

BAYESIAN BOOTSTRAPS FOR U-PROCESSES, HYPOTHESIS TESTS AND CONVERGENCE OF DIRICHLET U-PROCESSES

Dixin Zhang

Guangdong Commercial College

Abstract: Suppose \mathcal{F} is a class of bounded or unbounded functions. We construct a Bayesian bootstrapped U-process over \mathcal{F} and study the limiting behavior of the process. We obtain conditional central limit theorems for Bayesian bootstrapped U-processes and Dirichlet U-processes over \mathcal{F} , and discuss Bayesian bootstrap approximations for U-processes. Some problems concerning hypothesis testing in high-dimensional spaces are solved by combining the results in this paper with the Projection Pursuit method.

Key words and phrases: Bayesian bootstrap approximation, conditional central limit theorem, U-process.

1. Introduction

In recent years, theories of U-statistics and U-processes have generated an extensive literature. Much of this literature is surveyed by Serfling (1980) and Lee (1990). U-processes are useful for solving complex statistical problems. Examples are density estimation, non-parameter regression tests and goodness-of-fit tests. Recent studies by Nolan and Pollard (1987, 1988), Acrones and Giné (1992, 1993, 1995), Acrones and Yu (1994) on U-processes have led to further developments. In particular, Nolan and Pollard (1988) provide a central limit theorem for U-processes of order $m = 2$; Acrones and Giné (1993) extend the results of Nolan and Pollard (1987, 1988) to U-Processes of arbitrary degrees. However, it is difficult to apply these results in practice because the distributions and limiting distributions of U-processes depend on the unknown underlying distributions. To solve this problem, some authors study the bootstrap for U-statistics and U-processes: Acrones and Giné (1992) study the bootstrap of U and V statistics; Acrones and Giné (1994) present an Efron-type of bootstrap for U-processes, and prove an (almost sure) consistency result; Huskova and Janssen (1993a, b) prove the consistency of the Bayesian bootstrap for U-statistics. The main objective of this paper is to develop approximations for the distributions and limiting distributions of U-processes using the Bayesian bootstrap method due

to Rubin (1981) and Lo (1987). This method has been validated with first- and second-order limit theorems for a variety of statistics. See, for example, Lo (1987, 1993), Wang (1989), Mason and Newton (1992), Huskova and Janssen (1993a, b), Praestgaard and Wellner (1993), Aerts and Janssen (1995), Shao and Tu (1995), Barbe and Bertail (1995), James (1997) and Zhang(1997).

In Bayesian analysis, Dirichlet processes are used in a wide range of situations due to the computational convenience that results from conjugate prior properties of the process. Ferguson (1973) constructs a prior distribution, using Dirichlet processes, to study Bayesian nonparametric problems. Lo (1987) combines Dirichlet processes with the Bayesian bootstrap, and gives approximations for the empirical processes indexed by a class of indicator functions in R^1 . In subsequent sections, we demonstrate that Lo's (1987) results hold also for U-processes indexed by a class of unbounded functions in R^d , $d \geq 1$. Furthermore, it is shown that a framework useful for testing multivariate hypotheses can be constructed by combining the Bayesian bootstrap and Projection Pursuit (PP) methods.

The analysis conducted here relies on some important results of Pollard (1984) for empirical processes, and the approach we adopt is parallel to Pollard (1984, Chs. II and VII). There are, however, significant differences between the two approaches. Empirical processes are constructed from sums of independent random variables, whereas our Dirichlet U-processes cannot be written in such form. In the paper, we construct a Bayesian bootstrapped U-process. We show that Bayesian bootstrapped U-processes and Dirichlet U-processes conditionally converge in distribution to a common P-bridge process almost surely, i.e., for the same index set, Bayesian bootstrapped U-processes, Dirichlet U-processes and U-processes have the same limiting distribution almost surely. This implies that Bayesian bootstrapped U-processes and Dirichlet U-processes can be used to simulate the distributions and limiting distributions of U-processes.

The rest of this paper is organized as follows. Section 2 introduces Dirichlet U-Processes and Bayesian bootstrapped U-processes, and gives the main results for Bayesian bootstrap approximations of U-Processes. In Section 3, we discuss problems of hypothesis testing in high dimensional spaces by combining the results in this paper with Projection Pursuit (PP) methods. Section 4 contains some useful lemmas and proofs of the main results.

2. Bayesian Bootstrap Approximation of U-processes and Dirichlet U-processes

Assume that $\{X_i\}$ are i.i.d. R^d -valued random vectors, $d \geq 1$, with a common unknown probability measure P_0 . Let $\{Z_i\}$ be independent standard exponential random variables with mean 1. Furthermore, suppose $\{X_i\}$ and $\{Z_i\}$

are mutually independent random sequences on a probability space $(\Omega, \mathcal{T}, \mathbf{P})$. Let $f(x_1, \dots, x_m)$ be a function from R^{md} to $R, x_i \in R^d, 1 \leq i \leq m$. Let $1 \leq i \leq m \leq n$ and

$$P = \prod_{k=1}^m P_0, Pf = \int f(X_1, \dots, X_m) dP, V_i = Z_i / \sum_{j=1}^n Z_j. \tag{2.1}$$

Obviously, (V_1, \dots, V_{n-1}) follows the Dirichlet distribution $D(1, \dots, 1)$. Suppose \mathcal{F} is a class of symmetric functions from R^{md} to R . The U-statistic based on the sample X_1, \dots, X_n and a symmetric kernel f is defined as

$$U_n f = (n - m)! / n! \sum_{(i_1, \dots, i_m) \in I_m^n} f(X_{i_1}, \dots, X_{i_m}), \tag{2.2}$$

where $I_m^n = \{(i_1, \dots, i_m) : 1 \leq i_k \leq n, i_l \neq i_k \text{ if } 1 \leq l \neq k, \leq m\}$. See Serfling (1980) or Lee (1990) for details. The process $\{\sqrt{n}(U_n f - Pf) : f \in \mathcal{F}\}$ is called a U-process over \mathcal{F} (cf. Nolan and Pollard (1987)). Let P_n be the empirical probability measure with respect to the sample X_1, \dots, X_n ,

$$U_{n,z} f = (n - m)! / n! \sum_{I_m^n} \left(\prod_{k=1}^m Z_{i_k} \right) f(X_{i_1}, \dots, X_{i_m}), \tag{2.3}$$

$$D_{n,z} f = \sqrt{n}(U_{n,z} f - U_n f).$$

Definition 2.1. $\{D_{n,z}(f - P_n f) : f \in \mathcal{F}\}$ is called a Bayesian bootstrapped U-process over \mathcal{F} .

Define the Dirichlet U-statistic as

$$U_{n,v} f = \sum_{(i_1, \dots, i_m) \in I_m^n} \left(\prod_{k=1}^m V_{i_k} \right) f(X_{i_1}, \dots, X_{i_m}), \tag{2.4}$$

and call $\{D_{n,v} f : f \in \mathcal{F}\}$ a Dirichlet U-process over \mathcal{F} , where $D_{n,v} = \sqrt{n}(U_{n,v} - U_n)$.

Condition 1. \mathcal{F} is permissible (see Pollard (1984, p.196)), and there exists a positive function F with $PF^2 < \infty$, such that $|f| \leq F$ for any $f \in \mathcal{F}$.

Condition 2. The class of graphs of functions in \mathcal{F} has polynomial discrimination (see Pollard (1984, pp.17-27)).

