数理解析研究所講究録
1007 巻 1997 年 51-69

BOOTSTRAP TESTS FOR THE JOINT INDEPENDENCE OF VARIABLES

千葉大学大学院・自然科学 景 平 (Ping Jing)
中国科学院応用数学研究所 朱 力行 (Zhu Li-Xing)

Abstract. Testing the joint independence of variables and equality of covariance matrix has long been an interesting issue in statistics inference. To overcome the sparseness of data points in high-dimensional space and deal with the general cases, we suggest several projection pursuit type statistics. Some results on the limiting distributions of the statistics are obtained. Some properties of Bootstrap approximation are investigated. furthermore, for computational reasons an approximation for the statistics the based on Number theoretic method is applicated. Several simulation experiments are performed.

1. Introduction

Suppose that n multivariate observations $\mathbf{z}_1,\ldots,\ \mathbf{z}_n$ are collected, and that $\mathbf{z}_i = (\mathbf{z}_i^{(1)},\ldots,\mathbf{z}_i^{(d)})$, where $\mathbf{z}_i^{(j)}$, $j = 1,\ldots,d$, are made up of p_j components respectively and $\sum_{j=1}^{d} p_j = p$. A common issue is to test the joint independence of d sets of variables $\mathbf{z}_i^{(1)},\ldots,\mathbf{z}_i^{(d)}$. Several tests based on the empirical measure have been proposed. When $p = 2$ and $p_1 = p_2 = 1$, for instance, the chi-square test is available. In general case, Blum, Kiefer and Rosenblatt (1961) proposed a nonparametric distance test(B-K-R test). They suggested using

$$D_n = \sqrt{n}\sup_{\mathbf{t}\in R^p}\left|\frac{1}{n}\sum_{i=1}^{n} I(\mathbf{z}_i \leq \mathbf{t}) - \prod_{j=1}^{p}\left(\frac{1}{n}\sum_{i=1}^{n} I(z_i^{(j)} \leq t^{(j)})\right)\right|, \tag{1.1}$$

where $\mathbf{t} = (t^{(1)},\ldots,t^{(p)})$ and $\mathbf{z} \leq \mathbf{t}$ means that $z^{(j)} \leq t^{(j)}$ for $j = 1,\ldots,p$. Their test is under the assumption that the underlying distribution of $\mathbf{z}$ has a density function.

Clearly a similar version of B-K-R test can be applied to treat the problem of testing the joint independence of $\mathbf{z}^{(j)}, j = 1,\ldots,d$. However, when the dimension p is large, then sparseness problem of the sample points in high-dimensional space will be encountered unless the size of sample is gigantic. One can refer to Huber (1985) and references therewith. As projection pursuit technique is a very useful tool for overcoming such a problem of sparseness of sample points, our aim in this paper is to develop some tests based on the empirical measure and projection pursuit technique.

On the other hand, in order to determine the critical values, one need to know the properties of the sampling or limiting distributions of the test statistics proposed. Similar to B-K-R test, accurate expressions of the sampling and limiting distributions of our test statistics depend on the underlying distribution of $\mathbf{z}$ and are not tractable. In this paper, we use the bootstrap method introduced by Efron (1979) to estimate the null distributions of the test statistics.

Furthermore, the exact critical value of the test statistics, similar to that of B-K-R test, may be difficult to obtain because the test statistics proposed are the supremum and integration of function based on sample over uncountable sets in the Euclidean space, and may be hard to compute. Instead, one may have to resort to compute the values over a finite number of search sets. As in Beran and Miller (1986), a stochastic approximation can be used. We in this chapter also suggest another approximation derived by Number-theoretic method (e.g. Fang and Wang (1994)). This section is organized in such a way: Section 2 presents the construction of the test statistics. The bootstrap approximation is discussed in the same section. Number theoretic method is described in Section 3. Section 4 contains some simulation experiments and a real-life example to which the new tests are applied. Section 5 are Tests of Elliptical symmetry of Distribution.

2 Construction of Tests and Bootstrap Approximations

2.1 Test statistics and their asymptotic properties

Let $F_{\mathbf{z}}(\mathbf{t})$ be the distribution function of $\mathbf{z}$ and let $F_{\mathbf{z}^{(j)}}(\mathbf{t}^{(j)}), j = 1, \ldots, d$, be the distribution function of $\mathbf{z}^{(j)}$. As is known, $\mathbf{z}^{(j)}, j = 1, \ldots, d$, are jointly independent if and only if $F_{\mathbf{z}}(\cdot) \equiv \prod_{j=1}^{d} F_{\mathbf{z}^{(j)}}(\cdot)$. This is the basis of constructing B-K-R test. In order to construct our tests via projection pursuit technique, we give another version of necessary and sufficient condition of the joint independence of $\mathbf{z}^{(j)}$'s. Define

$$S_j = \{\mathbf{a}_j : \|\mathbf{a}_j\| = 1, \mathbf{a}_j \in R^{p_j}\}, \quad j = 1, \ldots, d, \tag{2.1}$$

where the notation $\|\cdot\|$ stands for the Euclidean norm in R^{p_j}.

LEMMA 2.1. *$\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(d)}$ are jointly independent of each other if and only if $\mathbf{a}_1^\tau \mathbf{z}^{(1)}, \ldots, \mathbf{a}_d^\tau \mathbf{z}^{(d)}$ are jointly independent of each other for all $\mathbf{a}_j \in S_j, j = 1, \ldots, d$.*

Proof. The necessity is obvious. We now show the sufficiency. Let $\varphi_{\mathbf{a}_j^\tau \mathbf{z}^{(j)}}(h_j)$ and $\varphi_{\mathbf{z}^{(j)}}(\mathbf{h}_j)$ be, respectively, the characteristic functions of $\mathbf{a}_j^\tau \mathbf{z}^{(j)}$ and $\mathbf{z}^{(j)}$. It is easy to see that $\varphi_{\mathbf{a}_j^\tau \mathbf{z}^{(j)}}(h_j) = \varphi_{\mathbf{z}^{(j)}}(h_j \mathbf{a}_j)$ for $j = 1, \cdots, d$. Furthermore, let $\varphi_{\mathbf{a}_1^\tau \mathbf{z}^{(1)}, \cdots, \mathbf{a}_d^\tau \mathbf{z}^{(d)}}(h_1, \cdots, h_d)$ and $\phi_{\mathbf{z}^{(1)}, \cdots, \mathbf{z}^{(d)}}(\mathbf{h}_1, \cdots, \mathbf{h}_d)$ be the characteristic functions of $(\mathbf{a}_1^\tau \mathbf{z}^{(1)}, \cdots, \mathbf{a}_d^\tau \mathbf{z}^{(d)})$ and $(\mathbf{z}^{(1)}, \cdots, \mathbf{z}^{(d)})$ respectively. We can also get that

$$\varphi_{\mathbf{a}_1^\tau \mathbf{z}^{(1)}, \cdots, \mathbf{a}_d^\tau \mathbf{z}^{(d)}}(h_1, \cdots, h_d) = \varphi_{\mathbf{z}^{(1)}, \cdots, \mathbf{z}^{(d)}}(h_1 \mathbf{a}_1, \cdots, h_d \mathbf{a}_d). \tag{2.2}$$

All we need to do is to prove that, for showing the sufficiency,

$$\varphi_{\mathbf{z}^{(1)}, \cdots, \mathbf{z}^{(d)}}(\mathbf{h}_1, \cdots, \mathbf{h}_d) = \prod_{j=1}^{d} \varphi_{\mathbf{z}^{(j)}}(\mathbf{h}_j). \tag{2.3}$$

When $(\mathbf{a}_1^\tau \mathbf{z}^{(1)}, \cdots, \mathbf{a}_d^\tau \mathbf{z}^{(d)})$ are jointly independent for every group of $(\mathbf{a}_1, \cdots, \mathbf{a}_d)$, we then have

$$\varphi_{\mathbf{a}_1^\tau \mathbf{z}^{(1)}, \cdots, \mathbf{a}_d^\tau \mathbf{z}^{(d)}}(h_1, \cdots, h_d) = \prod_{j=1}^{d} \varphi_{\mathbf{a}_j^\tau \mathbf{z}^{(j)}}(h_j). \tag{2.4}$$

Based on the above discussion, we can derive that

$$\varphi_{\mathbf{z}^{(1)}, \cdots, \mathbf{z}^{(d)}}(h_1 \mathbf{a}_1, \cdots, h_d \mathbf{a}_d) = \prod_{j=1}^{d} \varphi_{\mathbf{z}^{(j)}}(h_j \mathbf{a}_j) \tag{2.5}$$

holds for every group of $\mathbf{a}_j, j = 1, \cdots, d$. Note that $\mathbf{h}_j$ can be expressed as $\|\mathbf{h}_j\| \cdot \mathbf{h}_j / \|\mathbf{h}_j\|$ where $\mathbf{h}_j / \|\mathbf{h}_j\| \in S_j$. Hence (2.3) is showed, which completes the proof.

