^{\mathrm{i}t\sum_{k=1}^{\infty}\lambda_k(\tilde{Z}_k^2-1)})| \\ &\leqslant E|\mathrm{e}^{\mathrm{i}t\sum_{k=N+1}^{\infty}\lambda_k(\tilde{Z}_k^2-1)} - 1| \end{aligned}$$

$$\begin{aligned}
&\leqslant |t|E\left|\sum_{k=N+1}^{\infty}\lambda_k(\tilde{Z}_k^2-1)\right| \\
&\leqslant |t|\sum_{k=N+1}^{\infty}|\lambda_k|(E\tilde{Z}_k^2+1) \\
&\leqslant C|t|\sum_{k=N+1}^{\infty}|\lambda_k| \\
&\leqslant \frac{\varepsilon}{3} \quad \text{(by (2.3) and (2.18))}.
\end{aligned} \tag{2.22}$$

Observe that

$$\begin{aligned}
E\left(n^{-1/2}\sum_{i=1}^{n}e_k(X_i)\right)^2 &= E\frac{1}{n}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}e_k(X_i)e_k(X_j)\right) \\
&= \frac{1}{n}E\left(\sum_{i=1}^{n}e_k^2(X_i)+\sum_{1\leqslant i\neq j\leqslant n}e_k(X_i)e_k(X_j)\right) \\
&= 1+\frac{1}{n}\sum_{1\leqslant i\neq j\leqslant n}E(e_k(X_i)e_k(X_j)) \\
&\leqslant 1+\frac{1}{n}\sum_{1\leqslant i\neq j\leqslant n}\left(\sup_x|e_k'(x)|\right)^2\operatorname{Cov}(X_i,X_j) \\
&\leqslant 1+\frac{C}{n}\sum_{j=2}^{n}(n-j)\operatorname{Cov}(X_1,X_j) \quad \text{(using stationarity)} \\
&\leqslant C \quad \text{(using (2.2))}.
\end{aligned} \tag{2.23}$$

Hence,

$$\begin{aligned}
E|nU_n-nU_{n,N}| &= E\left|\frac{n}{n-1}\sum_{k=N+1}^{\infty}\lambda_k\left\{\left(\frac{\sum_{i=1}^{n}e_k(X_i)}{\sqrt{n}}\right)^2-\frac{1}{n}\sum_{i=1}^{n}e_k^2(X_i)\right\}\right| \\
&\leqslant \frac{n}{n-1}\sum_{k=N+1}^{\infty}\left|\lambda_k|E|\left\{\left(\frac{\sum_{i=1}^{n}e_k(X_i)}{\sqrt{n}}\right)^2-\frac{1}{n}\sum_{i=1}^{n}e_k^2(X_i)\right\}\right| \\
&\leqslant \frac{n}{n-1}\sum_{k=N+1}^{\infty}|\lambda_k|\left\{E\left(\frac{\sum_{i=1}^{n}e_k(X_i)}{\sqrt{n}}\right)^2+\frac{1}{n}\sum_{i=1}^{n}E(e_k^2(X_i))\right\} \\
&\leqslant \frac{n}{n-1}(C+1)\sum_{k=N+1}^{\infty}|\lambda_k|.
\end{aligned} \tag{2.24}$$

Hence, we have

$$\begin{aligned}
|\eta_n(t)-\eta_{n,N}(t)| &= |E(\mathrm{e}^{\mathrm{i}tnU_n}-\mathrm{e}^{\mathrm{i}tnU_{n,N}})| \\
&\leqslant E|\mathrm{e}^{\mathrm{i}tn(U_n-U_{n,N})}-1|
\end{aligned}$$

$$\leqslant |t| E|nU_n - nU_{n,N}|$$
$$< \frac{\varepsilon}{3} \tag{2.25}$$

for large n, N depending on t.

The result now follows by combining (2.19), (2.20), (2.22) and (2.25).

Theorem 2.2. *Let $\{X_n, n \geqslant 1\}$ be a stationary sequence of associated random variables. Suppose U_n is a non-degenerate U-statistic corresponding to the kernel ψ with the eigenfunction expansion* (1.3). *Assume that*

$$\sup_j \sup_x |e_j'(x)| < \infty \tag{2.26}$$

and

$$\sum_{j=1}^{\infty} \mathrm{Cov}(X_1, X_j) < \infty. \tag{2.27}$$

Further assume that

$$\sum_{k=0}^{\infty} |\lambda_k| < \infty. \tag{2.28}$$

If U_n has a finite variance, then

$$n^{1/2}(U_n - EU_n) \xrightarrow{\mathscr{L}} \mathrm{N}(0, 4\sigma^2) \quad \text{as } n \to \infty. \tag{2.29}$$

where

$$\sigma^2 = \mathrm{Var}(g(X_1)) + 2\sum_{j=1}^{\infty} \mathrm{Cov}(g(X_1), g(X_{j+1})), \tag{2.30}$$

and

$$g(x) = \int_{-\infty}^{\infty} \psi(x, y)\, \mathrm{d}F(y), \tag{2.31}$$

provided $g(.)$ is monotone or g is Lipschitzian, that is,

$$|g(x) - g(y)| \leqslant C|x - y|, \quad x, y \in R \tag{2.32}$$

and $\sum_{k=0}^{\infty} |a_k^*| < \infty$, where $a_k^* = E(e_k(X_1))$.

Proof. Note that

$$g(x) = \int_{-\infty}^{\infty} \psi(x, y)\, \mathrm{d}F(y)$$
$$= \int_{-\infty}^{\infty} \left\{ \sum_{k=0}^{\infty} \lambda_k\, e_k(x) e_k(y) \right\} \mathrm{d}F(y)$$
$$= \sum_{k=0}^{\infty} \lambda_k e_k(x) a_k^*. \tag{2.33}$$

Then

$$|a_k^*| \leqslant E|e_k(X_1)| \leqslant \{E(e_k^2(X_1))\}^{1/2} = 1, \tag{2.34}$$

and

$$Eg(X_1) = \sum_{k=0}^{\infty} \lambda_k a_k^{*2}. \tag{2.35}$$

Using (2.35) and the fact that $e_k(x)$ are orthogonal functions, we get

$$E(g^2(X_1)) = \sum_{k=0}^{\infty} \lambda_k^2 a_k^{*2} \leqslant \sum_{k=0}^{\infty} \lambda_k^2 < \infty. \tag{2.36}$$

Define

$$\hat{\psi}(x,y) = \psi(x,y) - g(x) - g(y) + Eg(X_1). \tag{2.37}$$

It is easy to see that $\int_{-\infty}^{\infty} \hat{\psi}(x,y)\,\mathrm{d}F(y)$ is zero. Thus $\hat{\psi}$ is symmetric, square integrable and degenerate.

