ASYMPTOTIC DISTRIBUTIONS OF LB-STATISTICS AND V-STATISTICS FOR DEGENERATE KERNEL

Yamato, Hajime
Department of Mathematics and Computer Science, Kagoshima University

Toda, Koichiro
Graduate School of Science and Engineering, Kagoshima University

https://doi.org/10.5109/13501

出版情報：Bulletin of informatics and cybernetics. 33 (1/2), pp.27-42, 2001-12. Research Association of Statistical Sciences
バージョン：

Bulletin of Informatics and Cybernetics, Vol. 33, No. 1~2, 2001

ASYMPTOTIC DISTRIBUTIONS OF LB-STATISTICS AND V-STATISTICS FOR DEGENERATE KERNEL

By

Hajime YAMATO* and **Koichiro TODA**†

Abstract

Associated with an estimable parameter, we consider V-statistic and limit of Bayes estimate which we abbreviate LB-statistic. We assume that the kernel of the estimable parameter is degenerate. With respect to the expression of these statistics as a linear combination of U-statistics, we show the properties of its components which are unbiasedness and degeneracy. Using these properties, we give their asymptotic distributions. For the V-statistic, its asymptotic distribution is well-known (see, for example, Borovskikh (1996), p.113). But our expression is different from it. We confirm these are equivalent.

Key Words and Phrases: Linear combination of U-statistics, order of degeneracy, U-statistic, V-statistic.

1. Introduction

Let $\theta(F)$ be an estimable parameter of an unknown distribution F which has a symmetric kernel $g(x_1, ..., x_k)$ of degree k and $X_1, \ldots, X_n$ be a random sample of size n from the distribution F. U-statistic U_n corresponding to the kernel g is given by

$$U_n = \binom{n}{k}^{-1} \sum_{1 \le i_1 < \cdots < i_k \le n} g(X_{i_1}, \ldots, X_{i_k}) \tag{1.1}$$

where $\sum_{1 \le i_1 < \cdots < i_k \le n}$ denotes the summation over all integers $i_1, \ldots, i_k$ satisfying $1 \le i_1 < \cdots < i_k \le n$. V-statistic V_n corresponding to the kernel g is given by

$$V_n = \frac{1}{n^k} \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} g(X_{i_1}, \ldots, X_{i_k}). \tag{1.2}$$

(See, for example, Lee (1990), and Koroljuk and Borovskich (1994).)

We consider an estimator of $\theta(F)$ which is obtained by averaging the kernel g over all unordered arrangements of k X's chosen, allowing repetition, from $X_1, \ldots, X_n$. This estimator can be written as

$$B_n = \binom{n+k-1}{k}^{-1} \sum_{1 \le i_1 \le \cdots \le i_k \le n} g(X_{i_1}, \ldots, X_{i_k})$$

* Department of Mathematics and Computer Science, Kagoshima University, Kagoshima 890–0065, Japan.

† Graduate School of Science and Engineering, Kagoshima University, Kagoshima 890–0065, Japan.

where $\sum_{1\le i_1\le\cdots\le i_k\le n}$ denotes the summation over all integers $i_1,\dots,i_k$ satisfying $1\le i_1\le\cdots\le i_k\le n$. Alternatively this estimator is written as

$$B_n=\binom{n+k-1}{k}^{-1}\sum_{r_1+\cdots+r_n=k} g(X_1^{r_1},\dots,X_n^{r_n}),\tag{1.3}$$

where $\sum_{r_1+\cdots+r_n=k}$ denote the summation over all non-negative integers $r_1,\dots,r_n$ satisfying $r_1+\cdots+r_n=k$ and

$$g(x_1^{r_1},\dots,x_n^{r_n})=g(\underbrace{x_1,\dots,x_1}_{r_1},\dots,\underbrace{x_n,\dots,x_n}_{r_n}).$$

We call this statistic B_n as LB-statistic, because it was originally obtained as the limit of Bayes estimate of $\theta(F)$ (Yamato (1977)). It has the representation as a linear combination of U-statistics as follows.

$$B_n=\binom{n+k-1}{k}^{-1}\sum_{j=1}^{k}\binom{k-1}{j-1}\binom{n}{j}U_n^{(j)},\tag{1.4}$$

where for $j=1,\dots,k$, $U_n^{(j)}$ is the U-statistic corresponding to the kernel given by

$$g_{(j)}(x_1,x_2,\dots,x_j)=\binom{k-1}{j-1}^{-1}\sum\nolimits^{+}_{r_1+\cdots+r_j=k} g(x_1^{r_1},\dots,x_j^{r_j}),\tag{1.5}$$

where $\sum^{+}_{r_1+\cdots+r_j=k}$ denote the summation over all positive integers $r_1,\dots,r_j$ satisfying $r_1+\cdots+r_j=k$. Especially, we have

$$g_{(k)}(x_1,x_2,\dots,x_k)=g(x_1,x_2,\dots,x_k),\quad U_n^{(k)}=U_n,$$

$$g_{(k-1)}(x_1,x_2,\dots,x_{k-1})=\frac{1}{k-1}\{g(x_1,x_1,x_2,x_3,\dots,x_{k-1})$$
$$+g(x_1,x_2,x_2,x_3,x_4,\dots,x_{k-1})+\cdots+g(x_1,\dots,x_{k-3},x_{k-2},x_{k-1},x_{k-1})\}.$$

(See Nomachi and Yamato (2001)).

V-statistic V_n can be written as a linear combination of U-statistics as follows.

$$V_n=\frac{1}{n^k}\sum_{j=1}^{k}S(k,j)(n)_jU_n^{*(j)}\tag{1.6}$$

where $(n)_j=n(n-1)\cdots(n-j+1)$ for $j=1,\dots,k$, $U_n^{*(j)}$ is the U-statistic corresponding to the kernel given by

$$g^*_{(j)}(x_1,\dots,x_j)=\frac{1}{j!S(k,j)}\sum\nolimits^{+}_{r_1+\cdots+r_j=k}\frac{k!}{r_1!\cdots r_j!}g(x_1^{r_1},\dots,x_j^{r_j}),\tag{1.7}$$

and $S(k,j)$ are the Stirling numbers of the second kind (see, for example, Lee (1990) and Koroljuk and Borovskich (1994)) and $g^*_{(k)}=g_{(k)}=g$, $g^*_{(k-1)}=g_{(k-1)}$.

In this paper we consider the case that the kernel g is degenerate. For the kernel $g(x_1, \ldots, x_k)$, we put

$$\psi_j(x_1, \ldots, x_j) = E[g(X_1, \ldots, X_k) \mid X_1 = x_1, \ldots, X_j = x_j], \quad j = 1, ..., k$$

$$\sigma_j^2 = Var[\psi_j(X_1, \ldots, X_j)], \quad j = 1, ..., k.$$

We suppose that

$$\sigma_1^2 = \cdots = \sigma_{d-1}^2 = 0 \quad \text{and} \quad \sigma_d^2 > 0,$$

that is, the U-statistic and/or the kernel g is degenerate of order $d-1$. Hence $E\psi_d(X_1, \ldots, X_d) = \theta$ and with probability one (w.p.1) $\psi_1(X_1) = \theta, \ldots, \psi_{d-1}(X_1, \ldots, X_{d-1}) = \theta$.