Based on this fact, we can construct the tests for the joint independence of $\mathbf{z}^{(j)}, j = 1, \ldots, d$. Denote by $F_{n\mathbf{a}_j}(t_j)$ the empirical distribution determined by $\mathbf{a}_j^\tau \mathbf{z}_1^{(j)}, \ldots, \mathbf{a}_j^\tau \mathbf{z}_n^{(j)}$ and let $F_{n\mathbf{a}}(t_1, \ldots, t_d)$ be the empirical distribution based on $\mathbf{a}\mathbf{z}_1, \ldots, \mathbf{a}\mathbf{z}_n$, where $\mathbf{a}\mathbf{z}_i = (\mathbf{a}_1^\tau \mathbf{z}_i^{(1)}, \ldots, \mathbf{a}_d^\tau \mathbf{z}_i^{(d)})$ and $\mathbf{a} = (\mathbf{a}_1, \ldots, \mathbf{a}_d)$. Furthermore, let $\mathbf{t} = (t_1, \ldots, t_d)$ and $\mathbf{S} = S_1 \otimes, \cdots, \otimes S_d$. A B-K-R type test statistic is

$$KS_n = \sup_{\mathbf{a} \in \mathbf{S}, \mathbf{t} \in R^d} |K_n(\mathbf{a}, \mathbf{t})| = \sqrt{n} \sup_{\mathbf{a} \in \mathbf{S}, \mathbf{t} \in R^d} |F_{n\mathbf{a}}(\mathbf{t}) - \prod_{j=1}^{d} F_{n\mathbf{a}_j}(t_j)|. \tag{2.6}$$

Two other test statistics are the following Cramer-Von Mises type tests;

$$CVS_n = \sup_{\mathbf{a} \in \mathbf{S}} \int_{R^d} (K_n(\mathbf{a}, \mathbf{t}))^2 \prod_{j=1}^{d} dF_{n\mathbf{a}_j}(t_j), \tag{2.7}$$

and

$$CVA_n = \int_{\mathbf{S}} \int_{R^d} (K_n(\mathbf{a}, \mathbf{t}))^2 \prod_{j=1}^{d} dF_{n\mathbf{a}_j}(t_j) \prod_{j=1}^{d} d\mu_j(\mathbf{a}_j), \tag{2.8}$$

where $\mu_j(\cdot)$ is the uniform distribution on S_j.

In view of the construction procedure above, the test statistics are based on the low-dimensional projection of high-dimensional data. This is available for overcoming the sparseness of data in high-dimensional space. One can refer to Huber (1985) for details.

In order to present the asymptotic behavior of the test statistics, we define several centered continuous Gaussian processes $\mathbf{W} = \{W(\cdot, \mathbf{a}, \mathbf{t}) : \mathbf{a} \in \mathbf{S}, \mathbf{t} \in R^d\}$ and $\mathbf{W}_j = \{W_j(\cdot, \mathbf{a}_j, t_j) : \mathbf{a}_j \in S_j, t_j \in R^1\}, j = 1, \ldots, d$. These processes have the following properties;

1) each sample path of $W(\cdot, \mathbf{a}, \mathbf{t})$ is bounded and uniformly continuous with respect to the metric induced by the $L^2(P)$-semi-norm on $\mathfrak{F} = \{f = I(\mathbf{az} \leq \mathbf{t}) : \mathbf{a} \in \mathbf{S}, \mathbf{t} \in R^d\}$;

2) the covariance function of $W(\cdot, \mathbf{a}, \mathbf{t})$ is of the form

$$
\begin{aligned}
E\{W(\cdot, \mathbf{a}, \mathbf{t})W(\cdot, \mathbf{b}, \mathbf{s})\} &= \int I(\mathbf{az} \leq \mathbf{t})I(\mathbf{bz} \leq \mathbf{s})dP \\
&\quad - \int I(\mathbf{az} \leq \mathbf{t})dP \int I(\mathbf{bz} \leq \mathbf{s})dP;
\end{aligned} \tag{2.9}
$$

3) each sample path of $W_j(\cdot, \mathbf{a}_j, t_j)$ is bounded and uniformly continuous with respect to the metric induced by the $L^2(P^{(j)})$-semi-norm on $\mathfrak{F}_j = \{f = I(\mathbf{a}_j^\tau \mathbf{z}^{(j)} \leq t_j) : \mathbf{a}_j \in S_j, t_j \in R^1\}$;

4) the covariance function of $W_j(\cdot, \mathbf{a}_j, t_j)$ is of the form

$$
\begin{aligned}
E\{W_j(\cdot, \mathbf{a}_j, t_j)W_j(\cdot, \mathbf{b}_j, s_j)\} &= \int I(\mathbf{a}_j^\tau \mathbf{z}^{(j)} \leq t_j)I(\mathbf{b}_j^\tau \mathbf{z}^{(j)} \leq s_j)dP^{(j)} \\
&\quad - F_{\mathbf{a}_j}(t_j)F_{\mathbf{b}_j}(s_j);
\end{aligned} \tag{2.10}
$$

where P and $P^{(j)}$ are the probability measures of $\mathbf{z}$ and $\mathbf{z}^{(j)}$ respectively, and $F_{\mathbf{a}_j}(t_j)$ stands for the distribution function of $\mathbf{a}_j^\tau \mathbf{z}^{(j)}$.

We now present the asymptotic behavior of the test statistics defined in (2.6), (2.7) and (2.8). For the simplicity of notation, let

$$
\widehat{W}(\cdot, \mathbf{a}, \mathbf{t}) = W(\cdot, \mathbf{a}, \mathbf{t}) - \sum_{j=1}^{d} \prod_{i \neq j}^{d} F_{\mathbf{a}_i}(t_i) W_j(\cdot, \mathbf{a}_j, t_j). \tag{2.11}
$$

THEOREM 2.1. *Suppose that all $F_{\mathbf{a}_j}(t_j), j = 1, \ldots, d$, and $F_{\mathbf{a}}(\mathbf{t})$, the distribution of $\mathbf{az}$, are continuous with respect to $\mathbf{a}$ and $\mathbf{t}$. If $\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(d)}$ are jointly independent of each other, then*

$$
KS_n \Longrightarrow \sup_{\mathbf{a} \in \mathbf{S}, \mathbf{t} \in R^d} |\widehat{W}(\cdot, \mathbf{a}, \mathbf{t})|, \tag{2.12}
$$

$$(2.13) \qquad CVS_n \Longrightarrow \sup_{\mathbf{a}\in \mathbf{S}} \int_{R^d} (\widehat{W}(\cdot,\mathbf{a},\mathbf{t}))^2 \prod_{j=1}^{d} dF_{\mathbf{a}_j}(t_j),$$

and

$$(2.14) \qquad CVA_n \Longrightarrow \int_{\mathbf{S}} \int_{R^d} (\widehat{W}(\cdot,\mathbf{a},\mathbf{t}))^2 \prod_{j=1}^{d} dF_{\mathbf{a}_j}(t_j) \prod_{j=1}^{d} d\mu_j(\mathbf{a}_j),$$

where the notation"$\Longrightarrow$" means the weak convergence. If the null hypothesis is false, then the above three statistics tend to infinity with probability one as $n \to \infty$.

For simplicity of the notations, we consider $d = 2$. The general case can be proved in the same way.