Let " \xrightarrow{d} " denote convergence in distribution and $V_n f = \sqrt{n}(U_n f - Pf), f \in \mathcal{F}$. Acrones and Giné (1993) show that the result of Nolan and Pollard (1988) still holds for $m \geq 3$. Especially, Theorem 4.9 (Acrones and Giné (1993)) implies that if \mathcal{F} satisfies Conditions 1 and 2, then

(1) $V_n \xrightarrow{d} B_P$, where B_P is a P-bridge process over \mathcal{F} defined in Theorem 2.1 below (cf. Theorem 3.1 of Arcones and Yu (1994)).

Making use of (1) and Pollard (1984, Theorem 12, p.70), we have

$$(2) \sup_{\mathcal{F}} \sqrt{n}|U_n f - P f| \xrightarrow{d} \sup_{f \in \mathcal{F}} |B_P f|.$$

Note that the probability measure P is unknown, so both sides of (1) and (2) are actually unknown. These results are hard to use in practice. In Theorem 2.1, we prove that conditional distributions of Bayesian bootstrapped U-processes over \mathcal{F} converge almost surely to that of a P-bridge process B_P over \mathcal{F} . Corollary 1 below shows that Dirichlet U-processes over \mathcal{F} also conditionally converge in distribution to a P-bridge process B_P over \mathcal{F} almost surely. We can therefore use Bayesian bootstrapped U-processes or Dirichlet U-processes to simulate the distributions and limiting distributions of U-processes.

Let Q be any probability measure, and let

$$\begin{aligned} \mathbf{X} &= (X_1, X_2, \dots), \quad f_{(1)}(x) = E(f(X_1, \dots, X_m) | X_1 = x), \\ Qf &= \int f dQ \quad \text{and} \quad d_p(f, g, Q) = (Q(|f - g|^p))^{1/p}, \quad p = 1, 2. \end{aligned} \tag{2.5}$$

Suppose \mathcal{X} is the class of all bounded, real functions on \mathcal{F} . Equip \mathcal{X} with the metric generated by the uniform norm $\|\cdot\|$, $\|x\| = \sup_{\mathcal{F}} |x(f)|$ for $x \in \mathcal{X}$. Let $C(\mathcal{F}, P)$ be the set of all functionals $x(\cdot)$ in \mathcal{X} that are uniformly continuous with respect to the seminorm $d_2(\cdot, \cdot, P)$ on \mathcal{F} . That is, to each $\epsilon > 0$ and each $x \in \mathcal{X}$ there is a $\delta > 0$ for which $|x(f) - x(g)| < \epsilon$ whenever $d_2(f, g, P) = (P(f - g)^2)^{1/2} < \delta$. Given a sample sequence $\mathbf{X} = (X_1, X_2, \dots)$, let “ $\xrightarrow{d^*}$ ” denote convergence in $\mathbf{P}(\cdot | \mathbf{X})$ -distribution.

Theorem 2.1. *Assume that the stochastic process $D_{n,z}$ is indexed by \mathcal{F} and that \mathcal{F} satisfies Conditions 1 and 2. For almost all samples \mathbf{X} , if \mathbf{X} is fixed, then*

- (i) $\{D_{n,z}(f - P_n f) : f \in \mathcal{F}\} \xrightarrow{d^*} \{B_P f : f \in \mathcal{F}\}$,
- (ii) $\sup_{f \in \mathcal{F}} |D_{n,z}(f - P_n f)| \xrightarrow{d^*} \sup_{f \in \mathcal{F}} |B_P f|$,

where B_P is a P-bridge process over \mathcal{F} . Here B_P is a tight Gaussian random element of \mathcal{X} with sample paths in $C(\mathcal{F}, P)$, zero mean and covariance $\text{Cov}(B_P f, B_P h) = m^2 \text{Cov}(f_{(1)}(X_1), h_{(1)}(X_1))$ for all $f, h \in \mathcal{F}$ (cf. Pollard (1984, pp. 149-157)).

Corollary 1. *Under the conditions of Theorem 2.1, for almost all samples \mathbf{X} , if \mathbf{X} is fixed, then*

- (i) $D_{n,v} \xrightarrow{d^*} B_P$ and (ii) $\sup_{f \in \mathcal{F}} |D_{n,v} f| \xrightarrow{d^*} \sup_{f \in \mathcal{F}} |B_P f|$,
- where B_P is as defined in Theorem 2.1.

Corollary 2. *If the conditions of Theorem 2.1 are satisfied, then*

- (i) $D_{n,v} \xrightarrow{d} B_P$ and (ii) $\sup_{f \in \mathcal{F}} |D_{n,v} f| \xrightarrow{d} \sup_{f \in \mathcal{F}} |B_P f|$,

where B_P is as defined in Theorem 2.1.

Remark 1. Since empirical processes are special cases of U-processes, Corollary 2 implies Zhang’s (1997) Theorem 3.3. Furthermore, we can obtain the Bayesian bootstrap approximation of empirical processes over \mathcal{F} from Theorem 2.1 or Corollary 1.

3. Hypothesis Testing in High-Dimensional Spaces

The objective of the PP technique is to project high dimensional data into low dimensional subspaces. Thus, the problem of “curse of dimensionality” (Huber (1985)) is avoided. In this section, we discuss hypothesis testing in R^d by combining Corollary 1 with the PP method. Obviously, the problem can be treated similarly by Theorem 2.1 and the PP method.

3.1 Test of multivariate distribution

Consider the testing problem

$$H_0 : F(x) = F_0(x) \quad vs \quad H_a : F(x) \neq F_0(x), \tag{3.1.1}$$

where $F_0(x)$ is a known distribution function, $x \in R^d$, $d \geq 1$. Assume that X_1, \dots, X_n are i.i.d. R^d -valued random vectors and $X_1 \sim P$. The indicator function of set A is I_A and $\| \cdot \|$ is the Euclidean norm. Let $a, x \in R^d$ with $\|a\| = 1$ and $t \in R$. Denote the distribution functions of X_1 and $a^T X_1$ as $F(x)$ and $F^a(t)$ respectively. Let

$$F_n(x) = 1/n \sum_{i=1}^n I_{(-\infty, x)}(X_i), \quad F_n^a(t) = 1/n \sum_{i=1}^n I_{(-\infty, t)}(a^T X_i),$$

$$F_{n,v}(x) = \sum_{i=1}^n V_i I_{(-\infty, x)}(X_i) \text{ and } F_{n,v}^a(t) = \sum_{i=1}^n V_i I_{(-\infty, t)}(a^T X_i).$$

We consider Bayesian bootstrap approximations of the Cramér-Von Mises statistic and PP Cramér-Von Mises statistic. The approximations enable us to conduct goodness-of-fit tests for unknown multivariate distribution functions. The Cramér-Von Mises statistic is

$$M_n = n \int_{R^d} (F_n(x) - F(x))^2 dF(x), \quad n \geq 1. \tag{3.1.2}$$

Let the PP Cramér-Von Mises statistics be defined as

$$\hat{M}_n = \sup_{\|a\|=1} n \int (F_n^a(y) - F^a(y))^2 dF^a(y), \quad n \geq 1, \tag{3.1.3}$$

and set

$$M_{n,v} = n \int_{\mathbb{R}^d} (F_n(x) - F_{n,v}(x))^2 dF_n(x), \quad (3.1.4)$$

$$\hat{M}_{n,v} = \sup_{\|a\|=1} n \int (F_n^a(y) - F_{n,v}^a(y))^2 dF_n^a(y), \quad n \geq 1. \quad (3.1.5)$$

We call $M_{n,v}$ and $\hat{M}_{n,v}$ a Bayesian bootstrapped Cramér-Von Mises statistic and a Bayesian bootstrapped PP Cramér-Von Mises statistic, respectively.