For this degenerate kernel, under the conditions $E \mid g(X_1, \ldots, X_k) \mid^2 < \infty$, the asymptotic distribution of U_n is given by

$$n^{d/2}(U_n - \theta) \xrightarrow{\mathcal{D}} \binom{k}{d} J_d(\xi_{d,d}), \tag{1.8}$$

where $\xrightarrow{\mathcal{D}}$ means the convergence in distribution as $n \to \infty$ and $\xi_{d,d}(x_1, ..., x_d) = \psi_d(x_1, ..., x_d) - \theta$. $J_c(f)$ can be written by two expressions. Let W be the Gaussian random measure associated with the distribution F on the real line $(-\infty, \infty)$ such that $EW(A) = 0$ and $EW(A)W(B) = F(A \cap B)$ for any Borel sets A, B (see, for example, Kotani (1997), p.237). The one expression is the stochastic integral given by

$$J_c(f) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, ..., x_c) W(dx_1) \cdots W(dx_c).$$

For any function f_1 and f_2 such that $\int_{R^c} f_i(x_1, \ldots, x_c)^2 \prod_{j=1}^{c} dF(x_j) < \infty$ $(i = 1, 2)$, their inner product is given by

$$(f_1, f_2) = \int_{R^c} f_1(x_1, \ldots, x_c) f_2(x_1, \ldots, x_c) \prod_{j=1}^{c} dF(x_j).$$

By using an orthonormal basis $e_1, e_2, ...$ of $L_2(F)$, $J_c(f)$ is also written as

$$J_c(f) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_c=1}^{\infty} (f, e_{i_1} \cdots e_{i_c}) \prod_{l=1}^{\infty} H_{r_l(\mathbf{i})}(Z_l),$$

where H_r is the r-th Hermite polynomial, $\{Z_l\}_{l=1}^{\infty}$ is a sequence of independent standard normal random variables and $r_l(\mathbf{i})$ is the number of indices among $\mathbf{i} = (i_1, ..., i_d)$ equal to l (see, for example, Lee (1990), and Koroljuk and Borovskich (1994)). The above convergence is also shown under the following conditions: (i) $E \mid g^{(c)}(X_1, \ldots, X_d) \mid^{2c/(2c-d)} < \infty$ for $c = d, d+1, \ldots, k$ (Koroljuk and Borovskich (1994)). (ii) $E \mid g^{(d)}(X_1, \ldots, X_d) \mid^2 < \infty$ and $t^{2c/(2c-d)} P[\mid g^{(c)} \mid > t] \to 0$ $(t \to \infty)$ for $c = d+1, \ldots, k$. (Borovskich (1996)). Since we must consider the convergence of certain degenerate U-statistics jointly in this paper, we assume $E[g^2(X_{j_1}, X_{j_2}, ..., X_{j_k})] < \infty$ for all $j_1, j_2, ..., j_k$ $(1 \le j_1 \le j_2 \le \cdots \le j_k \le k)$.

In Section 2, we show the properties of the U-statistics $U_n^{(j)}$ having the kernel given by (1.5), which are unbiasedness and degeneracy . Using these and the relation (1.4) we give the asymptotic distribution of LB-statistic B_n.

In Section 3, we give the properties of the U-statistics $U_n^{*(j)}$ having the kernel given by (1.7), which are unbiasedness and degeneracy. Using these we give the asymptotic distribution of V-statistic V_n. This asymptotic distribution of V_n is derived using the linear combination of U-statistic given by (1.6).

In Section 4, we note the H-decomposition of U-statistic and the corresponding decomposition of the V-statistic. Using this decomposition we can also derive the asymptotic distribution of V-statistic V_n. It is well-known that the asymptotic distribution of V_n is given by (4.3) of Section 4. Our expression is different from this. So we confirm these two expressions are equivalent.

In Section 5, we give some examples for our results. In Section 6 which is Appendix, the proofs of some Lemmas and Propositions in the previous sections are given. Rubin and Vitale (1980) obtains the asymptotic distribution of symmetric statistic which includes V-statistic and LB-statistic. Their method is based on the fact that any function which belongs to $L_2(F)$ can be written by an orthonormal basis for $L_2(F)$. Differently from this, we obtain the asymptotic distribution of LB-statistic and V-statistic, using their expression by linear combinations of U-statistics.

2. Asymptotic distribution of LB-statistic

For the kernel $g_{(j)}(x_1,\dots,x_j)$ given by (1.5), we put for $c=1,\dots,j$ and $j=1,\dots,k$

$$\begin{aligned}\psi_{(j),c}(x_1,...,x_c) &= E[g_{(j)}(X_1,...,X_j)\mid X_1=x_1,...,X_c=x_c]\\ &= \binom{k-1}{j-1}^{-1}\sum\nolimits^{+}_{r_1+\cdots+r_j=k} Eg(x_1^{r_1},...,x_c^{r_c},X_{c+1}^{r_{c+1}},...,X_j^{r_j}),\end{aligned}$$

where on the right-hand side we use the notation used for (1.3). Using these notations, we show the properties of $U_n^{(j)}$ corresponding to the kernel $g_{(j)}$.

LEMMA 2.1.

$$E[U_n^{(j)}]=\theta,\quad k-\frac{d-1}{2}\le j\le k$$

or

$$E[U_n^{(k-j)}]=\theta,\quad 0\le j\le\frac{d-1}{2}.$$

From this lemma, the bias of B_n is

$$EB_n-\theta=\binom{n+k-1}{k}^{-1}\sum_{1\le j<k-\frac{d-1}{2}}\binom{k-1}{j-1}\binom{n}{j}[EU_n^{(j)}-\theta].$$

If $d=2l+1$ and l is a positive integer, then the summation of the right-hand side is taken over $j=1,2,...,k-l-1$ and $EB_n-\theta=O(n^{-l-1})$. If $d=2l$ and l is a positive integer, then the summation of the right-hand side is taken over $j = 1, 2, \ldots, k-l$ and $EB_n - \theta = O(n^{-l})$. Thus we have

$$EB_n - \theta = O(n^{-[\frac{d+1}{2}]}),$$

where $[x]$ denotes the greatest integer not exceeding x.