Proof of Theorem 2.1. Using the mondern theory of empirical processes, we can show

$$(2.15) \qquad \sup_{\mathbf{a}\in S,\mathbf{t}\in R^d} |\sqrt{n}(F_{n\mathbf{a}}(\mathbf{t}) - F_{\mathbf{a}}(\mathbf{t}))| \Longrightarrow \sup_{\mathbf{a}\in S,\mathbf{t}\in R^d} |W(\cdot,\mathbf{a},\mathbf{t})|, \quad a.s.$$

and

$$(2.16) \qquad \sup_{\mathbf{a}_j\in S_j,t_j\in R^1} |\sqrt{n}(F_{n\mathbf{a}_j}(t_j) - F_{\mathbf{a}_j}(t_j))| \Longrightarrow \sup_{\mathbf{a}_j\in S_j,t_j\in R^1} |W_j(\cdot,\mathbf{a}_j,t_j)|, \quad a.s.$$

where $F_{\mathbf{a}}(\mathbf{t})$ and $F_{\mathbf{a}_j}(t_j)$ are the distribution functions of $\mathbf{az} = (\mathbf{a}_1'\mathbf{z}^{(1)}, \mathbf{a}_2'\mathbf{z}^{(2)})$ and of $\mathbf{a}_j'\mathbf{z}^{(j)}$ respectively, $F_{n\mathbf{a}}(\mathbf{t})$ and $F_{n\mathbf{a}_j}(t_j)$ are the corresponding empirical distributions, and $W(\cdot,\mathbf{a},\,\mathbf{t})$ and $W_j(\cdot,\mathbf{a}_j,t_j)$ are the centered continuous Gaussian processes defined in Section 2 associated with the covariance functions in (2.9) and (2.10) respectively. Note that

$$\begin{aligned} F_{n\mathbf{a}}(\mathbf{t}) - F_{n\mathbf{a}_1}(t_1)F_{n\mathbf{a}_2}(t_2) = {} & (F_{n\mathbf{a}}(\mathbf{t}) - F_{\mathbf{a}}(\mathbf{t})) - F_{n\mathbf{a}_1}(t_1)(F_{n\mathbf{a}_2}(t_2) - F_{\mathbf{a}_2}(t_2)) \\ & - F_{\mathbf{a}_2}(t_2)(F_{n\mathbf{a}_1}(t_1) - F_{\mathbf{a}_1}(t_1)) \\ & + F_{\mathbf{a}}(\mathbf{t}) - F_{\mathbf{a}_1}(t_1)F_{\mathbf{a}_2}(t_2)). \end{aligned}$$

The formulas (2.12) — (2.14) are the direct consequence of (2.15) and (2.16) when the null hypothesis is true, namely, $F_{\mathbf{a}}(\mathbf{t}) - F_{\mathbf{a}_1}(t_1)F_{\mathbf{a}_2}(t_2) \equiv 0$. If the null is false, then $\sqrt{n}\sup_{\mathbf{a},\mathbf{t}} |F_{\mathbf{a}}(\mathbf{t}) - F_{\mathbf{a}_1}(t_1)F_{\mathbf{a}_2}(t_2)| \to \infty$. This leads that three statistics in (2.6) — (2.8) tend to infinity. The proof is finished.

2.2 Bootstrap tests

In previous subsection, we construct the test statistics and present the asymptotic behavior. A serious problem is that neither the sampling nor limiting distribution of the statistics is tractable because they depend on the underlying distribution of $\mathbf{z}$. This problem leads up to that the critical values may be difficult to be determined. We now apply the bootstrap tests to solve this problem.

Let $\mathbf{z}_{1*}^{(j)}, \ldots, \mathbf{z}_{n*}^{(j)}$ be bootstrap samples drawn from the empirical marginal distribution $F_n^{(j)}(\cdot)$ based on $\mathbf{z}_1^{(j)}, \ldots, \mathbf{z}_n^{(j)}$, $j = 1, \ldots, d$, and let $\mathbf{z}_{i*} = (\mathbf{z}_{i*}^{(1)}, \ldots, \mathbf{z}_{i*}^{(d)})$, that is, $\mathbf{z}_{1*}, \ldots, \mathbf{z}_{n*}$ is the bootstrap sample drawn from $\prod_{j=1}^{d} F_n^{(j)}(\cdot)$. Similar to the definition of $F_{n\mathbf{a}}(\mathbf{t})$ and $F_{n\mathbf{a}_j}(t_j)$, let $F_{n\mathbf{a}}^*(\mathbf{t})$ and $F_{n\mathbf{a}_j}^*(t_j), j = 1, \ldots, d$, be the empirical distributions determined by $\{\mathbf{a}\mathbf{z}_{1*}, \ldots, \mathbf{a}\mathbf{z}_{n*}\}$ and $\{\mathbf{a}_j^\tau \mathbf{z}_{1*}^{(j)}, \ldots, \mathbf{a}_j^\tau \mathbf{z}_{n*}^{(j)}\}$ respectively. Define

$$K_n^*(\mathbf{a}, \mathbf{t}|\mathbf{z}_1, \ldots, \mathbf{z}_n) = \sqrt{n}(F_{n\mathbf{a}}^*(\mathbf{t}) - \prod_{j=1}^{d} F_{n\mathbf{a}_j}^*(t_j)). \tag{2.17}$$

The bootstrap test statistics are defined by

$$KS_n^*(\mathbf{z}_1, \ldots, \mathbf{z}_n) = \sup_{\mathbf{a}\in\mathbf{S}, \mathbf{t}\in R^d} |K_n^*(\mathbf{a}, \mathbf{t}|\mathbf{z}_1, \ldots, \mathbf{z}_n)|, \tag{2.18}$$

$$CVS_n^*(\mathbf{z}_1, \ldots, \mathbf{z}_n) = \sup_{\mathbf{a}\in\mathbf{S}} \int_{R^d} (K_n^*(\mathbf{a}, \mathbf{t}|\mathbf{z}_1, \ldots, \mathbf{z}_n))^2 \prod_{j=1}^{d} dF_{n\mathbf{a}_j}^*(t_j), \tag{2.19}$$

and

(2.20)

$$CVA_n^*(\mathbf{z}_1, \ldots, \mathbf{z}_n) = \int_{\mathbf{S}} \int_{R^d} (K_n^*(\mathbf{a}, \mathbf{t}|\mathbf{z}_1, \ldots, \mathbf{z}_n))^2 \prod_{j=1}^{d} dF_{n\mathbf{a}_j}^*(t_j) \prod_{j=1}^{d} d\mu_j(\mathbf{a}_j).$$

The following theorem shows that the test statistics and their bootstrap versions have the same limit if the null hypothesis is true.

THEOREM 2.2. *If the null hypothesis holds, then, with probability one,*

$$KS_n^*(\mathbf{z}_1, \ldots, \mathbf{z}_n) \Longrightarrow \sup_{\mathbf{a}\in\mathbf{S}, \mathbf{t}\in R^d} |\widehat{W}(\cdot, \mathbf{a}, \mathbf{t})|, \tag{2.21}$$

$$CVS_n^*(\mathbf{z}_1, \ldots, \mathbf{z}_n) \Longrightarrow \sup_{\mathbf{a}\in\mathbf{S}} \int_{R^d} (\widehat{W}(\cdot, \mathbf{a}, \mathbf{t}))^2 \prod_{j=1}^{d} dF_{\mathbf{a}_j}(t_j), \tag{2.22}$$

$$(2.23) \qquad CVA_n^*(\mathbf{z}_1, \dots, \mathbf{z}_n) \Longrightarrow \int_{\mathbf{S}} \int_{R^d} (\widehat{W}(\cdot, \mathbf{a}, \mathbf{t}))^2 \prod_{j=1}^{d} dF_{\mathbf{a}_j}(t_j) \prod_{j=1}^{d} d\mu_j(\mathbf{a}_j).$$

If the null hypothesis is false, the above three statistics still converge weakly to the maximum or integration functionals of certain Gaussian processes, with probability one.