Let $\hat{U}_n$ be the U-statistic corresponding to $\hat{\psi}$. Note that

$$\hat{U}_n = \frac{1}{\binom{n}{2}} \sum_{k=0}^{\infty} \lambda_k \left\{ \sum_{1 \leqslant i < j \leqslant n} (e_k(X_i) - a_k^*)(e_k(X_j) - a_k^*) \right\}. \tag{2.38}$$

Using (2.37) and (2.38), we have

$$U_n - \sum_{k=0}^{\infty} \lambda_k a_k^{*2} = \hat{U}_n + \frac{2}{n} \sum_{i=1}^{n} g(X_i) - 2 \sum_{k=0}^{\infty} \lambda_k a_k^{*2}$$

$$= \hat{U}_n + \frac{2}{n} \sum_{i=1}^{n} (g(X_i) - Eg(X_i)). \tag{2.39}$$

Now,

$$E(e_k(X_1) - a_k^*) = 0,$$

$$E(e_k(X_1) - a_k^*)^2 = 1 - a_k^{*2}. \tag{2.40}$$

Then by (2.26)–(2.28) and an application of Theorem 2.1 with $e_k(x)$ replaced by $e_k(X) - a_k^*$, we get that

$$n\hat{U}_n \xrightarrow{\mathscr{L}} \sum_{k=0}^{\infty} \lambda_k (\hat{Z}_k^2 - 1 + a_k^{*2}) \quad \text{as } n \to \infty \tag{2.41}$$

where $\{\hat{Z}_k\}$ is a sequence of jointly normal random variables.

In view of (2.41), we have

$$n^{1/2}\hat{U}_n \xrightarrow{\mathrm{P}} 0 \quad \text{as } n \to \infty. \tag{2.42}$$

Furthermore,

$$\begin{aligned}\left|E(U_n)-\sum_{k=0}^{\infty}\lambda_k a_k^{*2}\right| &= \frac{2}{n(n-1)}\left|\sum_{k=0}^{\infty}\lambda_k\sum_{1\leqslant i<j\leqslant n}\{E(e_k(X_i)e_k(X_j))-a_k^{*2}\}\right|\\ &\leqslant \frac{2}{n(n-1)}\left|\sum_{k=0}^{\infty}\lambda_k\sum_{1\leqslant i<j\leqslant n}\{E(e_k(X_i)-a_k^{*})(e_k(X_j))-a_k^{*})\}\right|\\ &\leqslant \frac{2}{n(n-1)}\sum_{k=0}^{\infty}|\lambda_k|\sum_{1\leqslant i<j\leqslant n}|\mathrm{Cov}(e_k(X_i),e_k(X_j))|\\ &\leqslant \frac{2}{n(n-1)}\sum_{k=0}^{\infty}|\lambda_k|\sup_k\sup_x(e_k'(x))^2\sum_{1\leqslant i<j\leqslant n}\mathrm{Cov}(X_i,X_j)\\ &\leqslant \frac{C}{n}\sum_{k=0}^{\infty}|\lambda_k|\sum_{i=2}^{n}\mathrm{Cov}(X_1,X_i). \qquad (2.43)\end{aligned}$$

Therefore,

$$n^{1/2}\left|E(U_n)-\sum_{k=0}^{\infty}\lambda_k a_k^{*2}\right|\to 0\quad \text{as } n\to\infty \qquad (2.44)$$

from (2.27) and (2.28).

If $g(x)$ is monotone, then following Newman (1980) we have, as $n\to\infty$,

$$\frac{2}{\sqrt{n}}\sum_{i=1}^{n}(g(X_i)-Eg(X_i))\xrightarrow{\mathscr{L}}\mathrm{N}(0,4\sigma^2)\quad \text{as } n\to\infty. \qquad (2.45)$$

If $g(x)$ is not monotone then let $\tilde{g}(x)=Cx$. Since

$$\sum_{k=0}^{\infty}|a_k^{*}|<C<\infty, \qquad (2.46)$$

it follows from Remark 1.9 and (2.26) that $g\ll\tilde{g}$. Therefore (2.45) follows from Theorem 1.8.

The result follows by combining (2.42), (2.44) and (2.45).

3. Limit theorem for U-statistics of kernels of degree 3

Suppose that $\psi(x,y,z)$ is a symmetric function from $\mathbb{R}^3$ to $\mathbb{R}$ which is square integrable in the sense that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi^2(x,y,z)\,\mathrm{d}F(x)\,\mathrm{d}F(y)\,\mathrm{d}F(z)<\infty. \qquad (3.1)$$

Assume that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\psi(x,y,z)\,\mathrm{d}F(x)\,\mathrm{d}F(y)\,\mathrm{d}F(z)=0. \qquad (3.2)$$

Suppose there exists an orthonormal basis $\{e_k(x),\ k\geqslant 0\}$ with $e_0(x)=1$ such that

$$\psi(x,y,z)=\sum_{k,l,m\geqslant 0}\lambda_{klm}e_k(x)e_l(y)e_m(z), \qquad (3.3)$$

where

$$\sum_{k,l,m \geqslant 0} \lambda_{klm}^2 < \infty. \tag{3.4}$$

Note that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(x,y,z) e_k(x) e_\ell(y) e_m(z) \,\mathrm{d}F(x) \,\mathrm{d}F(y) \,\mathrm{d}F(z) = \lambda_{klm}. \tag{3.5}$$

Then the U-statistic corresponding to the kernel ψ of degree three is

$$U_n = \frac{1}{\binom{n}{3}} \sum_{1 \leqslant i_1 < i_2 < i_3 \leqslant n} \psi(X_{i_1}, X_{i_2}, X_{i_3}). \tag{3.6}$$

Since $e_0(x) = 1$, we get that $\lambda_{000} = 0$ from (3.5). Then,

$$\begin{aligned} \psi(x,y,z) = & \sum_{k>0} \lambda_{k00} e_k(x) + \sum_{\ell > 0} \lambda_{0\ell 0} e_\ell(y) \\ & + \sum_{m>0} \lambda_{00m} e_m(z) + \sum_{k,\ell > 0} \lambda_{k\ell 0} e_k(x) e_\ell(y) \\ & + \sum_{k,m>0} \lambda_{k0m} e_k(x) e_m(z) + \sum_{m,\ell > 0} \lambda_{0\ell m} e_m(z) e_\ell(y) \\ & + \sum_{k,\ell,m>0} \lambda_{k,\ell,m} e_k(x) e_\ell(y) e_m(z). \end{aligned} \tag{3.7}$$

Definition 3.1. The kernel ψ is said to be *first-order degenerate* if

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(x,y,z) \,\mathrm{d}F(x) \,\mathrm{d}F(y) = 0. \tag{3.8}$$

Remark 3.2. If the kernel is first-order degenerate, then

$$\lambda_{k00} = \lambda_{0\ell 0} = \lambda_{00m} = 0 \quad \text{for all } k, \ell, m.$$

Hence

$$\begin{aligned} \psi(x,y,z) = & \sum_{k,\ell > 0} \lambda_{k\ell 0} e_k(x) e_\ell(y) + \sum_{k,m>0} \lambda_{k0m} e_k(x) e_m(z) \\ & + \sum_{m,\ell > 0} \lambda_{0\ell m} e_m(z) e_\ell(y) + \sum_{k,\ell,m>0} \lambda_{k,\ell,m} e_k(x) e_\ell(y) e_m(z). \end{aligned} \tag{3.9}$$