Proposition 1. M_n and $M_{n,v}$ have the same limiting distribution almost surely. Specifically,

- (i) $M_n \xrightarrow{d} M = \int_{\mathbb{R}^d} (B_p I_{(-\infty, x)})^2 dP$ and
- (ii) $M_{n,v} \xrightarrow{d^*} M = \int_{\mathbb{R}^d} (B_p I_{(-\infty, x)})^2 dP$ for almost all samples \mathbf{X} .

Obviously, $M_n = n \int_{\mathbb{R}^d} (F_n(x) - F_0(x))^2 dF_0(x)$ can be used as a test statistic for (3.1.1). If the null hypothesis holds, then (i) of Proposition 1 implies that $M_n \xrightarrow{d} M$. Given that P is unknown, the distribution functions of M and M_n cannot be calculated when $d > 2$. However, (ii) indicates that for almost all samples \mathbf{X} , if \mathbf{X} is given and n is sufficiently large, then the distribution of $M_{n,v}$ is approximately equal to that of M . Hence, we can use the distribution of $M_{n,v}$ to simulate the distribution of M_n and approximate the distribution of M . A procedure for simulation is as follows.

Step 1. Suppose n and m are sufficiently large and let X_1, \dots, X_n be an independent sample of size n . Generate m independent random replications of (V_1, \dots, V_n) : $(V_1^{(k)}, \dots, V_n^{(k)})$, $k = 1, \dots, m$. Calculate $M_{n,v}$ at $(V_1^{(k)}, \dots, V_n^{(k)})$ and denote the outcome by $M_{n,v}^{(k)}$, i.e.,

$$M_{n,v}^{(k)} = \sum_{j=1}^n \left(\sum_{i=1}^n (V_n^{(k)} - 1/n) I_{(-\infty, X_j)}(X_i) \right)^2, \quad 1 \leq k \leq m.$$

Step 2. Calculate the empirical distribution function with respect to $M_{n,v}^{(1)}, \dots, M_{n,v}^{(m)}$ and denote the outcome by $F_{m,n}$. Since $F_{m,n}$ converges to the distribution function of $M_{n,v}$ with $\mathbf{P}(\cdot | \mathbf{X})$ -probability one, we replace the distribution of M_n by $F_{m,n}$, and use it to approximate the distribution of M . For any given $\alpha \in (0, 1)$, a confidence region of level $1 - \alpha$ can be constructed by using $F_{m,n}$. Thus, a test for (3.1.1) can be conducted.

In order to avoid ‘‘curse of dimensionality’’ when testing (3.1.1), we employ the PP method. We have the following result.

Proposition 2. If H_0 in (3.1.1) holds, then

- (i) $\hat{M}_n \xrightarrow{d} \sup_{\|a\|=1} \int (B_P I_{(-\infty, y)}(a^\tau x))^2 dP$ and
- (ii) $\hat{M}_{n,v} \xrightarrow{d^*} \sup_{\|a\|=1} \int (B_P I_{(-\infty, y)}(a^\tau x))^2 dP$ for almost all \mathbf{X} .

3.2. Testing multivariate location

Suppose X_1, \dots, X_n are i.i.d. random vectors with $\mathbf{P}(X_1 < x) = F(x - \theta)$, and $F(y) = F(-y)$ for any $y \in R^d$. Furthermore, suppose that $F^a(t)$ is continuous for any unit vector $a \in R^d$, where $F^a(t) = \mathbf{P}(a^\tau X_1 < t)$. We wish to test

$$H_0 : \theta = 0 \quad \text{vs} \quad H_a : \theta \neq 0. \tag{3.2.1}$$

Let $h(\cdot)$ be a Borel measurable function on R^k , $h^a(x_1, \dots, x_k) = h(a^\tau x_1, \dots, a^\tau x_k)$, $x_i \in R^d$, $1 \leq i \leq k$, and $\mathcal{F}_h = \{h^a(x_1, \dots, x_k) : a \in R^d, \|a\| = 1\}$. We call $\{U_n f : f \in \mathcal{F}_h\}$ and $\{D_{n,v} f : f \in \mathcal{F}_h\}$ a PP U-process and a PP Bayesian bootstrapped U-process, respectively, and term $h(\cdot)$ the kernel function. Wilcoxon tests are based on U-statistics, which can be applied to location testing problems. Here we consider the multivariate case and give Bayesian bootstrap approximations.

Choose $m = 2$ in Corollary 1. The kernel of the Wilcoxon statistic is $h(x_1, x_2) = I_{\{x_1+x_2>0\}}$. If H_0 holds, then for any $h^a \in \mathcal{F}_h$, $Ph^a = 1/2$ and $h_{(1)}^a(x) = E(h(a^\tau X_1, a^\tau X_2)) | X_1 = x) = F^a(a^\tau x)$. Our test statistics are $D_n = \sup_{\|a\|=1} \sqrt{n} |U_n h^a - 1/2| = \sup_{h^a \in \mathcal{F}_h} \sqrt{n} |U_n h^a - Ph^a|$, $n \geq 1$. Obviously, $\sup_{f \in \mathcal{F}_h} |D_{n,v} f| = \sup_{\|a\|=1} \sqrt{n} |U_{n,v} h^a - U_n h^a|$. If $\theta = 0$, combining Theorem 12 of Pollard (1984, p.70) with the result of Nolan and Pollard (1988), and Corollary 1, we have

$$D_n \xrightarrow{d} \sup_{\mathcal{F}_h} |B_P f|, \quad \text{and} \quad \sup_{\mathcal{F}_h} |D_{n,v} f| \xrightarrow{d^*} \sup_{\mathcal{F}_h} |B_P f| \quad \text{for almost all } \mathbf{X}, \tag{3.2.2}$$

where $\text{Cov}(B_P h^a, B_P h^b) = 4 \int F^a(a^\tau y) F^b(b^\tau y) dF(y) - 1$ and $E(B_P h^a) = 0$ for any $h^a, h^b \in \mathcal{F}_h$. Using (3.2.2), a test of (3.2.1) can be readily conducted.

Suppose $\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_n$ are mutually independent, where ξ_1, \dots, ξ_n are i.i.d. random vectors with distribution function $F(x)$ and η_1, \dots, η_n are i.i.d. random vectors with distribution function $F(x - \mu)$, $x, \mu \in R^d$. Furthermore, suppose $F^a(t) = \mathbf{P}(a^\tau \xi_1 < t)$, $t \in R$, is continuous for any unit vector a in R^d . Our goal is to test

$$H_0 : \mu = 0 \quad \text{vs} \quad H_a : \mu \neq 0. \tag{3.2.3}$$

Let $X_i = (\xi_i, \eta_i)$, $1 \leq i \leq n$, $x = (x_1, x_2)$ and $u = (u_1, u_2)$, where $x_1, x_2, u_1, u_2 \in R^d$. Let $h((s_1, s_2), (t_1, t_2)) = I_{\{s_1 < t_2\}} + I_{\{t_1 \leq s_2\}}$, $s_1, s_2, t_1, t_2 \in R$. For any $a \in R^d, \|a\| = 1$, let $h^a(x, u) = h((a^\tau x_1, a^\tau x_2), (a^\tau u_1, a^\tau u_2))$. Let $m = 2$ and $\mathcal{F} = \mathcal{F}_h = \{h^a(\cdot, \cdot) : a \in R^d, \|a\| = 1\}$ in Corollary 1. If H_0 holds, then $Ph^a = E h^a(X_1, X_2) = 1$ for any $h^a \in \mathcal{F}_h$. Thus, $D_n = \sup_{\|a\|=1} \sqrt{n} |U_n(h^a) - 1|$

can be used to test (3.2.3). Obviously, $h_{(1)}^a(x) = 1 - F^a(a^\tau x_1) + F^a(a^\tau x_2)$, $D_n \xrightarrow{d} \sup_{\mathcal{F}_h} |B_P h^a|$, and, for almost all \mathbf{X} ,

$$\sup_{\|a\|=1} |D_{n,v} h^a| \xrightarrow{d^*} \sup_{\mathcal{F}_h} |B_P h^a|, \tag{3.2.4}$$

where $\text{Cov}(B_P(h^a), B_P(h^b)) = 8(\int F^a(a^\tau y)F^b(b^\tau y) dF(y) - 1/4)$ and $E(B_P(h^a)) = 0$ for any $h^a, h^b \in \mathcal{F}_h$. Therefore, we can use $D_{n,v}$ to simulate the distribution of D_n and test (3.2.3) from (3.2.4).