Proof of Theorem 2.2. First recall that the bootstrap sample $(\mathbf{z}_{1*}, \dots, \mathbf{z}_{n*})$ is an i.i.d. sample from the distribution $\prod_{j=1}^2 P_n^{(j)}(\cdot)$, and the components of $\mathbf{z}_{i*}$ are conditionally independent of each other, given $\{\mathbf{z}_1, \dots, \mathbf{z}_n\}$, where $P_n^{(j)}(\cdot)$ is the probability measure based on $(\mathbf{z}_1^{(j)}, \dots, \mathbf{z}_n^{(j)})$. Note that

(2.24)

$$F_{n\mathbf{a}}^*(\mathbf{t}) - \prod_{j=1}^{2} F_{n\mathbf{a}_j}^*(t_j) = \{F_{n\mathbf{a}}^*(\mathbf{t}) - \prod_{j=1}^{2} F_{n\mathbf{a}_j}(t_j)\} - \{\prod_{j=1}^{2} F_{n\mathbf{a}_j}^*(t_j) - \prod_{j=1}^{2} F_{n\mathbf{a}_j}(t_j)\}.$$

Now we apply Corollary 2.7 of Giné and Zinn (1991, p.771) to obtain the asymptotic distribution of the first term in the right hand side of (2.24). To check the conditions there, first notice that $\mathfrak{F}$ is a measurable finitely uniformly pregaussian class (see Giné and Zinn (1991, p.761 and p.778)). Next, let $\mathfrak{F}' = \{f - g : f, g \in \mathfrak{F}\}$, $(\mathfrak{F}')^2 = \{(f-g)^2 : f, g \in \mathfrak{F}\}$ and $\mathfrak{G} = \mathfrak{F} \cup \mathfrak{F}' \cup (\mathfrak{F}')^2$. Then it turns out that

$$(2.25) \qquad \sup_{g \in \mathfrak{G}} | \prod_{j=1}^{2} P_n^{(j)}(g) - \prod_{j=1}^{2} P^{(j)}(g)| \to 0, \quad a.s.$$

where $P^{(j)}(\cdot)$ is the associated probability measure of the distribution $F_{\mathbf{z}^{(j)}}(\cdot)$, since the metric entropy of $\mathfrak{G}$ (Giné and Zinn (1986, p.53)) is finite. Consequently it follows from Corollary 2.7 of Giné and Zinn (1991) that, with probability one,

$$\sqrt{n}(F_{n\mathbf{a}}^*(\mathbf{t}) - \prod_{j=1}^{2} F_{n\mathbf{a}_j}(t_j)) \Longrightarrow W^*(\cdot, \mathbf{a}, \mathbf{t}),$$

where $W^*(\cdot, \mathbf{a}, \mathbf{t})$ is a Gaussian process with the zero mean and covariance kernel:

$$E\{W^*(\cdot, \mathbf{a}, \mathbf{t}) W^*(\cdot, \mathbf{b}, \mathbf{s})\} = \prod_{j=1}^{2} \int I(\mathbf{a}_j' \mathbf{z}^{(j)} \leq t_j) I(\mathbf{b}_j' \mathbf{z}^{(j)} \leq s_j) dP^{(j)}$$
$$- \prod_{j=1}^{2} F_{\mathbf{a}_j}(t_j) \prod_{j=1}^{2} F_{\mathbf{b}_j}(s_j),$$

for $\mathbf{a} = (\mathbf{a}_1, \mathbf{a}_2), \mathbf{b} = (\mathbf{b}_1, \mathbf{b}_2)$ with $\mathbf{a}_j, \mathbf{b}_j \in S_j$ and $\mathbf{t} = (t_1, t_2), \mathbf{s} = (s_1, s_2)$, with real numbers t_j's and s_j's. Comparing this covariance kernel with the one of W

in (2.11), it is clear that, when the null hypothesis is true, W^* is just W in (2.11). Also note that the term inside the second curly parenthesis in (2.24) equals

$$F^*_{n\mathbf{a}_1}(t_1)(F^*_{n\mathbf{a}_2}(t_2) - F_{n\mathbf{a}_2}(t_2)) + F_{n\mathbf{a}_2}(t_2)(F^*_{n\mathbf{a}_1}(t_1) - F_{n\mathbf{a}_1}(t_1)).$$

Invoking again Corollary 2.7 of Giné and Zinn (1991) and checking the conditions there, we can derive that the bootstrap empirical process associated with $F_{n\mathbf{a}_1}(t_1)F_{n\mathbf{a}_2}(t_2)$ converges weakly to a process associated with the second term in (2.11). Then when the null hypothesis is true, we have

$$K^*_n(\mathbf{a},\mathbf{t}|\mathbf{z}_1,\dots,\mathbf{z}_n) = \sqrt{n}(F^*_{n\mathbf{a}}(\mathbf{t}) - \prod_{j=1}^{2} F^*_{n\mathbf{a}_j}(t_j)) \Longrightarrow \widehat{W}(\cdot,\mathbf{a},\mathbf{t}). \tag{2.26}$$

On the other hand, if the null is false, the bootstrap empirical process still converges weakly to $W^*(\cdot,\mathbf{a},\mathbf{t}) - \sum_{j=1}^{2}\prod_{i\neq j}^{2} F_{\mathbf{a}_i}(t_i)W_j(\cdot,\mathbf{a}_j,t_j)$. The formulas (2.21) — (2.23) are the direct consequence of (2.26), which completes the proof.

3 The Approximations for The Bootstrap Test

In view of the bootstrap test statistics, they are the supremum or the integration over the Euclidean space. For the computational reason, one may have to resort to compute $K^*_n(\mathbf{a},\mathbf{t}|\mathbf{z}_1,\dots,\mathbf{z}_n)$ over a finite number of points. For given $\mathbf{a}_i^1, \dots, \mathbf{a}_i^m \in S_i$, $i = 1, \dots, d$, we define

$$KS^*_n(\mathbf{z}_1,\dots,\mathbf{z}_n) = \max_{\mathbf{S}^f}\sup_{\mathbf{t}} |K^*_n(\mathbf{a},\mathbf{t}|\mathbf{z}_1,\dots,\mathbf{z}_n)|, \tag{3.1}$$

$$CVS^*_n(\mathbf{z}_1,\dots,\mathbf{z}_n) = \max_{\mathbf{S}^f}\int_{R^d}(K^*_n(\mathbf{a},\mathbf{t}|\mathbf{z}_1,\dots,\mathbf{z}_n))^2\prod_{i=1}^{d} dF^*_{n\mathbf{a}_i}(t_i), \tag{3.2}$$

and

$$CVA^*_n(\mathbf{z}_1,\dots,\mathbf{z}_n) = \frac{1}{m^d}\sum_{\mathbf{S}^f}\int_{R^d}(K^*_n(\mathbf{a},\mathbf{t}|\mathbf{z}_1,\dots,\mathbf{z}_n))^2\prod_{i=1}^{d} dF^*_{n\mathbf{a}_i}(t_i), \tag{3.3}$$

where $\mathbf{S}^f = \{\mathbf{a} = (\mathbf{a}_1,\dots,\mathbf{a}_d) : \mathbf{a}_i \in \{\mathbf{a}_i^1,\dots,\mathbf{a}_i^m\} \in S_i; i = 1,\dots,d\}$. On the above definition, maximum and summation are taken over all different sets of S^f. A stochastic approximation (cf. Beran and Miller (1986)) is, of course, a choice. That is, let $\mathbf{a}_i^1,\dots,\mathbf{a}_i^m$, $i = 1,\dots,d$, be i.i.d. unit random vectors distributed uniformly on S_i. We know that the uniformity of $\{\mathbf{a}_i^1,\dots,\mathbf{a}_i^m\}$ on S_i is important for this kind of approximation. We now suggest another method of choosing $\{\mathbf{a}_i^1,\dots,\mathbf{a}_i^m\}$