Definition 3.3. The kernel ψ is said to be *second-order degenerate* if

$$\int_{-\infty}^{\infty} \psi(x,y,z) \,\mathrm{d}F(x) = 0. \tag{3.10}$$

Remark 3.4. If the kernel ψ is second-order degenerate, then

$$\lambda_{k\ell 0} = \lambda_{0\ell m} = \lambda_{k0m} = 0 \quad \text{for all } k, \ell, m$$

and

$$\psi(x,y,z) = \sum_{k,\ell,m=1}^{\infty} \lambda_{k\ell m} e_k(x) e_\ell(y) e_m(z). \tag{3.11}$$

Theorem 3.5. *Let $\{X_n,\ n \geqslant 1\}$ be a stationary sequence of associated random variables. Let U_n be a* U*-statistic of degree three where the kernel $\psi(.,.,.)$ is second-order degenerate and satisfies* (3.1). *Assume that the eigenfunctions $e_k(x)$ are differentiable and*

$$\sup_j \sup_x |e_j(x)| < \infty, \tag{3.12}$$

and

$$\sup_j \sup_x |e'_j(x)| < \infty, \tag{3.13}$$

Furthermore, assume that

$$\sum_{j=n+1}^{\infty} \mathrm{Cov}(X_1, X_j) = \mathrm{O}(n^{-\theta}) \quad \text{for } \theta > 1 \tag{3.14}$$

and assume that

$$\sum_{k,\ell,m=1}^{\infty} |\lambda_{k\ell m}| < \infty.$$

Then

$$n^{3/2} U_n \xrightarrow{\mathscr{L}} \sum_{k,\ell,m=1}^{\infty} \lambda_{k\ell m} \tilde{Z}_k \tilde{Z}_\ell \tilde{Z}_m \quad \text{as } n \to \infty$$

where $\{\tilde{Z}_k\}$ is a sequence of mean zero correlated jointly normal random variables with covariances given by (2.17).

Proof. Since the kernel ψ satisfies (3.1) and it is second-order degenerate, we have

$$\psi(x,y,z) = \sum_{k,\ell,m=1}^{\infty} \lambda_{k\ell m} e_k(x) e_\ell(y) e_m(z). \tag{3.15}$$

Let A be the set $\{1 \leqslant k, \ell, m \leqslant N\}$. Let B_N denote the complement of set A. Given $\varepsilon > 0$, there exists a positive integer N such that if

$$\psi_N(x,y,z) = \sum_{1 \leqslant k,\ell,m \leqslant N} \lambda_{k\ell m} e_k(x) e_\ell(y) e_m(z), \tag{3.16}$$

then

$$\int\int |\psi(x,y,z) - \psi_N(x,y,z)|^2 \,\mathrm{d}F(x)\,\mathrm{d}F(y)\,\mathrm{d}F(z) = \sum_{B_N} \lambda^2_{k\ell m} < \varepsilon. \tag{3.17}$$

Let $U_{n,N}$ be the U-statistic based on the kernel $\psi_N(x,y,z)$. Then

$$U_{n,N} = \frac{1}{\binom{n}{3}} \sum_{1 \leqslant i<j<r \leqslant n} \psi_N(X_i, X_j, X_r)$$

$$= \frac{1}{\binom{n}{3}} \sum_{1 \leqslant k,\ell,m \leqslant N} \left(\sum_{1 \leqslant i<j<r \leqslant n} \lambda_{k\ell m} e_k(X_i) e_\ell(X_j) e_m(X_r) \right)$$

$$= \frac{1}{n(n-1)(n-2)} \sum_{1 \leqslant k,\ell,m \leqslant N} \lambda_{k\ell m} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{r=1}^{n} e_k(X_i) e_\ell(X_j) e_m(X_r) \right.$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n} e_k(X_i)e_\ell(X_i)e_m(X_j) - \sum_{i=1}^{n}\sum_{j=1}^{n} e_k(X_i)e_\ell(X_j)e_m(X_i)$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n} e_k(X_j)e_\ell(X_i)e_m(X_i) + 2\sum_{i=1}^{n} e_k(X_i)e_\ell(X_i)e_m(X_i)\Bigg\}. \tag{3.18}$$

Therefore,

$$n^{3/2}U_{n,N} = \frac{n^3}{n(n-1)(n-2)} \sum_{1\leqslant k,\ell,m\leqslant N} \lambda_{k\ell m} \Bigg\{ \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{r=1}^{n} \frac{e_k(X_i)}{\sqrt{n}}\frac{e_\ell(X_j)}{\sqrt{n}}\frac{e_m(X_r)}{\sqrt{n}}$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{e_k(X_i)e_\ell(X_i)}{n}\frac{e_m(X_j)}{\sqrt{n}} - \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{e_k(X_i)e_m(X_i)}{n}\frac{e_\ell(X_j)}{\sqrt{n}}$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{e_k(X_j)}{\sqrt{n}}\frac{e_\ell(X_i)e_m(X_i)}{n} + \frac{2}{n^{3/2}}\sum_{i=1}^{n} e_k(X_i)e_\ell(X_i)e_m(X_i)\Bigg\}. \tag{3.19}$$

Because of (3.12), (3.13) and the Strong law of large numbers for associated random variables (Bagai and Prakasa Rao, 1995), we get that

$$\frac{1}{n}\sum_{i=1}^{n} e_k(X_i)e_m(X_i) \to 0 \quad \text{a.s. as } n\to\infty, \text{ for all } k,m, \tag{3.20}$$

and

$$\frac{1}{n^{3/2}}\sum_{i=1}^{n} e_k(X_i)e_\ell(X_i)e_m(X_i) \to 0 \quad \text{a.s. as } n\to\infty, \text{ for all } k,\ell,m. \tag{3.21}$$

From (2.11)–(2.15), (3.19) and the discussion for deriving (2.15), we get that

$$n^{3/2}U_{n,N} \xrightarrow{\mathscr{L}} \sum_{1\leqslant k,\ell,m\leqslant N} \lambda_{k\ell m}\tilde{Z}_k\tilde{Z}_\ell\tilde{Z}_m \quad \text{as } n\to\infty. \tag{3.22}$$

Let $\eta_n^*(t)$, $\eta(t)$, $\eta_N(t)$ and $\eta_{n,N}^*(t)$ be the characteristic functions of $n^{3/2}U_n$, $\sum_{k,\ell,m=1}^{\infty}\lambda_{k\ell m}\tilde{Z}_k\tilde{Z}_\ell\tilde{Z}_m$, $\sum_{1\leqslant k,\ell,m\leqslant N}\lambda_{k\ell m}\tilde{Z}_k\tilde{Z}_\ell\tilde{Z}_m$, and $n^{3/2}U_{n,N}$, respectively.
Then

$$|\eta_n^*(t)-\eta(t)| \leqslant |\eta_n^*(t)-\eta_{n,N}^*(t)| + |\eta_{n,N}^*(t)-\eta_N(t)| + |\eta_N(t)-\eta(t)|. \tag{3.23}$$