3.3. Tests about multivariate dispersion

A. One-sample case. Assume that X_1, \dots, X_n are i.i.d. R^d -valued random vectors with $\text{Var}(X_1) = V$. Let $V_0 > 0$ be a positive definite $d \times d$ matrix. We wish to test

$$H_0 : V = V_0 \quad \text{vs} \quad H_a : V \neq V_0. \tag{3.3.1}$$

Since $V = V_0$ is equivalent to $a^\tau(V - V_0)a = 0$ for all unit vectors $a \in R^d$, we can use $D_n = \sup_{\|a\|=1} \sqrt{n}|a^\tau(\hat{V} - V_0)a|$ as a test statistic, where $\hat{V} = \sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})^\tau / (n-1)$ is the sample covariance matrix. Let $h(x_1, x_2) = (x_1 - x_2)^2 / 2$ and $\mu = EX_1$. If the null hypothesis holds and $E\|X_1\|^4 < +\infty$, then $D_n = \sup_{\mathcal{F}_h} |U_n f - P f|$ and $h_{(1)}^a(x) = E(2^{-1}(a^\tau X_1 - a^\tau X_2)^2 | X_1 = x) = (a^\tau(x - \mu))^2 / 2 + a^\tau V_0 a / 2$. Therefore,

$$D_n \xrightarrow{d} \sup_{\mathcal{F}_h} |B_P f|, \text{ and } \sup_{\mathcal{F}_h} |D_{n,v} f| \xrightarrow{d^*} \sup_{\mathcal{F}_h} |B_P f| \text{ for almost all } \mathbf{X}, \tag{3.3.2}$$

where $\text{Cov}(B_P h^a, B_P h^b) = E(a^\tau(X_1 - \mu)b^\tau(X_1 - \mu))^2 - a^\tau V_0 a b^\tau V_0 b$ and $E(B_P h^a) = 0$ for any $h^a, h^b \in \mathcal{F}_h$. It is straightforward to see that H_0 can be tested using (3.3.2).

B. Two-sample case. Lehmann (1951) proposed a statistic for a two-sample scale test as follows. Suppose $\{X_{1,1}, \dots, X_{1,n_1}\}$ and $\{X_{2,1}, \dots, X_{2,n_2}\}$ are independent sequences of one-dimensional random variables, where $X_{1,1}, \dots, X_{1,n_1}$ are i.i.d. with distribution function $F(x)$ and $X_{2,1}, \dots, X_{2,n_2}$ are i.i.d. with distribution function $F((x - \mu)/\sigma)$. The U -statistic with kernel $h(s_1, s_2, t_1, t_2) = I_{\{|s_1 - s_2| > |t_1 - t_2|\}}$ was used to test $\sigma = 1$ against $\sigma \neq 1$. Here we give an alternative test statistic for the multivariate case using Bayesian bootstrap and PP methods.

Suppose ξ_1, \dots, ξ_n are i.i.d. random vectors in R^d with density function $G(x)$ and η_1, \dots, η_n are i.i.d. random vectors in R^d with density function $G_1(x) = \det(V)G(V(x - \mu))$, where $x, \mu \in R^d$ and V is a $d \times d$ positive definite matrix. Furthermore, suppose ξ_1, \dots, ξ_n and η_1, \dots, η_n are mutually independent and

$F^a(t) = \mathbf{P}(a^\tau \xi_1 < t)$ is continuous in R for any unit vector $a \in R^d$. We wish to test

$$H_0 : V = I_{d \times d} \quad vs \quad H_a : V \neq I_{d \times d}, \tag{3.3.3}$$

where $I_{d \times d}$ is a $d \times d$ identity matrix. Let $X_i = (\xi_i, \eta_i)$, $1 \leq i \leq n$, $x = (x_1, x_2)$, $u = (u_1, u_2)$ and $h^a(x, u) = I_{\{|a^\tau x_1 - a^\tau u_1| > |a^\tau x_2 - a^\tau u_2|\}}$, where $x_1, x_2, u_1, u_2, a \in R^d$. Choose $m = 2$ and $\mathcal{F} = \mathcal{F}_h = \{h^a : a \in R^d, \|a\| = 1\}$ in Corollary 1. If H_0 holds, then $a^\tau(\xi_1 - \xi_2)$ and $a^\tau(\eta_1 - \eta_2)$ have the same distribution and $Ph^a = 1/2$ for any unit vector $a \in R^d$. We use $D_n = \sup_{\|a\|=1} \sqrt{n} |1/[n(n-1)] \sum_{1 \leq i \neq j \leq n} I_{\{|a^\tau \xi_i - a^\tau \xi_j| > |a^\tau \eta_i - a^\tau \eta_j|\}} - 1/2|$ as the test statistic for (3.3.3). It is easy to show that

$$\begin{aligned} D_n &= \sup_{\mathcal{F}_h} \sqrt{n} |U_n(h^a(x, y)) - Ph^a|, D_n \xrightarrow{d} \sup_{\mathcal{F}_h} |B_P h^a|, \\ &\sup_{\mathcal{F}_h} |D_{n,v} h^a| \xrightarrow{d^*} \sup_{\mathcal{F}_h} |B_P h^a| \quad \text{for almost all } \mathbf{X}, \end{aligned} \tag{3.3.4}$$

where $\text{Cov}(B_P h^a, B_P h^b) = 4(E(h_{(1)}^a(\xi_1, \eta_1) h_{(1)}^b(\xi_1, \eta_1)) - 1/4)$ and $E(B_P h^a) = 0$ for any $h^a, h^b \in \mathcal{F}_h$.

3.4. Test of independence

To test the independence of one-dimensional random variables X and Y , Kendall (1938) proposed a method based on the U -statistic K_n with kernel function

$$h((s_1, t_1), (s_2, t_2)) = I_{\{(s_2 - s_1)(t_2 - t_1) > 0\}} - I_{\{(s_2 - s_1)(t_2 - t_1) \leq 0\}}. \tag{3.4.1}$$

Its reject on region is of the form $\{\sqrt{n}|K_n| > \beta\}$, we consider a multivariate counterpart.