, the Number-theoretic method (NTM) (e.g. Hua and Wang (1981), or Fang and Wang (1994)). It is well known that the Kolmogorov distance for $\{\mathbf{a}_i^1, \ldots, \mathbf{a}_i^m\}$ is evaluated as

$$\sup_{\delta\in\Delta} |\frac{1}{m}\sum_{k=1}^{m} I(\mathbf{a}_i^k \in \delta) - \mu(\delta)| = O_p(m^{-1/2}),$$

where $\Delta = \{\delta(\mathbf{v}) : \mathbf{v} \in [0,1]^{p_i-1}\}$ and δ is a set of the form

$$\begin{aligned}\delta(v_1, \ldots, v_{p_i-1}) = \{\mathbf{a}_i = [&\cos(f_1(u_1)), \sin(f_1(u_1))\cos(f_2(u_2)), \ldots,\\ &(\prod_{j=1}^{p_i-2} \sin(f_j(u_j)))\cos(2\pi u_{p_i-1}), (\prod_{j=1}^{p_i-2} \sin(f_j(u_j)))\sin(2\pi u_{p_i-1})] :\\ &0 \le u_j \le v_j \le 1, \quad j = 1, \ldots, p_i - 1\}\end{aligned}$$

and $f_j(u) = F_j^{-1}(u)$, $j = 1, \ldots, p_i - 2$. Here $F_j(y)$ is the distribution with the density function $g_j(y) = c(j)(\sin y)^{p_i-j-1}$, where $y \in [0, \pi]$ and $c(j)$ being a normalized constant. In the sense of the Kolmogorov distance, $\{\dot{\mathbf{a}}_i^1, \ldots, \dot{\mathbf{a}}_i^m\}$, chosen by Number-theoretic method will enjoy better uniformity on S_i than that of the above set of random vectors $\{\mathbf{a}_i^1, \ldots, \mathbf{a}_i^m\}$, that is,

$$\sup_{\delta\in\Delta} |\frac{1}{m}\sum_{k=1}^{m} I(\dot{\mathbf{a}}_i^k \in \delta) - \mu(\delta)| = O(m^{-1}(\log m)^{p_i-1}).$$

Recently, Fang and Wang (1994) gave a systematic study on application of NTM in Statistics and gave the corresponding algorithm that is so-called TFWW algorithm. They pick up the good lattice point (glp)method, one of NTM, to generate an NT-net on $[0,1]^{p_i-1}$ as follows: For given integer m larger than $p_i - 1$, choose an integer vector $\mathbf{h} = (h_1, \ldots, h_{p_i-1})$ satisfying $1 \le h_j \le m$, $h_j \ne h_l$ for $j \ne l$. Denote $c_{kj} = \{(2kh_j - 1)/2m\}$ and $\mathbf{c}_k = (c_{k1}, \ldots, c_{k(p_i-1)})^\tau$ for $k = 1, \ldots, m$ and $j = 1, \ldots, p_i - 1$, where $\{x\}$ denotes the fraction part of x. A choice of $\mathbf{h} = (h_1, \ldots, h_{p_i-1})$ can be found in the appendix of Hua and Wang (1981) or Fang and Wang (1993). We can use the TFWW algorithm to generate the desired points $\{\dot{\mathbf{a}}_i^1, \cdots, \dot{\mathbf{a}}_i^m\}$ on S_i corresponding to $\{\mathbf{c}_1, \ldots, \mathbf{c}_m\}$.

4 Simulations and an Example

To demonstrate the power of the proposed test for the joint independence of variables, we in this section apply it to several simulated data sets and a real-life example. In simulation experiments, we consider that $\mathbf{z}$ is 6-dimensional and both $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$ are 3-dimensional, that is, $\mathbf{z} = (\mathbf{z}^{(1)}, \mathbf{z}^{(2)})$. It is expected that our tests would be also powerful in higher dimensional cases. The first we consider

$\mathbf{z}^{(1)} \sim \mathbf{U}_3$;$\mathbf{z}^{(2)} \sim \mathbf{U}_3$ and $\mathbf{z}^{(1)}$ and $\mathbf{z}^{(2)}$ are independent, where $\mathbf{U}_3$ is 3-dimensional distribution with i.i.d. marginal distributions each having the uniform distribution for $[0, 1]$. Several simulated data sets are generated from the multivariate normal distributions with different covariance matrices. For these distributions, the various sample sizes, say $n = 20$ and $n = 30$, are investigated. In each case, 5 and 7 projection directions are chosen by Number-theoretic method and 1000 replications are performed. In order to simulate the critical values under the significance level 1%, 5%, 10%, 500 bootstrap samples are generated for each replication. Let $N(0, \mathbf{V}_6^{(i)})$ be the normal distribution with mean zero and covariance matrix $\mathbf{V}_6^{(i)}$. The following $\mathbf{V}_6^{(i)}$'s are considered, and $\mathbf{V}_6^{(9)}$ and $\mathbf{V}_6^{(10)}$ are concerned with the real example. The simulation results are summarized in the Tables 2.0, 2.1, 2.2, and 2.3 below. In the tables, the notation "*dir*" stands for the number of projective directions.

$$\mathbf{V}_6^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{V}_6^{(2)} = \begin{pmatrix} 1 & 0.5 & 0 & 0 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 & 1 \end{pmatrix},$$

$$\mathbf{V}_6^{(3)} = \begin{pmatrix} 1 & 0.5 & 0 & 0 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 1 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 1 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0 & 0 & 0.5 & 1 \end{pmatrix},$$

$$\mathbf{V}_6^{(4)} = \begin{pmatrix} 1 & 0.5 & 0.5 & 0 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 & 1 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0 & 0.5 & 0.5 & 1 \end{pmatrix},$$

$$\mathbf{V}_6^{(5)} = \begin{pmatrix} 1 & 0.5 & 0.5 & 0.5 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 & 0.5 & 1 & 0.5 \\ 0 & 0 & 0.5 & 0.5 & 0.5 & 1 \end{pmatrix},$$

$$\mathbf{V}_6^{(6)} = \begin{pmatrix} 1 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 1 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 1 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 1 \end{pmatrix},$$

$$\mathbf{V}_6^{(7)} = \begin{pmatrix} 1 & 0.5 & 0.5^2 & 0.5^3 & 0.5^4 & 0.5^5 \\ 0.5 & 1 & 0.5 & 0.5^2 & 0.5^3 & 0.5^4 \\ 0.5^2 & 0.5 & 1 & 0.5 & 0.5^2 & 0.5^3 \\ 0.5^3 & 0.5^2 & 0.5 & 1 & 0.5 & 0.5^2 \\ 0.5^4 & 0.5^3 & 0.5^2 & 0.5 & 1 & 0.5 \\ 0.5^5 & 0.5^4 & 0.5^3 & 0.5^2 & 0.5 & 1 \end{pmatrix},$$

$$\mathbf{V}_6^{(8)} = \begin{pmatrix} 1 & 0.9 & 0.8 & 0.7 & 0.6 & 0.5 \\ 0.9 & 1 & 0.9 & 0.8 & 0.7 & 0.6 \\ 0.8 & 0.9 & 1 & 0.9 & 0.8 & 0.7 \\ 0.7 & 0.8 & 0.9 & 1 & 0.9 & 0.8 \\ 0.6 & 0.7 & 0.8 & 0.9 & 1 & 0.9 \\ 0.5 & 0.6 & 0.7 & 0.8 & 0.9 & 1 \end{pmatrix},$$

$$\mathbf{V}_6^{(9)} = \begin{pmatrix} 1 & 0.486 & 0.678 & 0.366 & 0.448 & 0.486 \\ 0.486 & 1 & 0.857 & 0.636 & 0.403 & 0.417 \\ 0.678 & 0.857 & 1 & 0.681 & 0.520 & 0.558 \\ 0.366 & 0.636 & 0.681 & 1 & 0.345 & 0.367 \\ 0.448 & 0.403 & 0.520 & 0.345 & 1 & 0.820 \\ 0.486 & 0.417 & 0.558 & 0.367 & 0.820 & 1 \end{pmatrix},$$

$$\mathbf{V}_6^{(10)} = \begin{pmatrix} 1 & 0.737 & 0.676 & 0.476 & 0.483 & 0.540 \\ 0.737 & 1 & 0.627 & 0.339 & 0.392 & 0.446 \\ 0.676 & 0.627 & 1 & 0.441 & 0.447 & 0.440 \\ 0.476 & 0.339 & 0.441 & 1 & 0.452 & 0.535 \\ 0.483 & 0.392 & 0.447 & 0.452 & 1 & 0.663 \\ 0.540 & 0.446 & 0.440 & 0.535 & 0.663 & 1 \end{pmatrix}.$$