LEMMA 2.2. *The order of degeneracy of $U_n^{(k-j)}$ is at least $d-2j-1$ for $1 \le j \le (d-1)/2$ and*

$$\psi_{(k-j),d-2j}(x_1, \ldots, x_{d-2j}) = \binom{k-1}{j}^{-1} \{[\binom{k-1}{j} - \binom{k-d+j}{j}]\theta \tag{2.1}$$

$$+\binom{k-d+j}{j}\varphi_{d,d-2j}(x_1, \ldots, x_{d-2j})\},$$

where for $1 \le j \le (d-1)/2$

$$\varphi_{d,d-2j}(x_1, \ldots, x_{d-2j}) = E[\psi_d(x_1, \ldots, x_{d-2j}, X_{d-2j+1}, X_{d-2j+1}, \ldots, X_{d-j}, X_{d-j})]. \tag{2.2}$$

Since $E\varphi_{d,d-2j}(X_1, \ldots, X_{d-2j}) = EE[\psi_d(X_1, \ldots, X_{d-2j}, X_{d-2j+1}, X_{d-2j+1}, \ldots, X_{d-j}, X_{d-j} \mid X_{d-2j+1}, \ldots, X_{d-j}] = E\psi_{2j}(X_{d-2j+1}, X_{d-2j+1}, \ldots, X_{d-j}, X_{d-j})$, we have for $1 \le j \le (d-1)/2$

$$E\varphi_{d,d-2j}(X_1, \ldots, X_{d-2j}) = \theta. \tag{2.3}$$

Before stating the asymptotic distribution of B_n, we note the followings which are the results of Lemmas 2.1 and 2.2: If $d = 2l+1$ and l is a positive integer, then $EU_n^{(k)} = EU_n^{(k-1)} = \cdots = EU_n^{(k-l+1)} = EU_n^{(k-l)} = \theta$. The orders of degeneracy of $U_n^{(k-1)}, \ldots, U_n^{(k-l+1)}$, $U_n^{(k-l)}$ are at least $2(l-1), \ldots, 2, 0$, respectively. If $d = 2l$ and l is a positive integer, then $EU_n^{(k)} = EU_n^{(k-1)} = \cdots = EU_n^{(k-l+2)} = EU_n^{(k-l+1)} = \theta$. The orders of degeneracy of $U_n^{(k-1)}, \ldots, U_n^{(k-l+2)}, U_n^{(k-l+1)}$ are at least $2l-3, \ldots, 3, 1$, respectively.

PROPOSITION 2.3. *We suppose that $d = 2l+1$ where l is a positive integer and $E[g^2(X_{j_1}, X_{j_2}, \ldots, X_{j_k})] < \infty$ for all $j_1, j_2, \ldots, j_k$ $(1 \le j_1 \le j_2 \le \cdots \le j_k \le k)$. Then we have*

$$n^{d/2}(B_n - \theta) \xrightarrow{\mathcal{D}} \frac{k!}{(k-d)!}\sum_{j=0}^{l}\frac{1}{(d-2j)!j!}J_{d-2j}(\xi_{d,d-2j}), \tag{2.4}$$

where

$$\xi_{d,d-2j}(x_1, \ldots, x_{d-2j}) = \varphi_{d,d-2j}(x_1, \ldots, x_{d-2j}) - \theta \quad (0 \le j \le \frac{d-1}{2}).$$

LEMMA 2.4. *In case of $d = 2l$,*

$$EU_n^{(k-l)} - \theta = \binom{k-1}{l}^{-1}\binom{k-l}{k-d}[E\psi_d(X_1, X_1, \ldots, X_l, X_l) - \theta]. \tag{2.5}$$

PROPOSITION 2.5. *We suppose that $d = 2l$ where l is a positive integer and $E[g^2(X_{j_1}, X_{j_2}, ..., X_{j_k})] < \infty$ for all $j_1, j_2, ..., j_k$ $(1 \le j_1 \le j_2 \le \cdots \le j_k \le k)$. Then we have*

$$n^{d/2}(B_n - \theta) \xrightarrow{\mathcal{D}} \frac{k!}{(k-d)!}\{\sum_{j=0}^{l-1} \frac{1}{(d-2j)!j!} J_{d-2j}(\xi_{d,d-2j}) + \frac{1}{l!}[E\psi_d(X_1, X_1, ..., X_l, X_l) - \theta]\}. \tag{2.6}$$

3. Asymptotic distribution of V-statistic

For the kernel $g^*_{(j)}(x_1, ..., x_j)$ given by (1.7), we put for $c = 1, ..., j$ and $j = 1, ..., k$

$$\psi^*_{(j),c}(x_1, ..., x_c) = E[g^*_{(j)}(X_1, ..., X_j) \mid X_1 = x_1, ..., X_c = x_c] \tag{3.1}$$

$$= \frac{1}{j!S(k,j)} \sum\nolimits^{+}_{r_1+\cdots+r_j=k} \frac{k!}{r_1! \cdots r_j!} g(x_1^{r_1}, ..., x_c^{r_c}, X_{c+1}^{r_c+1}, ..., X_j^{r_j}),$$

where on the right-hand side we use the notation used for (1.3). By the same methods as the proof of Lemmas 2.1 and 2.2, we can show the properties of $U_n^{*(j)}$ corresponding to the kernel $g^*_{(j)}$, which are the following Lemmas 3.1 and 3.2.

LEMMA 3.1.

$$E[U_n^{*(j)}] = \theta, \quad k - \frac{d-1}{2} \le j \le k$$

or

$$E[U_n^{*(k-j)}] = \theta, \quad 0 \le j \le \frac{d-1}{2}.$$

We can prove Lemma 3.2, using the relation (3.1) replacing j with $k - j$ for $c = 1, 2, ..., d - 2j$.

LEMMA 3.2. *The order of degeneracy of $U_n^{*(k-j)}$ is at least $d - 2j - 1$ for $1 \le j \le (d-1)/2$ and*

$$\psi^*_{(k-j),d-2j}(x_1, ..., x_{d-2j}) = \frac{1}{(k-j)!S(k,k-j)}\{[(k-j)!S(k,k-j) - \frac{k!}{2^j}\binom{k-d+j}{j}]\theta$$
$$+ \frac{k!}{2^j}\binom{k-d+j}{j}\varphi_{d,d-2j}(x_1, ..., x_{d-2j})\}. \tag{3.2}$$

$\varphi_{d,d-2j}(x_1, ..., x_{d-2j})$ is given by (2.2). Similarly to B_n, the bias of V_n is $EV_n - \theta = n^{-k} \sum_{1 \le j < k-(d-1)/2} S(k, k-j)(n)_{k-j}\ [EU_n^{*(k-j)} - \theta]$ and $EV_n - \theta = O(n^{-[(d+1)/2]})$.

Before stating the asymptotic distribution of V_n, we note the followings which are the results of Lemmas 3.1 and 3.2: If $d = 2l + 1$ and l is a positive integer, then $EU_n^{*(k)} = EU_n^{*(k-1)} = \cdots = EU_n^{*(k-l+1)} = EU_n^{*(k-l)} = \theta$. The order of degeneracy of $U_n^{*(k-1)}$,...,$U_n^{*(k-l+1)}$,$U_n^{*(k-l)}$ are at least $2(l-1), ..., 2, 0$, respectively. If $d = 2l$ and l is a positive integer, then $EU_n^{*(k)} = EU_n^{*(k-1)} = \cdots = EU_n^{*(k-l+2)} = EU_n^{*(k-l+1)} = \theta$. The order of degeneracy of $U_n^{*(k-1)}$,...,$U_n^{*(k-l+2)}$,$U_n^{*(k-l+1)}$ are at least $2l-3, ..., 3, 1$, respectively.