Relation (3.22) implies that, given $\varepsilon>0$, for large n depending on N, ε and t,

$$|\eta_{n,N}^*(t)-\eta_N(t)| \leqslant \frac{\varepsilon}{3}. \tag{3.24}$$

Using Cauchy–Schwartz inequality and the fact that $\tilde{Z}_k$ has normal distribution with mean zero, we get that

$$\begin{aligned}(E|\tilde{Z}_k\tilde{Z}_\ell\tilde{Z}_m|)^2 &\leqslant E(\tilde{Z}_k)^2E|\tilde{Z}_\ell\tilde{Z}_m|^2\\ &\leqslant E(\tilde{Z}_k)^2\{E(\tilde{Z}_\ell)^4E(\tilde{Z}_m)^4\}^{1/2}\\ &\leqslant 3E(\tilde{Z}_k)^2\{(E(\tilde{Z}_\ell)^2)^2(E(\tilde{Z}_m)^2)^2\}^{1/2}\\ &\leqslant C \text{ (by using (2.18))}\end{aligned} \tag{3.25}$$

uniformly in k,ℓ,m. Hence,

$$E\left|\sum_{B_N}\lambda_{k\ell m}\tilde{Z}_k\tilde{Z}_\ell\tilde{Z}_m\right| \leqslant C\sum_{B_N}|\lambda_{k\ell m}| < \varepsilon, \tag{3.26}$$

for large N. It follows that, for large N, depending on t

$$\begin{aligned}|\eta_N(t)-\eta(t)| &= |E(\mathrm{e}^{\mathrm{i}t\sum_{1\leqslant k,\ell,m\leqslant N}\lambda_{k\ell m}\tilde{Z}_k\tilde{Z}_\ell\tilde{Z}_m} - \mathrm{e}^{\mathrm{i}t\sum_{k,\ell,m=1}^{\infty}\lambda_{k\ell m}\tilde{Z}_k\tilde{Z}_\ell\tilde{Z}_m})| \\ &\leqslant E|\mathrm{e}^{\mathrm{i}t\sum_{B_N}\lambda_{k\ell m}\tilde{Z}_k\tilde{Z}_\ell\tilde{Z}_m}-1| \\ &\leqslant |t|E\left|\sum_{B_N}\lambda_{k\ell m}\tilde{Z}_k\tilde{Z}_\ell\tilde{Z}_m\right| \\ &\leqslant \frac{\varepsilon}{3} \text{ (by using (3.4) and (3.25)).}\end{aligned} \tag{3.27}$$

Let us now consider

$$\begin{aligned}&E|n^{3/2}U_n - n^{3/2}U_{n,N}| \\ &= E\left|\frac{n^3}{n(n-1)(n-2)}\sum_{1\leqslant k,\ell,m\leqslant N}\lambda_{k\ell m}\left\{\sum_{i=1}^n\sum_{j=1}^n\sum_{r=1}^n\frac{e_k(X_i)}{\sqrt{n}}\frac{e_\ell(X_j)}{\sqrt{n}}\frac{e_m(X_r))}{\sqrt{n}}\right.\right. \\ &\quad -\sum_{i=1}^n\sum_{j=1}^n\frac{e_k(X_i)e_\ell(X_i)}{n}\frac{e_m(X_j))}{\sqrt{n}} - \sum_{i=1}^n\sum_{j=1}^n\frac{e_k(X_i)e_m(X_i)}{n}\frac{e_\ell(X_j)}{\sqrt{n}} \\ &\quad \left.\left. -\sum_{i=1}^n\sum_{j=1}^n\frac{e_k(X_j)}{\sqrt{n}}\frac{e_\ell(X_i)e_m(X_i)}{n} + \frac{2}{n^{3/2}}\sum_{i=1}^n e_k(X_i)e_\ell(X_i)e_m(X_i)\right\}\right|\end{aligned} \tag{3.28}$$

Observe that, because of (3.12),

$$\frac{1}{n^{3/2}}E\left|\sum_{i=1}^n e_k(X_i)e_\ell(X_i)e_m(X_i)\right| \leqslant C. \tag{3.29}$$

Using Cauchy–Schwartz inequality, (2.23), (3.12)–(3.14) and the stationarity of the random variables, we get

$$\begin{aligned}&E\left|\left(\frac{1}{n}\sum_{i=1}^n e_k(X_i)e_\ell(X_i)\right)\left(\frac{1}{\sqrt{n}}\sum_{j=1}^n e_m(X_j)\right)\right| \\ &\leqslant \left\{E\left(\frac{1}{n}\sum_{i=1}^n e_k(X_i)e_\ell(X_i)\right)^2\right\}^{1/2}\left\{E\left(\frac{1}{\sqrt{n}}\sum_{j=1}^n e_m(X_j)\right)^2\right\}^{1/2} \\ &\leqslant CE\left(\frac{1}{n^2}\sum_{i=1}^n e_k^2(X_i)e_\ell^2(X_i) + \frac{1}{n^2}\sum_{1\leqslant i\neq j\leqslant n}e_k(X_i)e_\ell(X_i)e_k(X_j)e_\ell(X_j)\right) \\ &\leqslant C\left\{\frac{1}{n}+\frac{1}{n^2}\sum_{1\leqslant i\neq j\leqslant n}\mathrm{Cov}(e_k(X_i)e_\ell(X_i), e_k(X_j)e_\ell(X_j))\right\}\end{aligned}$$

$$\leqslant C\left\{\frac{1}{n}+\frac{1}{n^2}\sum_{1\leqslant i\neq j\leqslant n}\mathrm{Cov}(X_i,X_j)\right\}$$

$$\leqslant C\left\{\frac{1}{n}+\frac{1}{n^2}\sum_{j=1}^{n}(n-j)\mathrm{Cov}(X_i,X_j)\right\}$$

$$\leqslant C. \tag{3.30}$$

Further using (3.12)–(3.14) and Rosenthal-type inequality for associated sequences (Shao and Yu, 1996), we get

$$E\left|\sum_{i=1}^{n}\frac{e_k(X_i)}{n^{1/2}}\right|^4\leqslant C. \tag{3.31}$$