TABLE 2.1 **Proportion of rejecting** H_0 ($dir = 5, \quad n = 20$)

distribution	KS_n	CV_n	CVA_n
	1% 5% 10%	1% 5% 10%	1% 5% 10%
$N(0, \mathbf{V}_6^{(1)})$	0.012;0.062;0.128	0.004;0.044;0.110	0.004;0.052;0.108
$N(0, \mathbf{V}_6^{(2)})$	0.012;0.063;0.126	0.006;0.045;0.114	0.005;0.057;0.112
$N(0, \mathbf{V}_6^{(3)})$	0.013;0.066;0.135	0.006;0.047;0.112	0.005;0.059;0.119
$N(0, \mathbf{V}_6^{(4)})$	0.205;0.464;0.629	0.165;0.490;0.649	0.196;0.526;0.702
$N(0, \mathbf{V}_6^{(5)})$	0.237;0.539;0.681	0.217;0.551;0.713	0.263;0.579;0.735
$N(0, \mathbf{V}_6^{(6)})$	0.252;0.569;0.710	0.236;0.584;0.742	0.271;0.593;0.752
$N(0, \mathbf{V}_6^{(7)})$	0.035;0.146;0.254	0.016;0.085;0.186	0.013;0.090;0.207
$N(0, \mathbf{V}_6^{(8)})$	0.650;0.867;0.930	0.712;0.925;0.963	0.691;0.930;0.969
$N(0, \mathbf{V}_6^{(9)})$	0.274;0.616;0.736	0.260;0.602;0.767	0.257;0.595;0.744
$N(0, \mathbf{V}_6^{(10)})$	0.179;0.489;0.604	0.165;0.442;0.600	0.184;0.468;0.626

TABLE 2.2 **Proportion of rejecting** H_0 ($dir = 5, \quad n = 20$)

distribution	KS_n	CV_n	CVA_n
	1% 5% 10%	1% 5% 10%	1% 5% 10%
$\mathbf{U}_3$ and $\mathbf{U}_3$	0.012;0.063;0.115	0.010;0.037;0.108	0.012;0.047;0.097

TABLE 2.3 **Proportion of rejecting** H_0 ($dir = 5, \quad n = 30$)

distribution	KS_n	CV_n	CVA_n
	1% 5% 10%	1% 5% 10%	1% 5% 10%
$N(0, \mathbf{V}_6^{(1)})$	0.010;0.059;0.118	0.009;0.044;0.099	0.003;0.049;0.108
$N(0, \mathbf{V}_6^{(2)})$	0.011;0.061;0.117	0.008;0.048;0.104	0.008;0.051;0.110
$N(0, \mathbf{V}_6^{(3)})$	0.012;0.068;0.126	0.011;0.054;0.125	0.010;0.059;0.122
$N(0, \mathbf{V}_6^{(4)})$	0.372;0.645;0.779	0.397;0.742;0.866	0.431;0.753;0.872
$N(0, \mathbf{V}_6^{(5)})$	0.479;0.676;0.810	0.412;0.750;0.881	0.447;0.783;0.894
$N(0, \mathbf{V}_6^{(6)})$	0.528;0.818;0.871	0.564;0.841;0.908	0.536;0.815;0.917
$N(0, \mathbf{V}_6^{(7)})$	0.038;0.156;0.264	0.030;0.128;0.236	0.020;0.118;0.218
$N(0, \mathbf{V}_6^{(8)})$	0.892;0.986;0.992	0.970;0.994;1.000	0.974;0.994;1.000
$N(0, \mathbf{V}_6^{(9)})$	0.530;0.828;0.912	0.606;0.865;0.956	0.581;0.850;0.932
$N(0, \mathbf{V}_6^{(10)})$	0.398;0.671;0.794	0.341;0.670;0.828	0.462;0.731;0.826

TABLE 2.4 **Proportion of rejecting** H_0 ($dir = 7$, $n = 20$)

distribution	KS_n	CV_n	CVA_n
	1% 5% 10%	1% 5% 10%	1% 5% 10%
$N(0, \mathbf{V}_6^{(1)})$	0.010;0.057;0.111	0.006;0.044;0.098	0.006;0.052;0.106
$N(0, \mathbf{V}_6^{(2)})$	0.010;0.058;0.113	0.008;0.046;0.100	0.007;0.053;0.108
$N(0, \mathbf{V}_6^{(3)})$	0.014;0.073;0.145	0.010;0.053;0.123	0.006;0.061;0.131
$N(0, \mathbf{V}_6^{(4)})$	0.234;0.545;0.667	0.223;0.546;0.701	0.196;0.497;0.647
$N(0, \mathbf{V}_6^{(5)})$	0.338;0.655;0.768	0.307;0.664;0.812	0.400;0.783;0.892
$N(0, \mathbf{V}_6^{(6)})$	0.359;0.682;0.793	0.315;0.672;0.816	0.477;0.823;0.902
$N(0, \mathbf{V}_6^{(7)})$	0.030;0.198;0.323	0.016;0.115;0.199	0.018;0.128;0.214
$N(0, \mathbf{V}_6^{(8)})$	0.758;0.938;0.982	0.778;0.966;0.982	0.806;0.989;0.998
$N(0, \mathbf{V}_6^{(9)})$	0.462;0.753;0.854	0.334;0.708;0.847	0.461;0.797;0.898
$N(0, \mathbf{V}_6^{(10)})$	0.268;0.541;0.682	0.216;0.436;0.611	0.311;0.624;0.724

In view of the Tables 2.1, 2.3, and 2.4, we can find that the performances of the tests proposed are encouraging. At first, the tests hold their level very well. For the cases of $\mathbf{V}_6^{(3)}$, the powers are poor because two vectors, $\{z_1, z_2, z_3\}$ and $\{z_4, z_5, z_6\}$ have very weak correlation. As the correlation between two sets of variables increases according as $\mathbf{V}_6^{(4)}$ through $\mathbf{V}_6^{(6)}$, we enjoy higher power. Another finding is that, a little bit surprise, the performance of KS_n is the worst among three tests. In intuition, it would be the most powerful. On the contrary, CVA_n seems to be the most recommendable test among the tests here. For the case of $\mathbf{V}_6^{(7)}$, the situation is similar to the weak correlation cases. For $\mathbf{V}_6^{(8)}$, the power is good. On the other hand, the tests have, in natural, higher power with larger size of sample. However it will involve heavy computational work-load. The cases where the covariance matrices are $\mathbf{V}_6^{(9)}$ and $\mathbf{V}_6^{(10)}$, concern with a real-life example in which the new tests are employed. The data set pertains to the derivation of standardization of dress of Chinese men in 1976, and is taken from Fang, Yuan and Bentler (1992). There are 12 measurements of the body, including x_1: above the waist, x_2: under the waist, x_3: height, y_1:arm length, y_2: the front waist length, y_3: the back waist length. We now want to know whether or not $\mathbf{z}_1 = (x_1, x_2, x_3)$ and $\mathbf{z}_2 = (y_1, y_2, y_3)$ are independent. Applying the new tests to this example, the conclusion "rejected" is obtained. Similarly, put x_4: bust, x_5: waist, x_6: buttocks, y_4: shoulder length, y_5: the front part of chest, y_6: the back part of chest, and $\mathbf{z}_3 = (x_4, x_5, x_6)$, $\mathbf{z}_4 = (y_4, y_5, y_6)$. All three bootstrap tests reject the null hypothesis which is that $\mathbf{z}_3$ and $\mathbf{z}_4$ are independent. Clearly, the above mentioned two results correspond to objective reality.