By Lemmas 3.1 and 3.2, using (1.6) we can show the following two propositions. The methods of proofs are as same as that of Propositions 2.3 and 2.5. Here we use the relation

$$n^{d/2}(V_n - \theta) = \sum_{j=0}^{k-1} \mathcal{S}(k, k-j) \frac{(n)_{k-j}}{n^k} n^{d/2}[U_n^{*(k-j)} - \theta],$$

which is obtained from (1.6).

PROPOSITION 3.3. *We suppose that* $d = 2l + 1$ *where* l *is a positive integer and* $E[g^2(X_{j_1}, X_{j_2}, ..., X_{j_k})] < \infty$ *for all* $j_1, j_2, ..., j_k$ $(1 \le j_1 \le j_2 \le \cdots \le j_k \le k)$. *Then we have*

$$n^{d/2}(V_n - \theta) \xrightarrow{\mathcal{D}} \frac{k!}{(k-d)!} \sum_{j=0}^{l} \frac{1}{(d-2j)! j! 2^j} J_{d-2j}(\xi_{d,d-2j}).$$

The following lemma is used for Proposition 3.5.

LEMMA 3.4. *In case of* $d = 2l$,

$$EU_n^{*(k-l)} - \theta = \frac{k!}{\mathcal{S}(k, k-l) l! (k-d)! 2^l} [E\psi_d(X_1, X_1, ..., X_l, X_l) - \theta].$$

PROPOSITION 3.5. *We suppose that* $d = 2l$ *where* l *is a positive integer and* $E[g^2(X_{j_1}, X_{j_2}, ..., X_{j_k})] < \infty$ *for all* $j_1, j_2, ..., j_k$ $(1 \le j_1 \le j_2 \le \cdots \le j_k \le k)$. *Then we have*

$$n^{d/2}(V_n - \theta)$$

$$\xrightarrow{\mathcal{D}} \frac{k!}{(k-d)!} \{\sum_{j=0}^{l-1} \frac{1}{(d-2j)! j! 2^j} J_{d-2j}(\xi_{d,d-2j}) + \frac{1}{l! 2^l} [E\psi_d(X_1, X_1, ..., X_l, X_l) - \theta]\}.$$

4. Decompositions of statistics

It is well-known that U-statistic of degree k can be written as a linear combination of U-statistics of degrees $1, 2, ..., k$. We put

$$h^{(1)}(x_1) = \psi_1(x_1) - \theta$$

and

$$h^{(c)}(x_1, ..., x_c) = \psi_c(x_1, ..., x_c) - \sum_{j=1}^{c-1} \sum_{(c,j)} h^{(j)}(x_{i_1}, ..., x_{i_j}) - \theta, \quad c = 2, ..., k, \tag{4.1}$$

where the summation $\sum_{(c,j)}$ is taken over all subsets $1 \le i_1 < \cdots < i_j \le c$ of $\{1, 2, ..., c\}$. For $j = 1, 2, ..., k$, let $H_n^{(j)}$ be the U-statistic with the kernel $h^{(j)}$. Then $U_n = \theta + \sum_{j=1}^{k} \binom{k}{j} H_n^{(j)}$. The degree of $H_n^{(j)}$ and/or $h^{(j)}$ is j and the order of degeneracy is $j - 1$, that is $H_n^{(j)}$ and/or $h^{(j)}$ is completely degenerate. Let $K_n^{(j)}$ be the V-statistic based on the kernel $h^{(j)}$. Then

$$V_n = \theta + \sum_{j=1}^{k} \binom{k}{j} K_n^{(j)}.$$

If the order of degeneracy of U_n is $d-1$, then

$$U_n = \theta + \sum_{j=d}^{k} \binom{k}{j} H_n^{(j)}, \quad V_n = \theta + \sum_{j=d}^{k} \binom{k}{j} K_n^{(j)}. \tag{4.2}$$

(See, for example, Lee (1990) and Borovskikh (1995).) Now, as we stated in Section 1, we consider the case that $E\psi_d(X_1, ..., X_d) = \theta$ and

$$\psi_1(X_1) = \theta, ..., \psi_{d-1}(X_1, ..., X_{d-1}) = \theta \quad \text{w.p.1},$$

that is, the U-statistic and/or the kernel g is degenerate of order $d-1$.
In the following, we suppose that $E[g^2(X_{j_1}, X_{j_2}, ..., X_{j_k})] < \infty$ for all $j_1, j_2, ..., j_k$ $(1 \leq j_1 \leq j_2 \leq \cdots \leq j_k \leq k)$. Then it holds that $E[\psi_c^2(X_{j_1}, X_{j_2}, ..., X_{j_c})] < \infty$ for all $j_1, j_2, ..., j_c$ $(1 \leq j_1 \leq j_2 \leq \cdots \leq j_c \leq c \leq k)$ and therefore $E[h^{(c)}(X_{j_1}, X_{j_2}, ..., X_{j_c})]^2 < \infty$ for all $j_1, j_2, ..., j_c$ $(1 \leq j_1 \leq j_2 \leq \cdots \leq j_c \leq c)$.

It is well-known that the asymptotic distribution of V-statistic is given by

$$\binom{k}{d} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \{\psi_d - \theta\} Q(dx_1) \cdots Q(dx_d), \tag{4.3}$$

where Q is a centered Gaussian random measure with covariance function $EQ(A)Q(B) = F(A \cap B) - F(A)F(B)$ for any Borel sets A, B (see, for example, Borovskikh (1996), p.113). This is derived by using the decomposition of V_n given by (4.2) and the convergence of the empirical measure. We can derive it not using this convergence and confirm the two form are equivalent.

Differently from the method using the convergence of the empirical measure, we use Propositions 3.3 and 3.5 in case of completely degenerate kernel. Since for $c = d, d+1, ..., k$ $h^{(c)}$ is completely degenerate, by applying Propositions 3.3 and 3.5 to $K_n^{(c)}$ we can get the convergence of $n^{c/2}K_n^{(c)}$ in distribution for $c = d, d+1, ..., k$. Therefore for $c = d+1, ..., k$, $n^{d/2}K_n^{(c)}$ converges to zero in probability as $n \to \infty$. Since the kernel of $K_n^{(d)}$ is $\psi_d - \theta$, the asymptotic distribution of $n^{d/2}K_n^{(d)}$ is given by Propositions 3.3 and 3.5 replacing k by d. Thus by (4.2), we know that the asymptotic distribution of V-statistic is given by Propositions 3.3 and 3.5. Hence we could confirm that the two approaches give the same asymptotic distribution. Furthermore we know that the asymptotic distribution of $K_n^{(d)}$ does not depend on the terms associated with the kernels of the lower degrees among the representation of $K_n^{(d)}$ by a linear combination of U-statistics.