Finally, using Cauchy–Schwartz inequality, (2.23) and (3.31), we get

$$E\left|\sum_{i=1}^{n}\frac{e_k(X_i)}{n^{1/2}}\sum_{j=1}^{n}\frac{e_\ell(X_j)}{n^{1/2}}\sum_{r=1}^{n}\frac{e_m(X_r)}{n^{1/2}}\right|$$

$$\leqslant E\left\{\left(\sum_{i=1}^{n}\frac{e_k(X_i)}{n^{1/2}}\right)^2\left(\sum_{j=1}^{n}\frac{e_\ell(X_j)}{n^{1/2}}\right)^2\right\}^{1/2}\left\{E\left(\sum_{r=1}^{n}\frac{e_m(X_r)}{n^{1/2}}\right)^2\right\}^{1/2}$$

$$\leqslant C\left\{E\left(\sum_{i=1}^{n}\frac{e_k(X_i)}{n^{1/2}}\right)^4E\left(\sum_{j=1}^{n}\frac{e_\ell(X_j)}{n^{1/2}}\right)^4\right\}^{1/4}$$

$$\leqslant C. \tag{3.32}$$

Hence, by using (3.4), (3.30) and (3.32), we get that

$$E|n^{3/2}U_n-n^{3/2}U_{n,N}|\leqslant C. \tag{3.33}$$

Hence, we have

$$\begin{aligned}|\eta_n^*(t)-\eta_{n,N}^*(t)| &= |E(\mathrm{e}^{\mathrm{i}tn^{3/2}U_n}-\mathrm{e}^{\mathrm{i}tn^{3/2}U_{n,N}})| \\ &\leqslant E|\mathrm{e}^{\mathrm{i}tn^{3/2}(U_n-U_{n,N})}-1| \\ &\leqslant |t|E|n^{3/2}U_n-n^{3/2}U_{n,N}| \\ &< \frac{\varepsilon}{3}\end{aligned} \tag{3.34}$$

for large n,N, depending on t. The result follows by combining (3.23), (3.27) and (3.34).

Remark 3.4. Results in this section can be extended to kernels of arbitrary degree by similar methods.

4. Two sample case

Let $\{X_m,\ m\geqslant 1\}$ be a stationary sequence of associated random variables with one-dimensional marginal distribution function F and $\{Y_n,\ n\geqslant 1\}$ be another stationary sequence of associated random variables with one-dimensional marginal distribution function G. The problem of interest is to test the hypothesis H: $F=G$. We now discuss a limit theorem useful in the study of such problems.

Let $\phi(x,y)$ be a function of two variables which is square integrable in the sense that

$$\int_{R^2} \phi^2(x,y)\,\mathrm{d}F(x)\,\mathrm{d}G(y) < \infty. \tag{4.1}$$

Define

$$U_{m,n} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\phi(X_i,Y_j). \tag{4.2}$$

Under (4.1), there exist systems of functions $\{f_k(x)\}$ and $\{g_k(y)\}$ (with $f_0(x)=g_0(x)=1$ which are complete and orthonormal on the spaces of square integrable functions of X and Y, respectively, so that

$$\phi(x,y) = \sum_{k=0}^{\infty}\lambda_k f_k(x) g_k(y), \tag{4.3}$$

where

$$\sum_{k=0}^{\infty}\lambda_k^2 < \infty, \tag{4.4}$$

and the series in (4.3) converges in mean square with respect to the product measure generated by the distribution function FG.

The functions $\{f_k(x)\}$ and $\{g_k(y)\}$ are eigenfunctions and $\{\lambda_k\}$ are the eigenvalues of ϕ in the sense that, for all $k\geqslant 0$,

$$\int_{-\infty}^{\infty} f_k(x)\phi(x,y)\,\mathrm{d}F(x) = \lambda_k g_k(y), \tag{4.5}$$

and

$$\int_{-\infty}^{\infty} g_k(y)\phi(x,y)\,\mathrm{d}G(y) = \lambda_k f_k(x). \tag{4.6}$$

Definition 4.1. The statistic U_{mn} and its kernel ϕ are called degenerate if

$$\int \phi(x,y)\,\mathrm{d}F(x) = 0 \tag{4.7}$$

for all y and

$$\int \phi(x,y)\,\mathrm{d}G(y) = 0 \tag{4.8}$$

for all x.

Theorem 4.2. *Let U_{mn} be a degenerate two sample U-statistic based on two independent stationary sequences of associated random variables. Suppose that the corresponding kernel ϕ is square integrable. Assume that the eigenfunctions $f_k(x)$ and $g_k(y)$ given by* (4.3) *are differentiable and*

$$\sup_k \sup_x |f_k'(x)| < \infty. \tag{4.9}$$

$$\sup_k \sup_y |g_k'(y)| < \infty. \tag{4.10}$$

Furthermore, assume that

$$\sum_{j=1}^{\infty} \mathrm{Cov}(X_1, X_j) < \infty, \tag{4.11}$$

$$\sum_{j=1}^{\infty} \mathrm{Cov}(Y_1, Y_j) < \infty \tag{4.12}$$

and

$$\sum_{k=1}^{\infty} |\lambda_k| < \infty. \tag{4.13}$$

Let

$$\delta_m = \frac{1}{m} \sum_{i=1}^{m-1} \mathrm{Cov}(X_1, X_{i+1}), \tag{4.14}$$

and

$$\eta_n = \frac{1}{n} \sum_{j=1}^{n-1} \mathrm{Cov}(Y_1, Y_{j+1}). \tag{4.15}$$

Assume that

$$mn\{\delta_m + \eta_n\} \to 0 \quad \text{as } m, n \to \infty. \tag{4.16}$$

Then

$$(mn)^{1/2} U_{mn} \xrightarrow{\mathscr{L}} \sum_{k=1}^{\infty} \lambda_k U_k V_k \quad \text{as } m \to \infty \text{ and } n \to \infty \tag{4.17}$$

where $\{U_k\}$ and $\{V_k\}$ are sequences of correlated zero mean normal random variables, and are independent of each other with

$$\mathrm{Cov}(U_k, U_{k'}) = \mathrm{Cov}(f_k(X_1), f_{k'}(X_1)) + 2\sum_{i=1}^{\infty} \mathrm{Cov}(f_k(X_1), f_{k'}(X_{1+i})),$$

and

$$\mathrm{Cov}(V_k, V_{k'}) = \mathrm{Cov}(g_k(Y_1), g_{k'}(Y_1)) + 2\sum_{i=1}^{\infty} \mathrm{Cov}(g_k(Y_1), g_{k'}(Y_{1+i})),$$

Proof. Given $\varepsilon > 0$, there exists N such that if

$$\phi_N(x,y) = \sum_{k=1}^{N} \lambda_k f_k(x) g_k(y) \tag{4.18}$$

then,

$$\int_{R^2} |\phi(x,y) - \phi_N(x,y)|^2 \, \mathrm{d}F(x) \, \mathrm{d}G(y) = \sum_{k=N+1}^{\infty} \lambda_k^2 < \varepsilon. \tag{4.19}$$

Let $U_{mn}^{(N)}$ be the U-statistic generated from the kernel $\phi_N(x,y)$. Then,

$$\begin{aligned} U_{mn}^{(N)} &= \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \phi_N(X_i, Y_j) \\ &= \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{N} \lambda_k f_k(X_i) g_k(Y_j) \\ &= \frac{1}{mn} \sum_{k=1}^{N} \lambda_k \left\{ \sum_{i=1}^{m} f_k(X_i) \right\} \left\{ \sum_{j=1}^{n} g_k(Y_j) \right\}. \end{aligned} \tag{4.20}$$