5. Application: Test of Elliptical symmetry of Distribution

5.1. Introduction

It is well known that the class of elliptically symmetric distributions(elliptical distribution for short) has played important roles in statistical theory and applications. Many kinds of distributions belong to this class. For example, multivariate normal, t-distributions. A random vector $\boldsymbol{x}$ has a p-dimensional elliptical distribution if there exist a non-singular matrix $\boldsymbol{\Sigma}$ and a constant vector $\boldsymbol{\mu}$ such that

$$\boldsymbol{\Gamma}\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{x}-\boldsymbol{\mu}) \stackrel{\mathrm{d}}{=} \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{x}-\boldsymbol{\mu}) \tag{5.1}$$

for every orthogonal matrix $\boldsymbol{\Gamma}$, where the notation "$\stackrel{\mathrm{d}}{=}$" means that two sides of the equality have the same distribution. We call $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$ the shape matrix and location vector of distribution respectively. In order to describe conveniently the following problem, we denote

$$\mathcal{F}_0 = \{P: \quad \text{(5.1) holds for both known shape matrix and location vector}\};$$

Without loss of generality, we consider $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$, in $\mathcal{F}_0$, are identical matrix and zero vector respectively. Thus given a high-dimensional distribution $P(\cdot)$, one is often required to test the following hypothesis:

$$\boldsymbol{H}_0 : P(\,\cdot\,) \in \mathcal{F}_0 \quad \text{versus} \quad \boldsymbol{H}_1 : P(\,\cdot\,) \notin \mathcal{F}_0, \tag{5.2}$$

Throughout this paper, the boldface symbols will represent vector or matrix if no special mentioning. In the following the notation "$|\cdot|$" stands for the Euclidean norm. Naturally if x is real-valued, $|x|$ is the absolute value of x. In order to present in the next section the proposed tests we now introduce two properties of a spherical distribution which will be the basis for constructing tests.

LEMMA 5.1. *$\boldsymbol{x}$ has a p-dimensional spherical distribution if and only if*

$|\boldsymbol{x}|$ and $\boldsymbol{x}|\boldsymbol{x}|^{-1}$ are independent and $\boldsymbol{x}|\boldsymbol{x}|^{-1}$ is uniformly distributed on $S^{p-1} = \{\boldsymbol{a} : \boldsymbol{a} \in R^p, |\boldsymbol{a}| = 1\}$.

For convenience, we present a necessary and sufficient condition for Lemma 5.1.

LEMMA 5.2. *That both* $|\boldsymbol{x}|$ *and* $\boldsymbol{x}|\boldsymbol{x}|^{-1}$ *are independent each other, and* $\boldsymbol{x}|\boldsymbol{x}|^{-1}$ *is uniformly distributed on* S^{p-1} *if and only if for every* $\boldsymbol{a} \in S^{p-1}$, $|\boldsymbol{x}|$ *and* $\boldsymbol{a}^\tau \boldsymbol{x}|\boldsymbol{x}|^{-1}$ *are independent, and* $\boldsymbol{a}^\tau \boldsymbol{x}|\boldsymbol{x}|^{-1}$*'s have the same distribution* $A(\,\cdot\,)$ *with the density function*

$$f(y) = \frac{\Gamma(p/2)}{\Gamma((p-1)/2)\pi^{1/2}}(1-y^2)^{\frac{(p-3)}{2}}, \qquad -1 \le y \le 1. \tag{5.3}$$

Lemma 5.2 is a direct consequence of the well known result (e.g. see Watson (1983) or Fang, Kotz and Ng, 1990, Theorem 2.2.5).

5.2. The Construction of the Statistics

Suppose that we have collected the i.i.d. p-dimensional sample $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$. Denote by P_n the empirical measure based on $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$. For convenience $P_n f$ will stand for the expectation value $\int f\, dP_n$ for function f on R^p, and $I(C)$ will mean the indicator function of set C.

Consider PP-sample, $\boldsymbol{a}^\tau \boldsymbol{x}_1, \cdots, \boldsymbol{a}^\tau \boldsymbol{x}_n$, where $\boldsymbol{a} \in S^{p-1}$. By making the use of Lemma 5.2, one can apply the following statistics for (5.2).

$$\begin{aligned} \widetilde{K}_n &= \sup_{a \in S^{p-1}} \sup_{t,s} \sqrt{n}|\widetilde{K}_n(\boldsymbol{a}, t, s)| \\ &= \sup_{\boldsymbol{a},t,s} \sqrt{n}\big|P_n I(|\boldsymbol{x}| \le t) I(\boldsymbol{a}^\tau \boldsymbol{x}|\boldsymbol{x}|^{-1} \le s) - P_n I(|\boldsymbol{x}| \le t)A(s)\big|, \end{aligned} \tag{5.4}$$

where $A(\cdot)$ is the distribution defined in (5.3), and

$$\widetilde{V}_n = \int \widetilde{K}_n^2(\boldsymbol{a}, t, s) dH_n(t) dA(s) d\mu(\boldsymbol{a}), \tag{5.5}$$

where $H_n(t)$ is the empirical distribution based on $\{|\boldsymbol{x}_1|, \cdots, |\boldsymbol{x}_n|\}$.

5.3. Asymptotics of Statistics

We in this section investigate the asymptotic behavior of statistics which includes all of the Kolmogorov type statistics and Cramer-von Mises type statistics in subsection 5.2. We do not consider the other statistics because we meet, unfortunately, some technical difficulties in mathematics. These should be considered in further research.

For considering the asymptotics of $\widetilde{K}_n$ and $\widetilde{V}_n$, let $\widetilde{\boldsymbol{W}} = \{\boldsymbol{W}_p(\boldsymbol{a}, t, s) : (\boldsymbol{a}, t, s) \in S^{p-1} \times R^2\}$ be a $\mathcal{L}^2(P)$-norm continuous, centered Gaussian process with covariance function

$$\begin{aligned}\widetilde{R}((\boldsymbol{a},t,s),(\boldsymbol{b},t_1,s_1)) =& \boldsymbol{E}\{\widetilde{W}_p(\boldsymbol{a},t,s)\widetilde{W}_p(\boldsymbol{b},t_1,s_1)\}\\ =& PI(|\boldsymbol{x}|\le t)I(|\boldsymbol{x}|\le t_1)I(\boldsymbol{a}^\tau \boldsymbol{x}|\boldsymbol{x}|^{-1}\le s)I(\boldsymbol{b}^\tau \boldsymbol{x}|\boldsymbol{x}|^{-1}\le s_1)\\ &-HI(|\boldsymbol{x}|\le t)HI(|\boldsymbol{x}|\le t_1)PI(\boldsymbol{a}^\tau \boldsymbol{x}|\boldsymbol{x}|^{-1}\le s)PI(\boldsymbol{b}^\tau \boldsymbol{x}|\boldsymbol{x}|^{-1}\le s_1),\end{aligned} \tag{5.6}$$

and sample paths being uniformly bounded, where $H(\cdot)$ is the distribution of $|\boldsymbol{x}|$. Denote by B the Brownian bridge.

THEOREM 5.1. *Suppose that the distribution of $|\boldsymbol{x}|$ is continuous. Then under $\boldsymbol{H}_0$ in (5.2)*

$$\widetilde{K}_n \Longrightarrow \widetilde{\boldsymbol{K}} = \sup_{b,t,s}|\widetilde{W}_p(\boldsymbol{b},t,s) - B(t)A(s)|, \tag{5.7}$$

$$\widetilde{\boldsymbol{V}}_n \Longrightarrow \widetilde{\boldsymbol{V}} = \iiint(\widetilde{W}_p(b,t,s) - B(t)A(s))^2 dH(t)dA(s)d\mu(\boldsymbol{b}). \tag{5.8}$$

5.4. Bootstrap Approximation

According to those results presented in subsection 5.3, we see that the distributions and the asymptotic distributions of the statistics depend on the unknown distribution P, which is the population distribution of sample, and are not tractable. Hence bootstrap approximations of the statistics are available for choosing critical values of the tests. Since we are here testing for the spherical symmetry of a distribution, a procedure for resampling data is suggested below. Let $\boldsymbol{U}_1,\cdots,\boldsymbol{U}_n$ come from the uniform distribution on the unit sphere which can be generated by computer and $\boldsymbol{y}_i = |\boldsymbol{x}_i|\boldsymbol{u}_i$. Let G_n be the empirical distribution based on $\boldsymbol{y}_i$'s and $\boldsymbol{y}_i{}^*$ come from G_n as the bootstrapping sample. Denote P_n^* empirical measure based on $\{\boldsymbol{y}_1^*,\cdots,\boldsymbol{y}_n^*\}$. The bootstrap approximations of the proposed statistics are as follows.

