Chin. Ann. of Math.
13B: 3(1992), 280—288.

ON PP TYPE ESTIMATED KOLMOGOROV STATISTICS

ZHU LIXING (朱力行)* CHENG PING (成平)**

Abstract

The authors construct PP type estimated Kolmogorov statistics by Projection Pursuit technique to deal with simultaneously the sparseness of sample points in high-dimensional space and the case that distribution occupies the unknown location parameter and dispersion matrix. Furthermore, the tali behavior of the limit of the statistics is investigated.

§ 1. Introduction

Suppose that $X_1, \cdots, X_n$ are i. i. d. sample with d-dimensional probability measure $Q(\cdot, \theta)$ whose parameter θ lies in a convex, open set in R^k. In order to test $Q(\cdot, \theta)=P(\cdot, \theta)$ whose function form is known, we can use the es timated Kolmogorov statistics

$$\sup_{x\in R^d} \sqrt{n}\,|P_nI(X\leqslant x)-P(x, \hat{\theta}_n)|, \tag{1.1}$$

where P_n and $\hat{\theta}_n$ are, repsectively, basing on $X_1, \cdots, X_n$, the empirical measure and the estimate of θ. But, the sparseness of data points in high-dimensional space is an obstacle for using effectively this sort of statistics. To avoid it, one can construct an estimated projection pursuit type Kolmogorov statistic

$$D_n=\sup_{t\in R^1}\sup_{a\in S_d}\sqrt{n}\,|P_nI(a^TX\leqslant t)-P_{\theta_n}I(a^TX\leqslant t)|, \tag{1.2}$$

where $S_d=\{a\in R^d, \|a\|=1\}$, $\|\cdot\|$ is the Euclidean norm, $I(A)$ is the indicator function of set A, P_nf and $P_{\theta_n}f$ are defined as, respectively, $\int f(x)dp_n(x)$ and $\int f(x)dp(x, \theta_n)$.

Similarly, for the case that dispersion matrix is unknown, that is, for distribution $P(\cdot, \Sigma)$, we can construct

$$D_n'=\sup_{a,t}\sqrt{n}\,|P_nI(a^TX\leqslant t)-P_{\hat{\Sigma}_n}I(a^TX\leqslant t)|. \tag{1.3}$$

More generally, consider $P(\cdot, \theta, \Sigma)$, where location parameter θ and dispersion matrix Σ are unknown. We construct the estimate

Manuscript received September 4, 1990, Revised April 4, 1991,
* Institute of Applied Mathematics, Academia Sinica, Beijing 100080, China.
** Institute of Sys tem Sciences, Academia Sinica, Beijing 100080, China,

$$D_n'' = \sup_{a,t} \sqrt{n}\, | P_n I(a^T X \leqslant t) - P_{\theta_n, \Sigma_n} I(a^T X \leqslant t) |. \tag{1.4}$$

It is well knewn that many useful properties of these statistics for one-dimensional case have been derived (cf. Sharock, Chapter V). Unfortunately, we have little knowledge in high-dimensional cases. In this paper, we investigate the tail behavior concerning the limit of statistics if the limit exists.

The paper is organized as follows: Section 2 provides some illustrations for the convergence of (1.2) in distribution, the case of location parameter is put in Section 3, Section 4 contains the case of dispersion matrix and some further discussion.

§ 2. The Convergence of the Statistic in Distribution

Does the statistics in (1.2) converge as n tends to infinite? What is the limit if the statistic converges? These are the problems we counter before studying the tail behavior of the limit statistics. So we first discuss these two problems.

According to the statistical usefulness, good estimates can often be regular estimates, that is,

$$\theta_n = \theta + n^{-1} \sum_{j=1}^{n} L(X_j) + o_p(n^{-\frac{1}{2}}), \tag{2.1}$$

for some k-dimensional function L satisfying $PL(X) = 0$ and $V = PL(X)L(X)^T$ is a positive definife matrix.

On the other hand, we provide P is uniformly differentiable to garantee the convergence of D_n as follows:

$$\sup_{a \in S_d} \sup_{t \in R^1} \| P_{\theta'} I(a^T X \leqslant t) - P_\theta I(a^T X \leqslant t) - (\theta' - \theta)^T A(a, t) \|_P$$
$$= o(\|\theta' - \theta\|) \text{ near } \theta \tag{2.2}$$

for some fixed k-dimensional function $A(\cdot, \cdot)$ with all its components in $C[S_d \times R^1]$ of all continuous functions, where the notation "$\|\cdot\|_p$" stands for $L^2(P)$-seminorm.