Note that for $k \geqslant 1$,

$$E f_k(X) = E g_k(Y) = 0, \tag{4.21}$$

$$E f_k^2(X) = E g_k^2(Y) = 1, \tag{4.22}$$

and

$$E f_k(X) g_k(Y) = 0. \tag{4.23}$$

Consider two linear combinations

$$T_1 = \sum_{k=1}^{N} \frac{c_k}{\sqrt{m}} \sum_{i=1}^{m} f_k(X_i) \tag{4.24}$$

and

$$T_2 = \sum_{k=1}^{N} \frac{d_k}{\sqrt{n}} \sum_{j=1}^{n} g_k(Y_j). \tag{4.25}$$

In view of (4.9)–(4.12) and following the arguments in Theorem 2.1, we have

$$T_1 \xrightarrow{\mathscr{L}} \mathrm{N}(0, \sigma_{1,N}^2) \quad \text{as } n \to \infty, \tag{4.26}$$

and

$$T_2 \xrightarrow{\mathscr{L}} \mathrm{N}(0, \sigma_{2,N}^2) \quad \text{as } n \to \infty, \tag{4.27}$$

where

$$\begin{aligned} \sigma_{1,N}^2 &= \sum_{k=1}^{N} c_k^2 + \frac{2}{m} \sum_{i=1}^{m-1} (m-i) \operatorname{Cov}(B_{1N}(X_1), B_{1N}(X_{1+i})), \\ \sigma_{2,N}^2 &= \sum_{k=1}^{N} d_k^2 + \frac{2}{n} \sum_{j=1}^{n-1} (n-j) \operatorname{Cov}(B_{2N}(X_1), B_{2N}(X_{1+j})), \end{aligned} \tag{4.28}$$

and

$$B_{1N}(X_i) = \sum_{k=1}^{N} c_k f_k(X_i),$$

$$B_{2N}(Y_j) = \sum_{k=1}^{N} d_k g_k(Y_j). \tag{4.29}$$

Therefore,

$$(mn)^{1/2} U_{mn}^{(N)} \xrightarrow{\mathscr{L}} \sum_{k=1}^{N} \lambda_k U_k V_k \quad \text{as } m, n \to \infty. \tag{4.30}$$

Note that

$$\operatorname{Cov}(U_k, U_{k'}) = \operatorname{Cov}(f_k(X_1), f_{k'}(X_1)) + 2\sum_{i=1}^{\infty} \operatorname{Cov}(f_k(X_1), f_{k'}(X_{1+i})), \tag{4.31}$$

and

$$\operatorname{Cov}(V_k, V_{k'}) = \operatorname{Cov}(g_k(Y_1), g_{k'}(Y_1)) + 2\sum_{i=1}^{\infty} \operatorname{Cov}(g_k(Y_1), g_{k'}(Y_{1+i})), \tag{4.32}$$

Let $\phi_{m,n}(t)$, $\phi^*(t)$, $\phi_N^*(t)$ and $\phi_{m,n,N}(t)$ be the characteristic functions of $(mn)^{1/2} U_{mn}$, $\sum_{k=1}^{\infty} \lambda_k U_k V_k$, $\sum_{k=1}^{n} \lambda_k U_k V_k$ and nU_{mn}^{N}, respectively.

In view of (4.30), for any $\varepsilon > 0$, there exists $m_0(N, \varepsilon, t)$ and $n_0(N, \varepsilon, t)$ such that for $m \geqslant m_0$ and $n \geqslant n_0$,

$$|\phi_{m,n,N}(t) - \phi_N^*(t)| \leqslant \frac{\varepsilon}{3}. \tag{4.33}$$

As in the proof of (2.18),

$$E(U_k^2) < C, \tag{4.34}$$

uniformly for $k \geqslant 1$, and

$$E(V_k^2) < C, \tag{4.35}$$

uniformly for $k \geqslant 1$. In view of (4.34), (4.35) and Cauchy–Schwartz inequality,

$$\begin{aligned} E\left(\sum_{k=N+1}^{\infty} \lambda_k U_k V_k\right)^2 &\leqslant \sum_{k=N+1}^{\infty} \lambda_k^2 E(U_k^2) E(V_k^2) \\ &\quad + \sum_{N+1 \leqslant k \neq k' < \infty} |\lambda_k||\lambda_{k'}| E|U_k U_{k'}| E|V_k V_{k'}| \\ &\leqslant C\left\{\sum_{k=N+1}^{\infty} |\lambda_k|\right\}^2 \\ &\to 0 \quad \text{as } N \to \infty. \end{aligned} \tag{4.36}$$

Therefore, for large N, depending on t,

$$\begin{aligned} |\phi_N^*(t) - \phi^*(t)| &\leqslant |t| E\left(\sum_{k=N+1}^{\infty} \lambda_k U_k V_k\right)^2 \\ &\leqslant \frac{\varepsilon}{3}. \end{aligned} \tag{4.37}$$

Let

$$\bar{\phi}(x,y) = \phi(x,y) - \phi_N(x,y). \tag{4.38}$$

Then, from Birkel (1986), it follows that

$$\begin{aligned} mnE|U_{m,n} - U_{m,n}^{(N)}|^2 &= \frac{1}{mn} E\left(\sum_{i=1}^{m}\sum_{j=1}^{n} \bar{\phi}(X_i,Y_j)\right)^2 \\ &= \frac{1}{mn} \sum_{1\leqslant i\neq i'\leqslant m}\ \sum_{1\leqslant j\neq j'\leqslant n} E(\bar{\phi}(X_i,Y_j)\bar{\phi}(X_i',Y_j')) \\ &\leqslant C\frac{1}{mn} \sum_{1\leqslant i\neq i'\leqslant m}\ \sum_{1\leqslant j\neq j'\leqslant n} \{\mathrm{Cov}(X_i,X_i') + \mathrm{Cov}(Y_j,Y_j')\} \\ &\leqslant C\frac{1}{mn}\left\{n^2 m \sum_{i=1}^{m-1}\mathrm{Cov}(X_1,X_{i+1}) + m^2 n \sum_{j=1}^{n-1}\mathrm{Cov}(Y_1,Y_{j+1})\right\} \\ &= Cmn\left\{\frac{1}{m}\sum_{i=1}^{m-1}\mathrm{Cov}(X_1,X_{i+1}) + \frac{1}{n}\sum_{j=1}^{n-1}\mathrm{Cov}(Y_1,Y_{j+1})\right\} \\ &= mn\{\delta_m + \eta_n\} \\ &\to 0 \quad \text{as } m,n\to\infty. \end{aligned} \tag{4.39}$$

Hence, for large $m \geqslant m_0$ and $n \geqslant n_0, m_0$ and n_0 depending on N, t and ε

$$|\phi_{m,n}(t) - \phi_{m,n,N}(t)| \leqslant \frac{\varepsilon}{3}. \tag{4.40}$$

The result follows by combining (4.33), (4.37) and (4.40).