Suppose ξ and η are d_1 - and d_2 -dimensional random vectors respectively, $d = d_1 + d_2$. Furthermore, suppose X_1, \dots, X_n are independent observations of (ξ, η) , where $X_i = (\xi_i, \eta_i)$, $1 \leq i \leq n$. We are interested in testing

$$H_0 : \xi \text{ and } \eta \text{ are independent} \quad vs \quad H_a : H_0 \text{ is not true.} \tag{3.4.2}$$

Let $a = (a_1, a_2) \in R^d$ with $\|a\| = 1$ and $a_r \in R^{d_r}$, $r = 1, 2$. Let the distribution functions of ξ and η be F and G , respectively. Furthermore, suppose $F^{a_1}(t)$ and $G^{a_2}(t)$ are continuous for any unit vector $a = (a_1, a_2)$, where $F^{a_1}(t) = \mathbf{P}(a_1^\tau \xi < t)$ and $G^{a_2}(t) = \mathbf{P}(a_2^\tau \eta < t)$. Let $x = (x_1, x_2)$, $u = (u_1, u_2)$, $x_r, u_r \in R^{d_r}$, $r = 1, 2$, and $h^a(x, u) = h((a_1^\tau x_1, a_2^\tau x_2), (a_1^\tau u_1, a_2^\tau u_2))$, where $h((s_1, t_1), (s_2, t_2))$ is defined by (3.4.1). Using Corollary 1 and choosing $m = 2$ and $\mathcal{F} = \mathcal{F}_h = \{h^a : a \in$

$R^d, \|a\| = 1\}$, it is easy to see that $\sup_{\|a\|=1} \sqrt{n}|U_n h^a|$ can be used as a test statistic for (3.4.2). If the null hypothesis is true, then

$$Ph^a = 0, D_n = \sup_{\|a\|=1} \sqrt{n}|U_n h^a - Ph^a| = \sup_{h^a \in \mathcal{F}_h} \sqrt{n}|U_n h^a|,$$

$$h_{(1)}^a(x, y) = 1 - 2F^{a_1}(a_1^\tau x) - 2G^{a_2}(a_2^\tau y) + 4F^{a_1}(a_1^\tau x)G^{a_2}(a_2^\tau y)$$

for any $h^a \in \mathcal{F}_h$. We have that

$$D_n \xrightarrow{d} \sup_{\mathcal{F}_h} |B_P f|, \text{ and } \sup_{\mathcal{F}_h} |D_{n,v} f| \xrightarrow{d^*} \sup_{\mathcal{F}_h} |B_P f| \text{ for almost all } \mathbf{X}, \quad (3.4.3)$$

where $E(B_P f) = 0$ and $\text{Cov}(B_P h^a, B_P h^b) = 4[1 - 4 \int F^{a_1}(a_1^\tau x) F^{b_1}(b_1^\tau x) dF(x)][1 - 4 \int G^{a_2}(a_2^\tau y) G^{b_2}(b_2^\tau y) dG(y)]$ for any $f, h^a, h^b \in \mathcal{F}_h$. Test (3.4.2) using (3.4.3).

4. Proofs of Main Results

Given a pseudometric space (\mathcal{F}, d) , the covering number and the covering integral are defined, respectively, by

$$N(u, d, \mathcal{F}) = \min\{m : \text{there are } f_1, \dots, f_m \in \mathcal{F} \text{ such that}$$

$$\sup_{f \in \mathcal{F}} \min_{1 \leq j \leq m} d(f, f_j) \leq u\},$$

$$J(u, d, \mathcal{F}) = \int_0^u (2 \log(N(t, \mathcal{F}, d)^2/t))^{1/2} dt \text{ for any } u > 0.$$

See Pollard (1984, p.143). Let $d_{n,p}(f, g) = (U_n(|f - g|^p))^{1/p}, p = 1, 2, n \geq 1$. Define $N_p(u, U_n, \mathcal{F})$ as the random covering numbers of $(\mathcal{F}, d_{n,p})$. In order to prove Theorem 2.1, we require the following lemmas.

Lemma 1. *Let \mathcal{F} be a permissible class of symmetric functions on R^{md} with an envelope $F > 0$ and $PF < \infty$. If $\log N_1(u, U_n, \mathcal{F}) = o_P(n)$, then*

- (i) $\sup_{\mathcal{F}} |U_{n,z^2} f - Pf| \rightarrow 0$ a.s. and
- (ii) $\sup_{\mathcal{F}} |U_{n,z^2} f - Pf| \rightarrow 0$ with $\mathbf{P}(\cdot | \mathbf{X})$ -probability one for almost all samples \mathbf{X} , where

$$U_{n,z^2} f = (n - m)! / (n!(2^m - 1)) \sum_{(i_1, \dots, i_m) \in I_m^n} \left(\prod_{k=1}^m Z_{i_k} - 1 \right)^2 f(X_{i_1}, \dots, X_{i_m}).$$

Proof. Let $\pi(z_1, \dots, z_m) = (\prod_{i=1}^m z_i - 1)^2 / (2^m - 1)$ and $f_\pi((z_1, x_1), \dots, (z_m, x_m)) = \pi(z_1, \dots, z_m) f(x_1, \dots, x_m)$, where $f \in \mathcal{F}, z_i \in R, x_i \in R^d$ and $1 \leq i \leq m$. Obviously, $U_{n,z^2} f$ is the U-statistic based on i.i.d. pairs $(Z_i, X_i), 1 \leq i \leq n$, with kernel function f_π . For any $\epsilon \in (0, 1)$, there is $K > 0$ such that $E(\pi(Z_1, \dots, Z_m) I_{\{\pi > K\}}) < \epsilon / (2(PF + 1))$. Let $\mathcal{F}_K = \{f_\pi I_{\{\pi \leq K\}} : f \in \mathcal{F}\}$

and $F_{\pi,K} = \pi F I_{\{\pi > K\}}$. Since $U_{n,z^2} f = U_n f_\pi$ and $P f_\pi = P f$ for any $f \in \mathcal{F}$, we have $\sup_{\mathcal{F}} |U_{n,z^2} f - P f| \leq \sup_{\mathcal{F}_K} |U_n f - P f| + U_n F_{\pi,K} + P(F_\pi I_{\{\pi > K\}})$. Obviously, $\log N_1(u, U_n, \mathcal{F}_K) \leq \log N_1(u/K, U_n, \mathcal{F}) = o_P(n)$ for any $u > 0$. Theorem 3.1 of Acrones and Giné (1993) and the Strong Law of Large Numbers for U-statistics imply that $\sup_{\mathcal{F}_K} |U_n f - P f| \rightarrow 0$ a.s. and $U_n F_{\pi,K} \rightarrow P(\pi F I_{\{\pi > K\}}) \leq \epsilon/2$ a.s. This completes the proof of result (i). Statement (ii) follows from result (i) by Fubini's theorem.