5. Examples

Propositions 2.3, 2.5, 3.3, and 3.5 with $l = 1$ give the followings: For the first-order of degeneracy, that is $\sigma_1^2 = 0$, $\sigma_2^2 > 0$ and $d = 2$, the asymptotic distribution of $n(B_n - \theta)$ is

$$k(k-1)\{\frac{1}{2}J_2(\xi_{2,2}) + [E\psi_2(X, X) - \theta]\}, \tag{5.1}$$

and the asymptotic distribution of $n(V_n - \theta)$ is

$$\frac{k(k-1)}{2}\{J_2(\xi_{2,2}) + [E\psi_2(X,X) - \theta]\} \tag{5.2}$$

(see, for example, Lee (1990) and Koroljuk and Borovskich (1994)).
In this case, the another form of $J_2(\xi_{2,2})$ is well-known. Let Z_1, Z_2,... be independently and identically distributed standard normal random variables and λ_j $(j = 1, 2, \ldots)$ be the eigenvalues of the integral equation

$$\int_{-\infty}^{\infty}(\psi_2(x_1, x_2) - \theta)f(x_2)dF(x_2) = \lambda f(x_1).$$

Then

$$J_2(\xi_{2,2}) = \sum_{j=1}^{\infty}\lambda_j(Z_j^2 - 1) \tag{5.3}$$

(see, for example, Lee (1990) and and Koroljuk and Borovskich (1994)).

For the second-order of degeneracy, that is $\sigma_1^2 = \sigma_2^2 = 0$, $\sigma_3^2 > 0$ and $d = 3$, the asymptotic distribution of $n^{3/2}(B_n - \theta)$ is

$$k(k-1)(k-2)\{\frac{1}{6}J_3(\xi_{3,3}) + J_1(\xi_{3,1})\},$$

and the asymptotic distribution of $n^{3/2}(V_n - \theta)$ is

$$\frac{k(k-1)(k-2)}{6}\{J_3(\xi_{3,3}) + 3J_1(\xi_{3,1})\}.$$

In the following Examples 1 and 2, we consider the kernels given by Lee (1990), p. 78 and p. 79, respectively.

Example 1. Let $g(x_1, x_2) = f(x_1)f(x_2)$ and assume $Ef(X) = 0$, $Ef^2(X) < \infty$. Then we have $\theta(F) = 0$, $\psi_1(x_1) = 0$ and $\psi_2(x_1, x_2) = f(x_1)f(x_2)$. Because of $d = k = 2$, from (2.2) we have $\xi_{2,2}(x_1, x_2) = \psi_{2,2}(x_1, x_2) = f(x_1)f(x_2)$. Hence,

$$\begin{aligned} J(\xi_{2,2}) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1)f(x_2)W(dx_1)W(dx_2) \\ &= \tau^2 H_2(\tau^{-1}\int f(x)W(dx)) = \tau^2(Z^2 - 1), \end{aligned}$$

where the random variable $Z = \tau^{-1}\int f(x)W(dx)$ has the standard normal distribution and $\tau^2 = Ef^2(X)$. Thus by (5.1) and (5.2), we get

$$nB_n \xrightarrow{\mathcal{D}} \tau^2(Z^2 - 1) + 2\sigma^2 = \tau^2(Z^2 + 1),$$

$$nV_n \xrightarrow{\mathcal{D}} \tau^2(Z^2 - 1) + \sigma^2 = \tau^2 Z^2.$$

Compare with $nU_n \xrightarrow{\mathcal{D}} \tau^2(Z^2 - 1)$ (Lee (1990), p.78).

Example 2. Let $g(x_1, x_2) = af(x_1)f(x_2)+bg(x_1)g(x_2)$. We assume that $Ef(X) = Eg(X) = Ef(X)g(X) = 0$ and $Ef^2(X) = Eg^2(X) = 1$. Then we have $\theta(F) = 0$, $\psi_1(x_1) = 0$, $\psi_2(x_1, x_2) = g(x_1, x_2)$ and $\sigma_2^2 = a^2 + b^2$ (see, Lee (1990), p.79). Because of $d = k = 2$, from (2.2) we have $\xi_{2,2}(x_1, x_2) = af(x_1)f(x_2) + bg(x_1)g(x_2)$ and

$$J(\xi_{2,2}) = a\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1)f(x_2)W(dx_1)W(dx_2)$$

$$+b\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1)g(x_2)W(dx_1)W(dx_2) = a(Z_1^2 - 1) + b(Z_1^2 - 1),$$

where $Z_1 = \int f(x)W(dx)$ and $Z_2 = \int g(x)W(dx)$ are the independent standard normal random variables. Thus by (5.1), (5.2) and $E\psi_2(X, X) = a + b$, we have

$$nB_n \xrightarrow{\mathcal{D}} a(Z_1^2 - 1) + b(Z_2^2 - 1) + 2(a + b) = a(Z_1^2 + 1) + b(Z_2^2 + 1),$$

$$nV_n \xrightarrow{\mathcal{D}} aZ_1^2 + bZ_2^2.$$

Compare with $nU_n \xrightarrow{\mathcal{D}} a(Z_1^2 - 1) + b(Z_2^2 - 1)$ (Lee (1990), p.79).

Example 3(Anderson-Darling Statistic.) We consider the kernel,

$$g(x, y) = \int_0^1 w(t)[I(x \le t) - t][I(y \le t) - t]dt,$$

where $w(t) = [t(1 - t)]^{-1}$, $I(x \le t) = 1$ if $x \le t$ and $= 0$ if $x > t$, and the distribution F is the uniform distribution on $(0, 1)$. Then $\theta(F) = 0$, $\psi_1(x_1) = 0$, $\psi_2(x_1, x_2) = g(x_1, x_2)$ and $E\psi_2(X, X) = 1$. The eigenvalues of $g(x, y)$ is

$$\lambda_j = \frac{1}{j(j + 1)}, \quad j = 1, 2, \dots .$$

(see, for example, Borovskikh (1985)). Thus by (5.1), (5.2) and (5.3), we have

$$nB_n \xrightarrow{\mathcal{D}} \sum_{j=1}^{\infty}\lambda_j(Z_j^2 - 1) + 2 = \sum_{j=1}^{\infty}\frac{1}{j(j + 1)}Z_j^2 + 1,$$

$$nV_n \xrightarrow{\mathcal{D}} \sum_{j=1}^{\infty}\frac{1}{j(j + 1)}Z_j^2.$$

The asymptotic distribution of nV_n is well-known (see, for example, Pettitt (1981)).

Example 4 (Cramér-von Mises Statistic.) We consider the kernel,

$$g(x, y) = \int_{-\infty}^{\infty}[I(x \le t) - F(t)][I(y \le t) - F(t)]dF(t).$$

Then $\theta(F) = 0$, $\psi_1(x_1) = 0$, $\psi_2(x_1, x_2) = g(x_1, x_2)$ and $E\psi_2(X, X) = 1/6$. The eigenvalues of $g(x, y)$ is

$$\lambda_j = \frac{1}{(j\pi)^2}, \quad j = 1, 2, \dots$$

(see, Lee (1990), p.160-161). Thus by (5.1), (5.2) and (5.3), we have

$$nB_n \xrightarrow{\mathcal{D}} \sum_{j=1}^{\infty} \frac{1}{j^2\pi^2}(Z_j^2 - 1) + \frac{1}{3} = \frac{1}{\pi^2}\sum_{j=1}^{\infty}\frac{1}{j^2}Z_j^2 + \frac{1}{6},$$

$$nV_n \xrightarrow{\mathcal{D}} \frac{1}{\pi^2}\sum_{j=1}^{\infty}\frac{1}{j^2}Z_j^2.$$

The asymptotic distribution of nV_n is well-known (see, for example, Pettitt (1981)).