Theorem 4.3. *Let U_{mn} be a non-degenerate two sample* U*-statistic based on stationary sequences of associated random variables. Let ϕ be the corresponding kernel. Assume that the eigenfunctions $f_k(x)$ and $g_k(y)$ given by* (4.3) *are differentiable and*

$$\sup_k \sup_x |f_k'(x)| < \infty, \tag{4.41}$$

$$\sup_k \sup_y |g_k'(y)| < \infty. \tag{4.42}$$

Furthermore, assume that,

$$\sum_{j=1}^{\infty} \mathrm{Cov}(X_1,X_j) < \infty. \tag{4.43}$$

and

$$\sum_{k=1}^{\infty} |\lambda_k| < \infty. \tag{4.44}$$

Let

$$\delta_m = \frac{1}{m}\sum_{i=1}^{m-1} \mathrm{Cov}(X_1, X_{i+1}) \tag{4.45}$$

and

$$\eta_n = \frac{1}{n}\sum_{j=1}^{n-1} \mathrm{Cov}(Y_1, Y_{j+1}). \tag{4.46}$$

Assume that

$$mn\{\delta_m + \eta_n\} \to 0 \quad \text{as } m,n \to \infty. \tag{4.47}$$

Let

$$g^*(x) = \int_{-\infty}^{\infty} \phi(x,y)\,\mathrm{d}G(y),$$

$$f^*(y) = \int_{-\infty}^{\infty} \phi(x,y)\,\mathrm{d}F(x). \tag{4.48}$$

Suppose that $g^*(.)$ *is monotone or* g^* *is Lipschitzian, and*

$$\sum_{k=0}^{\infty} |Eg_k^*(X_1)| < \infty, \tag{4.49}$$

and similar conditions hold for f^* *with*

$$\sum_{k=0}^{\infty} |Ef_k^*(Y_1)| < \infty, \tag{4.50}$$

If U_{mn} *has finite variance and* $m/n \to \lambda$ *as* $m,n \to \infty$, *then*

$$(m)^{1/2}(U_{mn} - EU_{mn}) \xrightarrow{\mathscr{L}} \mathrm{N}(0, \sigma_1^2 + \lambda\sigma_2^2), \tag{4.51}$$

where

$$\sigma_1^2 = \mathrm{Var}(g^*(X_1)) + 2\sum_{j=1}^{\infty} \mathrm{Cov}(g^*(X_1), g^*(X_{j+1})), \tag{4.52}$$

$$\sigma_2^2 = \mathrm{Var}(f^*(Y_1)) + 2\sum_{j=1}^{\infty} \mathrm{Cov}(f^*(Y_1), f^*(Y_{j+1})), \tag{4.53}$$

Proof. Note that

$$EU_{m,n} = E\phi(X_1, Y_1) = Eg^*(X) = Ef^*(Y). \tag{4.54}$$

Since $U_{m,n}$ has finite variance, we have

$$Eg^{*2}(X_1) = E[\phi(X_1, Y_1)\phi(X_1, Y_2)] < \infty, \tag{4.55}$$

and

$$Ef^{*2}(Y_1) = E[\phi(X_1, Y_1)\phi(X_2, Y_1)] < \infty. \tag{4.56}$$

Let

$$\hat{\phi}(x,y) = \phi(x,y) - g^*(x) - f^*(y) + Eg^*(X_1). \tag{4.57}$$

Then $\hat{\phi}$ is square integrable and degenerate.

Let $\hat{U}_{m,n}$ be the U-statistic based on $\hat{\phi}$. Then, from Theorem 4.2, $(mn)^{1/2}\hat{U}_{m,n}$ converges in distribution. Furthermore, using (4.54) and (4.57)

$$U_{m,n} - EU_{m,n} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{m}\{\phi(X_i, Y_j) - E\phi(X_i, Y_j)\}$$

$$= \hat{U}_{m,n} + \frac{1}{m}\sum_{i=1}^{m}\{g^*(X_i) - Eg^*(X_i)\}$$

$$+\frac{1}{n}\sum_{j=1}^{m}\{f^*(Y_j) - Ef^*(Y_j)\}. \tag{4.58}$$

Therefore,

$$m^{1/2}\{U_{m,n} - EU_{m,n}\} = m^{1/2}\hat{U}_{m,n} + \frac{1}{\sqrt{m}}\sum_{i=1}^{m}\{g^*(X_i) - Eg^*(X_i)\}$$

$$+\left(\frac{m}{n}\right)^{1/2}\frac{1}{\sqrt{n}}\sum_{j=1}^{m}\{f^*(Y_j) - Ef^*(Y_j)\}. \tag{4.59}$$

Therefore, from Theorem 4.2 we get

$$m^{1/2}\hat{U}_{m,n} \xrightarrow{p} 0 \quad \text{as } m, n \to \infty, \tag{4.60}$$

$$\frac{1}{\sqrt{m}}\sum_{i=1}^{m}\{g^*(X_i) - Eg^*(X_i)\} \xrightarrow{\mathscr{L}} \mathrm{N}(0, \sigma_1^2) \quad \text{as } n \to \infty, \tag{4.61}$$

and

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{m}\{f^*(Y_j) - Ef^*(Y_j)\} \xrightarrow{\mathscr{L}} \mathrm{N}(0, \sigma_2^2) \quad \text{as } n \to \infty. \tag{4.62}$$

The result now follows from relation (4.59).

5. V-statistics

Let $\psi(x_1,\ldots,x_k)$ be a real-valued function symmetric in its arguments. Let $E|\psi(x_{i_1},\ldots,x_{i_k})|^r < \infty$ for some positive integer r. Then, the U-Statistic of degree k based on the kernel ψ is defined as

$$U_n = \frac{1}{\binom{n}{k}}\sum_{(c)} \psi(x_{i_1},\ldots,x_{i_k}), \tag{5.1}$$

where $\sum_{(c)}$ denotes the summation over all subsets $1 \leqslant i_1 < \cdots < i_k \leqslant n$ of $\{1,\ldots,n\}$.

A V-statistic (Von Mises, 1947) based on the symmetric kernel ψ of degree k is defined by

$$V_n = n^{-k}\sum_{i_1=1}^{n}\cdots\sum_{i_1=1}^{n} \psi(X_{i_1},\ldots,X_{i_k}). \tag{5.2}$$

Then, one can express V_n as follows (see Lee, 1990, pp. 183):

$$V_n = n^{-k}\sum_{j=1}^{k} j!S_k^{(j)}\binom{n}{j}U_n^{(j)}, \tag{5.3}$$

where $U_n^{(j)}$ is a U-statistic of degree j, $S_k^{(j)}$ are Stirling's numbers of the second kind (see, e.g., Abramowitz and Stegun, 1965).

Also note that

$$n^k = \sum_{j=1}^{k} j! S_k^{(j)} \binom{n}{j}, \tag{5.4}$$

so that

$$n^k(V_n - \theta) = \sum_{j=1}^{k} j! S_k^{(j)} \binom{n}{j} (U_n^{(j)} - \theta). \tag{5.5}$$

The following theorem proves the asymptotic equivalence of the U-statistics and the V-statistics.