$$\begin{aligned}\widetilde{K}_n^*(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n) =& \sup_{a,t,s}\sqrt{n}|\widetilde{K}_n^*(\boldsymbol{a},t,s,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n)|\\ =& \sup_{a,t,s}\sqrt{n}|P_n^*I(|\boldsymbol{y}^*|\le t)I(\boldsymbol{a}^\tau\boldsymbol{y}^*|\boldsymbol{y}^*|\le s) - G_nI(|\boldsymbol{y}|\le t)\\ &\cdot I(\boldsymbol{a}^\tau\boldsymbol{U}\le s) - (P_n^*I(|\boldsymbol{y}^*|\le t)A(s) - G_nI(|\boldsymbol{y}|\le t)A(s))|.\\ \widetilde{V}_n^*(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n) =& n\iiint(\widetilde{K}_n^*(\boldsymbol{a},t,s,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n))^2 dG_n(t)dA(s)d\mu(\boldsymbol{a}).\end{aligned}$$

In practical implementation, for each set of $\{\boldsymbol{U}_1,\cdots,\boldsymbol{U}_n\}$ generated by computer, generate B groups of $\{\boldsymbol{y}_1^*,\cdots,\boldsymbol{y}_n^*\}$. Calculate, respectively, B values of $\widetilde{K}_n^*$ and $\widetilde{V}_n^*$, and then get the corresponding $(1-\alpha)$-quantile values, $\widetilde{K}_\alpha^*$ and $\widetilde{V}_\alpha^*$ say. Repeat this procedure c times to get $\{\widetilde{K}_{\alpha 1}^*, \cdots, \widetilde{K}_{\alpha c}^*\}$ and $\{\widetilde{V}_{\alpha 1}^*, \cdots, \widetilde{V}_{\alpha c}^*\}$. Finally, use the sample means $M\widetilde{K}_\alpha^* = \frac{1}{c}\sum_{j=1}^c \widetilde{K}_{\alpha j}^*$ and $M\widetilde{V}_\alpha^* = \frac{1}{c}\sum_{j=1}^c \widetilde{V}_{\alpha j}^*$ as the critical values. By making the use of result of Gine and Zinn (1990), we can easily derive the following result.

THEOREM 5.2. *Under the same conditions imposed in Theorem 5.1* $\widetilde{K}_n^*(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)$ *and* $\widetilde{V}_n^*(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)$ *have, in the almost sure sense, the same asymptotic distributions as those of* $\widetilde{K}_n$ *and* $\widetilde{V}_n$ *respectively.*

5.5 Simulations

Further insight into the applicability of four statistics, we conduct some simulations using the sample from 3-dimensional distributions. The samples are generated by Monte Carlo method. The l_n points $\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_{l_n}\}$ on S^2 are chosen by Number-theoretic method. We here, choose the sample size $n = 100$ and $l_n = 21$. The simulation experiments are performed for the tests $\widetilde{K}_n$, and $\widetilde{V}_n$ concerning the hypothesis (5.2). The critical values of tests are obtained by the bootstrap approximation. For each case, 500 replications of bootstrap samples are independently generated for determining the critical values. Furthermore, the basic experiment was replicated 1000 times for each case. The proportion of the statistic values exceeded the 95th percentile of the bootstrap statistic values in all cases studied here were recorded.

Since computation is intensive, we here only conduct simulation experiments concerning 3-dimensional cases. It is expected that the proposed tests should be powerful for higher dimensions and for more alternatives.

The simulated results are listed in the tables below. The data sets are generated from the following different distributions, where

1. N_l: the standard normal distribution $N(\mathbf{0}, \boldsymbol{I}_l)$.

2. T_3: the multivariate t-distribution in R^3 with the density function

$$c_3\left(1 + \frac{\boldsymbol{x}^T\boldsymbol{x}}{10}\right)^{-(10+3/2)}$$

where c_3 is the normalizing constant.

3. E_l: $l-$dimensional distribution with iid marginal distributions each having the standard exponential distribution.

4. χ_l^2: l−dimensional distribution with iid marginal distributions each having the chi-square distribution with degrees two of freedom.

5. U_3: the uniform distribution on the 3-dimensional unit sphere surface.

6. B_3: the beta distribution with parameters 3 and 1 respectively.

Furthermore, FG means that this distribution has two independent marginal distributions F and G For example, $\chi_2^2\ B_1$ means that the distribution has the independent marginal distributions χ_2^2 and B_1 respectively.

For the hypothesis (5.2), we consider that the location vectors $\boldsymbol{\mu}$ and the shape matrices $\boldsymbol{A}$ are known in the investigated distributions below. Hence we can make a transformation to get zero location and identical shape matrix. The simulation results are presented in the table 5.1.

TABLE 5.1 **Proportion of Rejecting the Null Hypothesis (5.2)**

Statistic	N_3	T_3	U_3	$\chi_2^2 B_1$	χ_3^2	E_3	B_3
$\tilde{K}_n$	0.047	0.035	0.038	1.000	1.000	1.000	1.000
$\tilde{V}_n$	0.045	0.032	0.038	1.000	1.000	1.000	1.000

In view of the tables, we can see that the proportion of rejecting the null are encouraging for the investigated cases here. On the other hand, the Cramer-von Mises type tests have the better power than the Kolmogorov type tests. On the other hand, in prectical use, there is a problem how many projection directions are chosen. This deserves further study.

Acknowledgements

The authors thank Professor Taguri, M. and Dr. Konno, Y., Chiba University, for their many valuable suggestions and critical comments.

References

[1] Beran, R. and Miller, P.W. (1986). Convergence sets for a multivariate distribution. *Ann. Statist.*, **14(2)** 431–443.

[2] Blum, J.R., Kiefer, J. and Rosenblatt, M. (1961). Distribution free tests of independence based on sample distribution function. *Ann. Math. Statist.*, **32(2)** 485–498.

[3] Dudley, R.M., and Philipp, W. (1983). Invariance principle for sums of Banach space valued random elements and empirical processes. *Zeit. Wahrscheinl.-theorie*, **62** 509–552.

[4] Efron, B. (1979). Bootstrap methods :Another look at the Jackknife. *Ann. Statist.*, **7(1)** 1–26.

[5] Fang, K.T. Kotz, S. and Ng, K.W. (1990). *Symmetric Multivariate and Related Distribution* ,Chapman and Hall London, Monographs on statistics and Applied Probability, No. 36.

[6] Fang, K.T., Yuan, K.H. and Bentler, P.M. (1992). Applications of sets of points uniformly distributed on a sphere to testing multinormality and robust estimation. *in Probability and Statistics*, (Jiang, Z.P., Yan, S.J., Cheng, P. and Wu, R., Eds.). World Scientific, 56–73.

[7] Fang, K.T. and Wang, Y. (1993). *Number-theoretic Methods in Statistics* ,Chapman and Hall London, Monographs on statistics and Applied Probability, No. 51.

[8] Fang, K.T., Zhu, L.X. and Bentler, P.M. (1993). A necessary test for sphericity of high-dimensional distribution. *J. Multivariate Anal.*, **45(1)** 34–55.

[9] Giné, E. and Zinn, J. (1986). Lectures on the central limit theorem for empirical processes. *Probability and Banach Spaces. Lecture Notes in Math.*, **1221** Springer, Berlin, 50–113.

[10] Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures. *Ann. Prob.*, **18** 851–869.

[11] Giné, E. and Zinn, J. (1991). Gaussian characterization of uniform Donsker classes of functions. *Ann. Prob.*, **19(2)** 758–782.

[12] Hua, L.G. and Wang, Y. (1981). *Applications of Number Theory to Numerical Analysis*, Springer-Verlag and Science Press, Berlin and Beijing.

[13] Huber, P.J. (1985). Projection Pursuit (with discussion). *Ann. Statist.*, **13(2)** 435–525.

[14] Jing, P and Zhu, L.X., (1996). Some Blum-Kiefer-Rosenblatt Type Tests for the Joint Independence of Variables. *Communications in Statistics, Theorem and Methods.* , **25(9)** 2127–2139.

[15] Jing, P and Zhu, L.X., (1996). On Some Tests Based Projection Pursuit for Elliptical Symmetry of a High-dimensional Distribution (submitted to Chinese Science Bulletin).