In the following Examples 5 and 6, we consider the kernel of degree k given by $g(x_1, ..., x_k) = x_1 \cdots x_k$ and assume that $E(X) = 0$, $E(X^2) = 1$. Then $\theta(F) = 0$, $\psi_j(x_1, ..., x_j) = 0$ $(j = 1, ..., k-1)$, and $\psi_k(x_1, ..., x_k) = x_1 \cdots x_k$. Thus this kernel g is degenerate of order $d - 1 = k - 1$.

Example 5 (LB-statistics): We shall derive the asymptotic distribution of LB-statistics corresponding to the kernel $g(x_1, ..., x_k) = x_1 \cdots x_k$. Because of $d = k$, from (2.1) we have

$$\varphi_{d,d-2j}(x_1, ..., x_{d-2j}) = E[\psi_d(x_1, ..., x_{d-2j}, X_{d-2j+1}, X_{d-2j+1}, ..., X_{d-j}, X_{d-j})]$$

$$= x_1 \cdots x_{d-2j}.$$

Since $J_r(f) \stackrel{d}{=} CH_r(Z)$ for $f(x_1, ..., x_r) = Cx_1 \cdots x_r$ (see, for example, Koroljuk and Borovskich (1994), p.66), we have

$$J_{d-2j}(\xi_{d,d-2j}) = J_{d-2j}(\varphi_{d,d-2j}) \stackrel{d}{=} H_{d-2j}(Z)$$

where Z is the standard normal random variable.

Thus if $d = 2l + 1$ (l is a positive integer), by Proposition 2.4 we have

$$n^{d/2}B_n \xrightarrow{\mathcal{D}} \sum_{j=0}^{l} \frac{d!}{(d-2j)!j!}H_{d-2j}(Z).$$

For example, in case of $d = 3$,

$$n^{3/2}B_n \xrightarrow{\mathcal{D}} H_3(Z) + 6H_1(Z) = Z^3 - 3Z + 6Z = Z^3 + 3Z,$$

which is also easily obtained by direct computation because of $n^{3/2}B_n = [(n-1)(n-2)/(n+1)(n+2)]n^{3/2}U_n + [6n^2/(n+1)(n+2)](\sum_{i=1}^n X_i^2/n)(\sum_{i=1}^n X_i/\sqrt{n})$.

If $d = 2l$ (l is a positive number), then by Proposition 2.5, $k = d = 2l$, and $E\psi_d(X_1^2, ..., X_l^2) = (EX^2)^l = 1$ we have

$$n^{d/2}B_n \xrightarrow{\mathcal{D}} \sum_{j=0}^{l-1} \frac{(2l)!}{(2l-2j)!j!}H_{2l-2j}(Z) + \frac{k!}{l!}.$$

For example, in case of $k=2$,

$$nB_n \xrightarrow{\mathcal{D}} H_2(Z) + 2 = Z^2 + 1,$$

which is also easily obtained by direct computation because of $nB_n = [n/(n+1)][(\sum_{i=1}^n X_i/\sqrt{n})^2 + \sum_{i=1}^n X_i^2/n]$.

Example 6 (V-statistics): We shall derive the asymptotic distribution of V-statistics corresponding to the kernel $g(x_1, ..., x_k) = x_1 \cdots x_k$. As stated in Example 5, we have $J_{d-2j}(\xi_{d,d-2j}) \stackrel{d}{=} H_{d-2j}(Z)$. Thus if $d = 2l+1$ (l is a positive integer), by Proposition 3.4 we have

$$n^{d/2} V_n \xrightarrow{\mathcal{D}} \sum_{j=0}^{l} \frac{d!}{j!(d-2j)!2^j} H_{d-2j}(Z).$$

If $d = 2l$ (l is a positive number), then as stated in Example 1 we have $E\psi_d(X_1^2, ..., X_l^2) = (EX^2)^l = 1$. Thus by Proposition 3.4 and $H_0(z) = 1$ we have

$$n^{d/2} V_n \xrightarrow{\mathcal{D}} \sum_{j=0}^{l} \frac{d!}{j!(d-2j)!2^j} H_{d-2j}(Z).$$

Therefore using the relations of Hermite polynomials such that $x^{2l+1} = (2l+1)! \sum_{j=0}^{l} H_{2l+1-2j}(x)/[j!(2l+1-2j)!2^j]$ and $x^{2l} = (2l)! \sum_{j=0}^{l} H_{2l-2j}(x)/[j!(2l-2j)!2^j]$, we have

$$n^{d/2} V_n \xrightarrow{\mathcal{D}} Z^d.$$

This convergence is also easily shown by noting $n^{d/2} V_n = (\sum_{i=1}^n X_i/\sqrt{n})^d$. Hence this example is the one to show the validity of Propositions 3.4 and 3.5.

6. Appendix

Proof of Lemma 2.1: In the description of the kernel $g_{(j)}$ given by (1.5), we consider the positive integers $r_1, ..., r_j$ satisfying $r_1 + \cdots + r_j = k$. Since $(r_1 - 1) + \cdots + (r_j - 1) = k - j$ and $j = (k-j) + (2j-k)$, the number of $r_1, ..., r_j$ equal to one is at least $2j - k$. By the assumption, $2j - k \geq k - d + 1$. So typically we consider the case of $r_{j-k+d} = r_{j-k+d+1} = \cdots = r_j = 1$. Then we have

$$Eg(X_1^{r_1}, ..., X_j^{r_j}) = EE[g(X_1^{r_1}, ..., X_{j-k+d-1}^{r_{j-k+d-1}}, X_{j-k+d}, ..., X_j)$$
$$\mid X_1, ..., X_{j-k+d-1}] = E\psi_{d-1}(X_1^{r_1}, ..., X_{j-k+d-1}^{r_{j-k+d-1}})$$

which equals θ, because of $\psi_{d-1} = \theta$ w.p.1. Since the number of positive integers $r_1, ..., r_j$ satisfying $r_1 + \cdots + r_j = k$ is $\binom{k-1}{j-1}$, we have $E[U_n^{(j)}] = \theta$ for $j \geq k - (d-1)/2$. □

Proof of Lemma 2.2: For $c = 1, 2, ..., d-2j$, we consider

$$\psi_{(k-j),c}(x_1, ..., x_c) = \binom{k-1}{k-j-1}^{-1} \sum\nolimits^{+}_{r_1+\cdots+r_{k-j}=k} E[g(x_1^{r_1}, ..., x_c^{r_c}, X_{c+1}^{r_c+1}, ..., X_{k-j}^{r_{k-j}})].$$

Since $k-(r_{c+1}+\cdots+r_{k-j})=r_1+\cdots+r_c \geq c$, we have $(r_{c+1}-1)+\cdots+(r_{k-j}-1)= 0,1,...,j$ where the number of terms of the left-hand side is $k-j-c=j+(k-2j-c)$. Thus the number of $r_{c+1},...,r_{k-j}$ equal to one is at least $k-2j-c$.