Lemma 2. *Suppose \mathcal{F} satisfies Condition 1 in Section 2 and there are $A > 0$ and $W > 0$ such that, for any probability measure Q with $QF^2 < \infty$ and for any $u > 0$, $N(u, d_2(\cdot, \cdot, Q), \mathcal{F}) \leq Au^{-W}$. Then for any $\epsilon > 0$ and $\eta > 0$ there is a $\delta > 0$ such that, for almost all samples \mathbf{X} $\limsup_{n \rightarrow \infty} \mathbf{P}\{\sup_{(f,g) \in [\delta]} |D_{n,z}(f - g)| > \eta \mid \mathbf{X}\} \leq \epsilon$, where $[\delta] = \{(f, g) : f, g \in \mathcal{F}, d_2(f, g, P) = (P(f - g)^2)^{1/2} \leq \delta\}$ and $\delta > 0$.*

Proof. For any $\eta > 0$, let $\delta \in (0, \eta/(4m))$. By Theorem 3.1 of Acrones and Giné (1993), for almost all samples \mathbf{X} there is a positive integer $N_{\mathbf{X}}$ such that, for any $n \geq N_{\mathbf{X}}$ and any $(f, g) \in [\delta]$, $\mathbf{P}(|D_{n,z}(f - g)| > \eta/2 \mid \mathbf{X}) \leq 2(\eta/2)^{-2} m^2 \delta^2 \leq 1/2$. Choose a sequence of i.i.d. random variables $\{\sigma_i : 1 \leq i \leq n\}$, $\mathbf{P}(\sigma_i = +1) = \mathbf{P}(\sigma_i = -1) = 1/2$, independent of the sequences $\{X_i\}$ and $\{Z_i\}$. From Lemma 8 and the Second Symmetrization Method in Pollard (1984, p.14),

$$\mathbf{P}\left(\sup_{(f,g) \in [\delta]} |D_{n,z}(f - g)| > \eta \mid \mathbf{X}\right) \leq 4\mathbf{P}\left(\sup_{(f,g) \in [\delta]} |D_{n,z}^0(f - g)| > \eta/4 \mid \mathbf{X}\right), \tag{4.0}$$

where $D_{n,z}^0 f = \sqrt{n}(n - m)!/n! \sum_{I_m^n} (\prod_{k=1}^m \sigma_{i_k}) (\prod_{k=1}^m Z_{i_k} - 1) f(X_{i_1}, \dots, X_{i_m})$. Let $d_{n,z}(f, g) = (U_{n,z^2}(f - g)^2)^{1/2}$ and $[\delta]_n = \{(f, g) : f, g \in \mathcal{F}, d_{n,z}(f, g) \leq \delta\}$. We have

$$\begin{aligned} & \mathbf{P}\left(\sup_{(f,g) \in [\delta]} |D_{n,z}^0(f - g)| > \eta/4 \mid \mathbf{X}\right) \\ & \leq \mathbf{P}\left(\sup_{(f,g) \in [2\delta]_n} |D_{n,z}^0(f - g)| > \eta/4 \mid \mathbf{X}\right) + \mathbf{P}([\delta] - [2\delta]_n \mid \mathbf{X}). \end{aligned} \tag{4.1}$$

By Lemma 1, $\mathbf{P}([\delta] - [2\delta]_n \mid \mathbf{X}) \leq \mathbf{P}(\sup_{(f,g) \in [\delta]} |U_{n,z^2}(f - g)^2 - P(f - g)^2| > \delta^2/2 \mid \mathbf{X}) \rightarrow 0$ a.s. For given $X(n) = (X_1, \dots, X_n)$ and $Z(n) = (Z_1, \dots, Z_n)$, if $c > 0$ and $f, g \in \mathcal{F}$, Hoeffding's Inequality gives

$$\mathbf{P}(|D_{n,z}^0(f - g)| \geq cd_{n,z}(f, g) \mid X(n), Z(n)) \leq 2 \exp\{-c^2/2\}, \quad n \geq 2^m.$$

For any $\epsilon \in (0, 1)$, choose $\epsilon_1 \in (0, \epsilon/8)$. Lemma 1 implies there exist an integer $N > 0$ and a set $E \in \mathcal{T}$ not depending on \mathbf{X} with $\mathbf{P}(E) > 1 - \epsilon_1$, such that, for

any $n \geq N$

$$1/2 < \eta_n := (n - m)! / ((2^m - 1)n!) \sum_{I_m^n} \left(\prod_{k=1}^m Z_{i_k} - 1 \right)^2 < 2 \text{ and}$$

$$U_{n,z^2} F^2 \leq 2PF^2 \text{ on } E \text{ uniformly.} \tag{4.2}$$

Let U_n^* put mass $(\prod_{k=1}^m Z_{i_k} - 1)^2 / (\sum_{I_m^n} (\prod_{k=1}^m Z_{i_k} - 1)^2)$ on each element $(X_{i_1}, \dots, X_{i_m}), (i_1, \dots, i_m) \in I_m^n$. Obviously, $U_n^*(f) = U_{n,z^2}(f) / \eta_n$. Let $d_2(f, g, U_n^*) = (U_n^*(f - g)^2)^{1/2}$. If $n \geq N$, (4.2) implies that for any $f, g \in \mathcal{F}$, $d_{n,z}(f, g) \leq 2d_2(f, g, U_n^*)$ uniformly on E . Consequently, for any $n \geq N$ and any $u > 0$, $N(2u, d_{n,z}, F) \leq N(u, d_2(\cdot, \cdot, U_n^*), \mathcal{F}) \leq Au^{-W}$ on E uniformly. This implies that $J_2(\delta, d_{n,z}, \mathcal{F}) \rightarrow 0$ on E uniformly as $\delta \rightarrow 0$. Therefore, there exists a positive number $\delta_1 \leq \min\{\epsilon/18, \eta/(4m)\}$ such that, for any $t \in (0, \delta_1]$ $J_2(t, d_{n,z}, \mathcal{F}) \leq \eta/208$ on E uniformly. Choose $n \geq N$ and $\delta \in (0, \delta_1/2]$. Let $\mathbf{P}_E(\cdot) = \mathbf{P}(\cdot \cap E) / \mathbf{P}(E)$. Making use of Lemma 9 of Pollard (1984, Chapter VII) on the probability space $(E, E \cap \mathcal{T}, \mathbf{P}_E)$,

$$\mathbf{P}_E(\sup_{[2\delta]_n} |D_{n,z}^0(f - g)| > 26J_2(2\delta, d_{n,z}, \mathcal{F}) \mid X(n), Z(n)) \leq 4\delta$$

and then

$$\begin{aligned} & \mathbf{P}(\sup_{[2\delta]_n} |D_{n,z}^0(f - g)| > \eta/4 \mid \mathbf{X}) \\ & \leq \mathbf{P}_E(\sup_{[2\delta]_n} |D_{n,z}^0(f - g)| > \eta/4 \mid \mathbf{X}) + \mathbf{P}(E^c) \leq \epsilon/4. \end{aligned} \tag{4.3}$$

Combining (4.3) with (4.0) and (4.1) completes the proof of Lemma 2.

Lemma 3. *Suppose $Pf^2 < \infty$ for any $f \in \mathcal{F}$. Then for any integer $k > 0$ and $f_i \in \mathcal{F}, 1 \leq i \leq k$, $(D_{n,v}f_1, \dots, D_{n,v}f_k) \xrightarrow{d^*} N(0, \Sigma)$ for almost all samples \mathbf{X} , where $\Sigma = (v_{i,j})_{k \times k}$, $v_{i,j} = m^2(P(f_{i,(1)}f_{j,(1)}) - Pf_iPf_j)$.*

Proof. Let $T_n f = m\sqrt{n} \sum_{i=1}^n (V_i - 1/n) f_{(1)}(X_i)$. Obviously, for almost all \mathbf{X} ,

$$T_n f \xrightarrow{d^*} N(0, m^2 \text{Var}(f_{(1)}(X_1))), \tag{4.4}$$

see Lo (1987). It is enough for us to show that (a) $E((D_{n,v}f - T_n f)^2 \mid \mathbf{X}) \rightarrow 0$ for almost all samples \mathbf{X} . Note that $E(\prod_{h=1}^{m_1} V_{i_h}^{k_h}) = \Gamma(n) \prod_{h=1}^{m_1} \Gamma(k_h + 1) / \Gamma(n + \sum_{h=1}^{m_1} k_h)$, where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \exp\{-x\} dx, \alpha > 0, k_h > 0, 1 \leq h \leq m_1 \leq n$. It follows from the Strong Law of Large Numbers for U-statistics that, for almost all samples \mathbf{X} ,

$$\begin{aligned} & E[(D_{n,v}f)^2 \mid \mathbf{X}] \rightarrow m^2 \text{Var}(f_{(1)}(X_1)), \quad E[(T_n f)^2 \mid \mathbf{X}] \rightarrow m^2 \text{Var}(f_{(1)}(X_1)), \\ & E[(T_n f)(D_{n,v}f) \mid \mathbf{X}] \rightarrow m^2 \text{Var}(f_{(1)}(X_1)). \end{aligned} \tag{4.5}$$

We obtain (a) from (4.5). Since $D_{n,v}$ is a linear operator, the lemma follows from (a), (4.4) and properties of characteristic functions.