Theorem 5.1. *Let $\{X_n, n \geqslant 1\}$ be a stationary associated sequence. Let U_n and V_n be the U-statistic and the V-statistic, respectively, based on a symmetric kernel $\psi(x_1,\ldots,x_k)$ of degree k. Assume that $\psi(x_1,\ldots,x_k)$ is monotonic in x_1. Further suppose that*

$$E[|\psi(X_1,\ldots,X_k)|^{r+\delta}] < \infty \quad for\ r > 2,\ \delta > 0, \tag{5.6}$$

and

$$u_n^* = 2 \sum_{j=n+1}^{\infty} \mathrm{Cov}(\psi(X_1,\ldots,X_k), \psi(X_{(j-1)k+1},\ldots,X_{jk})$$
$$= \mathrm{O}(n^{-[(r-2)(r+\delta)]/2\delta}). \tag{5.7}$$

Then

$$E|U_n - V_n|^r = \mathrm{O}(n^{-r/2}). \tag{5.8}$$

Proof. Let $k \geqslant 1$ and $p = [n/k]$, the greatest integer $\leqslant n/k$.

Define

$$W(x_1,\ldots,x_n) = \frac{1}{p}\{\psi(x_1,\ldots,x_k) + \psi(x_{k+1},\ldots,x_{2k}) + \cdots + \psi(x_{(p-1)k+1},\ldots,x_{pk})\} \tag{5.9}$$

Then

$$\sum_{(n)} W(x_{v_1},\ldots,x_{v_n}) = k!(n-k)! \sum_{(n,k)} \psi(x_{i_1},\ldots,x_{i_k}), \tag{5.10}$$

where $\sum_{(n)}$ is the summation over all the $n!$ permutations $(v_1,\ldots,v_n)$ of $\{1,\ldots,n\}$ and $\sum_{(n,k)}$ is summation over all the $\binom{n}{k}$ subsets $(i_1,\ldots,i_k)$ of $\{1,\ldots,n\}$.

Then, it is easy to see that

$$U_n - \theta = \frac{1}{n!} \sum_{(n)} \{W(X_{v_1},\ldots,X_{v_n}) - \theta\}. \tag{5.11}$$

By Minkowski's inequality and the symmetry property of W, we have

$$E[|U_n - \theta|^r] \leqslant E[|W(X_1,\ldots,X_n) - \theta|^r]. \tag{5.12}$$

Since ψ is monotone component wise, $W(X_1,\dots,X_n)$ is an average of p associated random variables.

Then, using (5.6), (5.7) and the moment inequalities in Birkel (1986), we have

$$E[|U_n-\theta|^r] = \mathrm{O}(p^{-r/2})$$
$$= \mathrm{O}(n^{-r/2}). \tag{5.13}$$

From (5.5), we get that

$$(V_n-\theta) = \sum_{j=1}^{k} a_j(U_n^{(j)}-\theta),$$

where

$$a_j = \frac{j!S_k^{(j)}\binom{n}{j}}{n^k}.$$

Using Minkowski's inequality, and (5.13), we have

$$E|V_n-\theta|^r \leqslant \left\{\sum_{j=1}^{k} a_j(E|U_n-\theta|^r)^{1/r}\right\}^r$$
$$\leqslant C\left\{\sum_{j=1}^{k} a_j n^{-1/2}\right\}^r$$
$$= Cn^{-r/2}\left\{\sum_{j=1}^{k} a_j\right\}^r. \tag{5.14}$$

It is easy to see that

$$a_j = S_k^{(j)}\mathrm{O}(n^{j-k}).$$

Therefore,

$$\sum_{j=1}^{k} a_j \leqslant \max_{1\leqslant j\leqslant k} S_k^{(j)}\mathrm{O}\left(\sum_{j=1}^{k} n^{j-k}\right)$$
$$\leqslant C \max_{1\leqslant j\leqslant k} S_k^{(j)}.$$

Hence,

$$E|V_n-\theta|^r = \mathrm{O}(n^{-r/2}). \tag{5.15}$$

The result follows using (5.13) and (5.15) by the C_r-inequality (Loeve, 1963).

Remark 5.2. If ψ is partially differentiable with bounded partial derivatives, then following Birkel (1988), condition (5.7) can be written as

$$\sum_{j=n+1}^{\infty}\sum_{i=1}^{k}\sum_{\ell=(j-1)k+1}^{jk} \mathrm{Cov}(X_i,X_\ell) = \mathrm{O}(n^{-[(r-2)(r+\delta)]/2\delta}). \tag{5.16}$$

Remark 5.3. The condition on component-wise monotonicity of ψ can possibly be dropped by extending the results of Shao and Yu (1996) to functions of random vectors of associated random variables.

Acknowledgements

The authors thank Prof. P.K. Sen and a referee for their suggestions to extend the results in Section 2 to kernels of higher degree and to V-statistics.

References

Abramowitz, M., Stegun, I., 1965. Handbook of Mathematical Functions. Dover, New York.

Bagai, I., Prakasa Rao, B.L.S., 1995. Kernel-type density and failure rate estimation for associated sequences. Ann. Inst. Statist. Math. 47, 253–266.

Birkel, T., 1986. Momentenabschatzungen und Grenzwertsatze fur Partialsummen assoziierter Zufalls-variablen. Ph.D. Dissertation. University of Koln.

Birkel, T., 1988. Moment bounds for associated sequences. Ann. Probab. 16, 1184–1193.

Eagleson, G.K., 1979. Orthogonal expansions and U-statistics. Austral. J. Statist. 21, 221–237.

Esary, J., Proschan, F., Walkup, D., 1967. Association of random variables with applications. Ann. Math. Statist. 38, 1466–1474.

Gregory, G.G., 1977. Large sample theory for U-statistics and tests of fit. Ann. Statist. 5, 110–123.

Hall, P., 1979. An invariance theorem for U-statistics. Stoch. Proc. Appl. 9, 163–174.

Lee, A.J., 1990. U-Statistics. Marcel Dekker, New York.

Loeve, M., 1963. Probability Theory. Van Nostrand, New York.

Newman, C.M., 1980. Normal fluctuations and the FKG inequalities. Comm. Math. Phys. 74, 119–128.

Newman, C.M., 1984. Asymptotic independence and limit theorems for positively and negatively dependent random variables. In: Tong, Y.L. (Ed.), Inequalities in Statistics and Probability. IMS, Hayward, pp. 127–140.

Prakasa Rao, B.L.S., Dewan, I., 2000. Associated sequences and related inference problems. In: Rao, C.R., Shanbhag, D.N. (Eds.), Handbook of Statistics, Stochastic Processes: Theory and Methods. North-Holland, Amsterdam.

Shao, Q.M., Yu, H., 1996. Weak convergences of weighted empirical processes of dependent sequences. Ann. Probab. 24, 2098–2127.

Von Mises, R., 1947. On the asymptotic distribution of differentiable statistical functions. Ann. Math. Statist. 18, 301–348.