At first we consider the case of $c=1,2,...,d-2j-1$. Typically we suppose $r_{j+c+1}=r_{j+c+2}=\cdots=r_{k-j}=1$, since the number of $r_{c+1},...,r_{k-j}$ equal to one is at least $k-2j-c$ as stated above. Then

$$Eg(x_1^{r_1},...,x_c^{r_c},X_{c+1}^{r_{c+1}},...,X_{c+j}^{r_{c+j}},X_{j+c+1},...,X_{k-j})=E\psi_{2j+c}(x_1^{r_1},...,x_c^{r_c},X_{c+1}^{r_{c+1}},...,X_{c+j}^{r_{c+j}}),$$

which is equal to θ because of $\psi_{2j+c}=\theta$ w.p.1 for $2j+c\leq d-1$. This is valid also for another $r_c,...,r_{k-j}$ which include at least the $k-2j-c$ components equal to 1. Thus we have $\psi_{(k-j),c}(x_1,...,x_c)=\theta$ w.p.1 for $1\leq c\leq d-2j-1$.

Now we consider the case of $c=d-2j$. Then the number of $r_{c+1},...,r_{k-j}$ equal to 1 is at least $k-2j-c=k-d$. If the number of $r_{c+1},...,r_{k-j}$ equal to 1 is more than $k-d$, then because the sum of r's not equal to 1 is less than d we have

$$Eg(x_1^{r_1},...,x_c^{r_c},X_{c+1}^{r_{c+1}},...,X_{k-j}^{r_{k-j}})=\theta.$$

If the number of $r_{c+1},...,r_{k-j}$ equal to 1 is $k-d$ exactly, then we consider typically the case of $r_{d-j+1}=r_{d-j+2}=\cdots=r_{k-j}=1$. Then $r_1+\cdots+r_c+r_{c+1}+\cdots+r_{d-j}=d$ because of $r_1+\cdots+r_{k-j}=k$, and $r_{c+1},...,r_{d-j}\geq 2$, where $c=d-2j$. Hence we have $r_1+\cdots+r_c\leq d-2j=c$ and $r_1=r_2=\cdots=r_c=1$. Therefore $r_{c+1}+\cdots+r_{d-j}=2j$ and we get $r_{c+1}=r_{c+2}=\cdots=r_{d-j}=2$. Thus as a typical term we have

$$Eg(x_1,...,x_c,X_{c+1}^2,...,X_{d-j}^2,X_{d-j+1},...,X_{k-j})=E\psi_d(x_1,...,x_c,X_{c+1}^2,...,X_{d-j}^2),$$

which is equal to $E\psi_d(x_1,...,x_c,X_{c+1},X_{c+1},...,X_{d-j},X_{d-j})$ by our notation. The number of these terms is $\binom{k-d+j}{j}=\binom{k-j-c}{d-j-c}$, which is the number of ways choosing $d-j-c$ places where the terms $X^2=(X,X)$ (with our notation) appear among $k-j-c$ places. All the other terms are equal to θ and its number is $\binom{k-1}{k-j-1}-\binom{k-d+j}{j}$. Hence we have (2.1) and the order of degeneracy of $U_n^{(k-j)}$ is at least $d-2j-1$ for $1\leq j\leq (d-1)/2$. □

Proof of Proposition 2.3: Since $E\psi_{(j),c}^2(X_1,...,X_c)\leq E[g_{(j)}^2(X_1,...,X_j)]$ and by Minkowski's inequality

$$\{E[g_{(j)}^2(X_1,...,X_j)]\}^{1/2}\leq \binom{k-1}{j-1}^{-1}\sum\nolimits^{+}_{r_1+\cdots+r_j=k}\{E[g^2(X_1^{r_1},...,X_j^{r_j})]\}^{1/2},$$

under the assumption we have $E\psi_{(j),c}^2(X_1,...,X_c)<\infty$ for $c=1,...,j$ and $j=1,...,k$. These assure the convergence of $U_n^{(j)}$ in distribution as stated in connection with (1.8). Since $\sum_{j=1}^k\binom{k-1}{j-1}\binom{n}{j}=\binom{n+k-1}{k}$, from (1.4) we have

$$n^{d/2}(B_n-\theta)=\sum_{j=0}^{l}\binom{k-1}{k-j-1}\frac{\binom{n}{k-j}}{\binom{n+k-1}{k}}n^{d/2}[U_n^{(k-j)}-\theta]+T_{1n}, \tag{6.1}$$

where

$$T_{1n} = \sum_{j=1}^{k-l-1} \binom{k-1}{j-1} \frac{\binom{n}{j}}{\binom{n+k-1}{k}} n^{d/2}[U_n^{(j)} - \theta].$$

For $1 \leq j \leq k-l-1$, we have $\binom{n}{j}n^{d/2}/\binom{n+k-1}{k} = O(n^{-1/2})$. Since $U_n^{(j)}$ converges to $EU_n^{(j)}$ as $n \to \infty$ w.p.1, T_{1n} converges to zero as $n \to \infty$ w.p.1. For $0 \leq j \leq l$, the order of degeneracy of $U_n^{(k-j)}$ is at least $d-2j-1$ by Lemma 2.2 and $E[U_n^{(k-j)}] = \theta$ by Lemma 2.1. Hence if the order of degeneracy of $U_n^{(k-j)}$ is $d-2j-1$, then by (1.8),

$$\frac{\binom{n}{k-j}}{\binom{n+k-1}{k}} n^{d/2}[U_n^{(k-j)} - \theta] = \frac{\binom{n}{k-j}n^j}{\binom{n+k-1}{k}} n^{(d-2j)/2}[U_n^{(k-j)} - \theta]$$

converges to

$$\frac{k!}{(k-j)!}\binom{k-j}{d-2j}\binom{k-1}{j}^{-1}\binom{k-d+j}{j} J_{d-2j}(\xi_{d,d-2j}) \tag{6.2}$$

in distribution as $n \to \infty$. Where we used the relation for the kernel (2.1) related with $U_n^{(k-j)}$;

$$\psi_{(k-j),d-2j}(x_1, ..., x_{d-2j}) - \theta = \binom{k-1}{j}^{-1}\binom{k-d+j}{j}\xi_{d,d-2j}(x_1, ..., x_{d-2j}).$$

If the order of degeneracy of $U_n^{(k-j)}$ is larger than $d-2j-1$ and is equal to $d-2j+i$ for some $i \geq 0$, then $n^{(d-2j+i+1)/2}[U_n^{(k-j)} - \theta]$ converges in distribution. Then, $\binom{n}{k-j}n^j n^{(d-2j)/2}[U_n^{(k-j)} - \theta]/\binom{n+k-1}{k}$ converges to zero in probability. In this case, $\varphi_{d,d-2j}$ is equal to θ by (2.3) since $\varphi_{d,d-2j}$ is constant. Therefore $\xi_{d,d-2j} = 0$ and we have $J_{d-2j}(\xi_{d,d-2j}) = 0$. Thus the corresponding term does not appear in the asymptotic distribution. Applying these convergence to the right-hand side of (6.1), we get (2.4). □

Proof of Lemma 2.4: The kernel of $U_n^{(k-l)}$ is given by (1.5) replacing j by $k-l$. So we consider $r_1, ..., r_{k-l}$ satisfying $r_1 + \cdots + r_{k-l} = k$, where $d = 2l$. Since $(r_1 - 1) + \cdots + (r_{k-l} - 1) = l$, noting $k - l = l + (k-d)$ we know that the number of r's, equal to 1 among $r_1, ..., r_{k-l}$, is at least $k-d$.