Proof of Theorem 2.1. Let $\mathcal{F}_0 = \{f - Pf : f, g \in \mathcal{F}\}$ and Q be any probability measure with $QF^2 < \infty$. Obviously, $d_2(f - Pf, g - Pg, Q) \leq d_2(f, g, Q) + d_2(f, g, P)$ for any $(f - Pf), (g - Pg) \in \mathcal{F}_0$. Conditions 1 and 2, Lemmas II.25 and 36 in Pollard (1984, pp.27-34) imply that there exist $A > 0$ and $W > 0$ such that, for all $u > 0$ and for all probability measures Q with $QF^2 < \infty$,

$$N(u, d_2(\cdot, \cdot, Q), \mathcal{F}_0) \leq N(u/2, d_2(\cdot, \cdot, Q), \mathcal{F})N(u/2, d_2(\cdot, \cdot, P), \mathcal{F}) \leq Au^{-W}. \tag{4.6}$$

This implies that Lemma 2 holds for the case of $\mathcal{F} = \mathcal{F}_0$. Obviously, $D_{n,z}(1) \rightarrow N(0, m^2)$. By Theorem II.24 in Pollard (1984, p.25), for almost all \mathbf{X} , $\sup_{\mathcal{F}} |D_{n,z}(f - Pf) - D_{n,z}(f - P_n f)| \leq \sup_{\mathcal{F}} |P_n f - Pf| |D_{n,z}(1)| \rightarrow 0$ in $\mathbf{P}(\cdot | \mathbf{X})$ -probability. Therefore, Lemma 2 also holds for the process $\{D_{n,z}(f - P_n f) : f \in \mathcal{F}\}$. Let $\bar{Z}_{n-k} = 1/(n-k) \sum_{i=1}^n Z_i$, $1 \leq k \leq m$, and $C_{n,m} = \prod_{k=0}^{m-1} \bar{Z}_{n-k}$. Since $D_{n,v} f = D_{n,v}(f - Pf) + D_{n,v}(Pf)$, $\sup_{\mathcal{F}} |D_{n,v} f - D_{n,z}(f - Pf)| \leq |D_{n,v}(1)|PF + |\sqrt{n}(C_{n,m} - 1)/C_{n,m}| \sup_{\mathcal{F}} |U_{n,z}(f - Pf)|$, where $U_{n,z} f$ is defined in (2.3). Since $E|D_{n,v}(1)| = \sqrt{n}E(1 - \sum_{I_m^n} \prod_{k=1}^m V_{i_k}) = O(n^{-1/2})$, $D_{n,v}(1)$ converges in probability to zero. By (4.6) and a similar proof to that of Lemma 1, for almost all samples \mathbf{X} , $\sup_{\mathcal{F}} |U_{n,z}(f - Pf)| \rightarrow 0$ with $\mathbf{P}(\cdot | \mathbf{X})$ -probability one. Note that

$$\sqrt{n}|C_{n,m} - 1| \leq \sum_{k=0}^{m-2} \left(|\sqrt{n}(\bar{Z}_{n-k} - 1)| \prod_{i=k+1}^{m-1} \bar{Z}_{n-k} \right) + |\sqrt{n}(\bar{Z}_{n-m+1} - 1)|.$$

By the Central Limit Theorem and Kolmogorov’s Strong Law of Large Numbers, $\sqrt{n}(\bar{Z}_{n-k} - 1)$ converges in distribution to $N(0, 1)$ and $\bar{Z}_{n-k} \rightarrow 1$ a.s. as $n \rightarrow \infty$ for any $0 \leq k \leq m - 1$. It follows from Slutsky’s Theorem that

$$\sup_{f \in \mathcal{F}} |D_{n,v} f - D_{n,z}(f - Pf)| \rightarrow 0 \text{ in } \mathbf{P}(\cdot | \mathbf{X})\text{-probability}. \tag{4.7}$$

It follows from (4.7) that, for almost all samples \mathbf{X} ,

$$\begin{aligned} \sup_{\mathcal{F}} |D_{n,v} f - D_{n,z}(f - P_n f)| &\leq \sup_{\mathcal{F}} |D_{n,v} f - D_{n,z}(f - Pf)| \\ &+ \sup_{\mathcal{F}} |P_n f - Pf| |D_{n,z}(1)| \rightarrow 0 \text{ in } \mathbf{P}(\cdot | \mathbf{X})\text{-probability}. \end{aligned} \tag{4.8}$$

The proof of (i) is almost the same as that of Theorem VII.21 of Pollard (1984). We therefore obtain result (i) from (4.8) and Lemmas 2 and 3. Then (ii) follows from (i) and Theorem 12 in Pollard (1984, p.70).

Proof of Corollary 1. Equation (4.8) and Lemma 2 imply that for almost all \mathbf{X} , $\{D_{n,v} f : f \in \mathcal{F}\}$ and $\{D_{n,z}(f - P_n f) : f \in \mathcal{F}\}$ have the same limit distribution. By Theorem 2.1, Corollary 1 holds.

Proof of Corollary 2. By the definition of weak convergence of sequences of random vectors and the Bounded Convergence Theorem, Theorem 2.1 implies that the finite-dimensional distributions of $\{D_{n,z}(f - P_n f) : f \in \mathcal{F}\}$ converge in $\mathbf{P}(\cdot)$ -distribution to that of B_P . By Fatou's Lemma, (4.8) and Lemma 2 also hold for the law $\mathbf{P}(\cdot)$. This completes the proof of Corollary 2.

Proof of Proposition 1. Let $E_n = \sqrt{n}(P_n - P)$ and $\mathcal{F} = \{I_{(-\infty,x)} : x \in \mathbb{R}^d\}$. Let $\|\cdot\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} |\cdot|$. Consider the stochastic processes $\{E_n f : f \in \mathcal{F}\}$ and $\{D_{n,v} f : f \in \mathcal{F}\}$. Obviously, $M_n = \int_{\mathbb{R}^d} (E_n I_{(-\infty,x)})^2 dP$ and $M_{n,v} = \int_{\mathbb{R}^d} (D_{n,v} I_{(-\infty,x)})^2 dP_n$. Theorem VII.21 of Pollard (1984, p.157) and Corollary 1 imply that $E_n \xrightarrow{d} B_P$ and, for almost all samples \mathbf{X} , $D_{n,v} \xrightarrow{d^*} B_P$. Theorem 13 in Pollard (1984, p.71) implies that for almost all \mathbf{X} , if \mathbf{X} is given, then there exist random elements $D'_{n,v} (n \geq 1)$ and B'_P such that

$$(1) B'_P \stackrel{d}{=} B_P ; D'_{n,v} \stackrel{d}{=} D_{n,v}, n \geq 1; (2) \|D'_{n,v} - B'_P\|_{\mathcal{F}} \longrightarrow 0 \text{ a.s.,}$$

where " $X \stackrel{d}{=} Y$ " denotes that X and Y have the same distribution. Let $M'_{n,v} = \int_{\mathbb{R}^d} (D'_{n,v} I_{(-\infty,x)})^2 dP_n$ and $M' = \int_{\mathbb{R}^d} (B'_P I_{(-\infty,x)})^2 dP$. By (1) and (2), for any $\epsilon > 0$, there exists a constant $\delta > 0$ such that if $(P(f - g)^2)^{1/2} \leq \delta$ for $f, g \in \mathcal{F}$, then

$$\|B'_P f - B'_P g\| \leq 4\epsilon \|B'_P\|_{\mathcal{F}} \text{ and } \|D'_{n,v}\|_{\mathcal{F}} \leq \|B'_P\|_{\mathcal{F}} + 1 < \infty \text{ a.s.} \quad (4.9)$$