At first let us assume that the number of r's equal to 1 is exactly $k-d$. For example, if we consider the case that $r_{l+1} = \cdots = r_{k-l} = 1$, then $(r_1 - 1) + \cdots + (r_l - 1) = l$ with $r_1 - 1, ..., r_l - 1 \geq 1$ and hence $r_1 = \cdots = r_l = 2$. The number of ways, such that the number of r's equal to 1 among $r_1, ..., r_{k-l}$ is exactly $k-d$, is $\binom{k-l}{k-d}$. The corresponding $Eg(X_1^{r_1}, ..., X_{k-l}^{r_{k-l}})$ is equal to $Eg(X_1, ..., X_{k-d}, X_{k-d+1}^2, ..., X_{k-d+l}^2) = E\psi_d(X_1^2, ..., X_l^2)$.

Now let us assume that the number of r's among $r_1, ..., r_{k-l}$ equal to 1 is more than or equal to $k-d+1$. For example, if we assume that $r_{d-l} = \cdots = r_{k-l} = 1$, then $r_1 + \cdots + r_{d-l-1} = d-1$ and $Eg(X_1^{r_1}, ..., X_{k-l}^{r_{k-l}}) = Eg(X_1^{r_1}, ..., X_{d-l-1}^{r_{d-l-1}}, X_{d-l}, ..., X_{k-l}) = E\psi_{d-1}(X_1^{r_1}, ..., X_{d-l-1}^{r_{d-l-1}}) = \theta$. This is still valid for the another $r_1, ..., r_{k-l}$ which include at least the $k-d+1$ components equal to 1.

Thus by (1.5) we have

$$EU_n^{(k-l)} = \binom{k-1}{l}^{-1}\{\binom{k-l}{k-d}E\psi_d(X_1^2,...,X_l^2) + [\binom{k-1}{l} - \binom{k-l}{k-d}]\theta\}.$$

By our notation, $\psi_d(X_1^2,...,X_l^2)$ means $\psi_d(X_1,X_1,...,X_l,X_l)$. Hence we get the desired relation. □

Proof of Proposition 2.5: By the same reason stated in the first paragraph of the proof of Proposition 2.3, under the assumption we have $E\psi^2_{(j),c}(X_1,...,X_c) < \infty$ for $c = 1,...,j$ and $j = 1,...,k$, and the convergence of $U_n^{(j)}$ are assured. From (1.4) we have

$$n^{d/2}(B_n - \theta) = \sum_{j=0}^{l-1}\binom{k-1}{k-j-1}\frac{\binom{n}{k-j}}{\binom{n+k-1}{k}}n^{d/2}[U_n^{(k-j)} - \theta] \tag{6.3}$$

$$+\binom{k-1}{k-l-1}\frac{\binom{n}{k-l}}{\binom{n+k-1}{k}}n^{d/2}[U_n^{(k-l)} - \theta] + T_{2n},$$

where

$$T_{2n} = \sum_{j=1}^{k-l-1}\binom{k-1}{j-1}\frac{\binom{n}{j}}{\binom{n+k-1}{k}}n^{d/2}[U_n^{(j)} - \theta].$$

For $1 \le j \le k-l-1$, we have $\binom{n}{j}n^{d/2}/\binom{n+k-1}{k} = O(n^{-1})$ and T_{2n} converges to zero as $n \to \infty$ w.p.1.

$\binom{n}{k-l}n^{d/2}/\binom{n+k-1}{k}$ converges to $k!/(k-l)!$ and $U_n^{(k-l)}$ to $EU_n^{(k-l)}$ w.p.1 as $n \to \infty$. Hence $\binom{k-1}{l}\binom{n}{k-l}n^{d/2}[U_n^{(k-l)} - \theta]/\binom{n+k-1}{k}$ converges w.p.1. as $n \to \infty$ to $\binom{k-1}{l}k!$ $[EU_n^{(k-l)} - \theta]/(k-l)!$, which equals $\{k!/[l!(k-d)!]\}\{E\psi_d(X_1^2,...X_l^2) - \theta\}$ by Lemma 2.5. By our notation, $\psi_d(X_1^2,...,X_l^2)$ means $\psi_d(X_1,X_1,...,X_l,X_l)$.

For $0 \le j \le l-1$, the order of degeneracy of $U_n^{(k-j)}$ is at least $d-2j-1$ and therefore by (1.8), $\binom{n}{k-j}n^{d/2}[U_n^{(k-j)} - \theta]/\binom{n+k-1}{k}$ converges to $(k!/(k-j)!)\binom{k-j}{d-2j}\binom{k-1}{j}^{-1}\binom{k-d+j}{j}J_{d-2j}$ $(\xi_{d,d-2j})$ in distribution as $n \to \infty$. Here we used the relation immediately after (6.2).

If the order of degeneracy of $U_n^{(k-j)}$ is larger than $d-2j-1$, then the corresponding term does not appear in the asymptotic distribution as stated at the last of the proof of Proposition 2.3. Applying these convergence to the right-hand side of (6.3), we get (2.5). □

Acknowledgement

The authors wish to express their thanks to the referee for his kind comments.

References

Borovskikh, Yu. V. (1985). Estimates of characteristic functions of some random variables with applications to ω^2-statistics. II, *Theory of Probability and its Applications,* **30**, 117–127.

Borovskikh, Yu.V. (1996). *U-statistics in Banach spaces.* VSP, Utrecht.

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

Kotani, S. (1997). *Measure and Probability 2* (In Japanese), Iwanami, Tokyo.

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

Nomachi, T. and Yamato, H. (2001). Asymptotic comparisons of U-statistics, V-statistics and Limits of Bayes estimates by deficiencies, *Journal of Japan Statistical Society,* **31**, No.1, 85–98.

Pettitt, A.N. (1981). Cramér-von Mises Statistic, in: Kotz, S. and Johnson, N.L., ed., *Encyclopedia of Statistical Science,* **2**, John Wiley and Sons, New York, 220–221.

Rubin, H. and Vitale, R.A. (1980). Asymptotic distribution of symmetric statistics, *The Annals of Statistics,* **8**, 165–170.

Yamato, H. (1977). Relations between limiting Bayes estimates and U-statistics for estimable parameters, *Journal of Japan Statistical Society,* **7**, 57–66.

Received February 1, 2001
Revised July 1, 2001
Re–Revised August 8, 2001