Lemmas 25 and 36 of Pollard (1984, pp.27-34) imply that there exists an integer $m \geq 1$ and $\{f_1, \dots, f_m\} \subset \mathcal{F}$ such that $\min_{1 \leq j \leq m} (P(f - f_j)^2)^{1/2} \leq \delta$ for any $f \in \mathcal{F}$. Let $A_0 = \emptyset$ and $A_i = \{I_{(-\infty,x)} : (P(I_{(-\infty,x)} - f_i)^2)^{1/2} \leq \delta, x \in \mathbb{R}^d\}$, $1 \leq i \leq m$. Let $A_i^* = A_i - \bigcup_{j=1}^{i-1} A_j$ and $B_i = \{x : x \in \mathbb{R}^d, I_{(-\infty,x)} \in A_i^*\}$, $i = 1, \dots, m$. Obviously, $\bigcup_{i=1}^m B_i = \mathbb{R}^d$ and $B_i \cap B_j = \emptyset$ for $1 \leq i \neq j \leq m$. Let $n \rightarrow \infty$ and then $\epsilon \rightarrow 0$. By (2), (4.9) and Kolmogorov's Strong Law of Large Numbers, for almost all samples \mathbf{X} , if \mathbf{X} is fixed, then

$$\begin{aligned} |M'_{n,v} - M'| &\leq (2\|B'_P\|_{\mathcal{F}} + 1)\|B'_P - D'_{n,v}\|_{\mathcal{F}} + \left| \sum_{i=1}^m (B'_P f_i)^2 \int_{B_i} d(P_n - P) \right| \\ &\quad + \left| \sum_{i=1}^m \int_{B_i} [(B'_P I_{(-\infty,x)})^2 - (B'_P f_i)^2] d(P_n - P) \right| \\ &\leq \epsilon + (2\|B'_P\|_{\mathcal{F}} + 1)\|B'_P - D'_{n,v}\|_{\mathcal{F}} + m\|B'_P\|_{\mathcal{F}}^2 \max_{1 \leq j \leq m} |P_n I_{B_j} - P I_{B_j}| \\ &\longrightarrow 0 \text{ with } \mathbf{P}(\cdot | \mathbf{X})\text{-probability one} \end{aligned}$$

By the definition of weak convergence of sequences of random variables and the Bounded Convergence Theorem, (ii) implies (i).

Proof of Proposition 2. The proof is the same as that of Proposition 1. The details are omitted.

Acknowledgement

The author is extremely grateful to an Associate Editor and two referees for many valuable suggestions and comments. Special thanks are due to Dr. Alan T. K. Wan and Prof. Ching-Shui Cheng whose careful correction of the language greatly improved the presentation of the paper. This work is supported partially by National Natural Science Foundations of China and funds from Guangdong Higher Education Committee and City University of Hong Kong.

References

- Aerts, M. and Janssen, P. (1995). Weighted bootstrapping for U-quantiles. *Statist. Probab. Lett.* **22**, 317-323.
- Arcones, M. A. and Giné, E. (1992). On Bootstrap of U and V statistics. *Ann. Statist.* **20**, 655-674.
- Arcones, M. A. and Giné, E. (1993). Limit theorems for U-processes. *Ann. Probab.* **21**, 1494-1542.
- Arcones, M. A. and Giné, E. (1994). U-processes indexed by Vapnik-Cervonenkis classes of functions with application to asymptotics and bootstrap of U-statistics with estimated parameters. *Stochastic Process. Appl.* **52**, 17-38.
- Arcones, M. A. and Giné, E. (1995). On the law of the iterated logarithm for canonical U-statistics and processes. *Stochastic Process. Appl.* **58**, 217-245.
- Arcones, M. A. and Yu, B. (1994) Central limit theorems for empirical and U-Processes of stationary mixing sequences. *J. Theoret. Probab.* **7**, 47-71.
- Barbe, P. and Bertail, P. (1995). *The Weighting Bootstrap*. Lecture Notes in Statistics. Springer-Verlag, New York.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**, 1-26.
- Efron B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Philadelphia.
- Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209-230.
- Giné and Zinn (1990). Bootstrapping empirical processes. *Ann. Probab.* **18**, 851-869.
- Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13-30.
- Huber, P. J. (1985). Projection pursuit. *Ann. Statist.* **13**, 435-475.
- Huskova, M. and Janssen, P. (1993a). Generalized bootstrap for studentized U-statistics: a rank based approach. *Statist. Probab. Lett.* **16**, 225-233.
- Huskova, M. and Janssen, P. (1993b). Consistency of generalized bootstrap for degenerate U-statistics. *Ann. Statist.* **21**, 1181-1823.
- James, L. F. (1997). A study of a class of weighted bootstraps for censored data. *Ann. Statist.* **25**, 1595-1621.
- Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika* **30**, 81-93.
- Lee, A. J. (1990). *U-Statistics, Theory and Practice*. Marcel Dekker, New York.
- Lehmann, E. L. (1951). Consistency and unbiasedness of certain nonparametric tests. *Ann. Math. Statist.* **22**, 165-179.
- Lo, A. Y. (1983). Weak convergence for Dirichlet processes. *Sankhyā Ser. A.* **45**, 105-111.
- Lo, A. Y. (1987). A large sample study for Bayesian bootstrap. *Ann. Statist.* **15**, 360-375.
- Lo, A. Y. (1993). Bayesian bootstrap for censored data. *Ann. Statist.* **21**, 100-123.

- Nolan, D. and Pollard, D. (1987). U-processes: rates of convergence. *Ann. Statist.* **15**, 780-799.
- Nolan, D. and Pollard, D. (1988). Functional limit theorems for U-processes. *Ann. Probab.* **16**, 1291-1298.
- Pollard, D. (1984). *Convergence of Stochastic Processes*. Springer-Verlag, New York.
- Praestgaard, J. and Wellner, J. A. (1993). Exchangeably weighted bootstraps of a general empirical process. *Ann. Probab.* **21**, 2053-2086.
- Rubin, D. B. (1981). The Bayesian bootstrap. *Ann. Statist.* **9**, 130-134.
- Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley, New York.
- Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap*. Springer, New York.
- Wang, C. S. (1989). On a secondary asymptotic property of Bayesian bootstrap mean. *Ann. Statist.* **17**, 705-710.
- Zhang, D. and Li, G. (1994). Bootstrap approximation for generalized U-processes. *Chinese Sci. Bull.* **39**, 1761-1765.
- Zhang, D. (1997). Convergence of Dirichlet empirical processes indexed by a class of functions. *Statist. Probab. Lett.* **32**, 25-33.

Department of Statistics, Guangdong Commercial College, Guangzhou, 510320, Guangdong, China.

E-mail: dixinz@ihw.com.cn

(Received January 1999; accepted September 2000)