Thus D_n could be written as

$$D_n = n^{\frac{1}{2}} \sup_{a \in S_d} \sup_{t \in R^1} | P_n I(a^T X \leqslant t) - P_\theta I(a^T X \leqslant t) - P_n L^T(X) A(a, t) | + o_p(1)$$
$$= n^{\frac{1}{2}} \sup_{a \in S_d} \sup_{t \in R^1} | w_n(a, t) - P_n L^T(X_j) A(a, t) | + o_p(1). \tag{2.3}$$

Applying multivariate central limit theorem, for any finite subset of $S_d \times R^1$, $\{(a_i, t_i),\ i = 1, 2, \cdots, k\}$ say,

$$\{n^{\frac{1}{2}} w_n(a_i, t_i),\ P_n L^T(X) A(a_i, t_i):\ i = 1, 2, \cdots, k\}$$
$$\to \{w(a_i, t_i),\ Z^T A(a_i, t_i):\ i = 1, 2, \cdots, k\},$$

whers the notation "$\to$" is defined as convergence in distribution, w is a Gaussian process, Z is distributed with $N(0, v)$. Employing the similar argument used by Pollard ([2], p. 157). we can show that

$$D_n \to \sup_{a\in S_d}\sup_{t\in R^1} |w(a, t) - Z^T A(a, t)| = D. \tag{2.4}$$

Moreover, we know that w is a zero mean Gaussian process where sample paths continue uniformly and boundedly with respect to $L^2(P(\cdot, \theta))$-seminorm, and covariance function is

$$R((a, t), (a_1, t_1)) = P_\theta I(a^T X \leqslant t) I(a_1^T X \leqslant t_1) - P_\theta I(a^T X \leqslant t) P_\theta I(a_1^T X \leqslant t_1) \tag{2.5}$$

and the covariance function of $w - Z^T A$ is

$$R'((a, t), (a_1, t_1)) = P_\theta(I(a^T X \leqslant t) - L^T(X) A(a, t))(I(a_1^T X \leqslant t_1) - L^T(X) A(a_1, t_1)) - P_\theta(I(a^T X \leqslant t) - L^T(X) A(a, t)) P_\theta(I(a_1^T X \leqslant t) - L^T(X) A(a_1, t_1)).$$

In the following sections, we investigate the tail behavior of supremum of $w - Z^T A$.

Without confusion, denote by c a constant throughout the following sections.

§ 3. Location Parameter is Unknown

Suppose $Q(\cdot, \theta) = F(\cdot - \theta)$, where F is spherically symmetric, and θ is a d-dimensional parameter belonging to an open, convex set Ω in R^d. We establish a tail probability bound for limit random variable of (1.2), $\sup |w - Z^T A|$.

Let us now see what is that of limiting (1.2) in terms of the discussion in Section 2. At first, the spherical symmetry of F implies that $a^T X - a^T \theta$ is distributed with $F_1(t)$, where F_1 is the marginal distribution of F not depending on projection direction a, and $F_1(t)$ must have the bounded density function f. We then have via taking $\theta_n = \frac{1}{n}\sum_{j=1}^{n} X_j$,

$$D_n \to D = \sup_{a,t} |w(a, t) - f(t - a^T\theta) a^T Z|. \tag{3.1}$$

where, according to (2.1), $L(X) = X - \theta$ and $V = \mathrm{cov}(X)$, the covariance matrix of X.

Now present the following result.

Theorem 3.1. *Assume density function $f(\cdot)$ satisfies*

i) *f is continuous uniformly:* (3.2)

ii) *there is a constant B such that for any $1 > \varepsilon > 0$ and any pair (t, t_1),*

$$|f(F_1^{-1}(t)) - f(F_1^{-1}(t_1))| \leqslant B|t - t_1|. \tag{3.3}$$

Then for $\lambda > 1$

$$P\{D > \lambda\} \leqslant c\lambda^{2d-1} \exp\left(-\frac{\lambda^2}{2\sigma^2}\right) \tag{3.4}$$

holds for some constant c, where

$$\sigma^2 = \sup_{a,t} P\{w(a, t) - f(t - a^T\theta)\ a^T Z\}^2.$$

The proof the theorem is devided into some lemmas.

Lemma 3. 2. (cf. Lemma 2.2, [3]). *For each $\gamma (0<\gamma<1)$, there exists a partition of S_d, $\{A_i(\gamma)\}$ say, such that*

$$\begin{cases} A_i^0(\gamma) \cap A_j^0(\gamma) = \phi, \\ \text{card } \{A_i(\gamma)\} \leqslant C\gamma^{-(d-1)}, \end{cases} \tag{3.5}$$

where for some $a_i \in S_d$

$$A_i(\gamma) = \{a \in S_d: \max_{1\leqslant j\leqslant d} |a^{(j)} - a_i^{(j)}| \leqslant \gamma\}, \tag{3.6}$$

$A_i^0(\gamma)$ is the set consisting of all the inner points in $A_i(\gamma)$.

Noticing that $A_i(\gamma)$ has anologous structure with S_d, we easily obtain the following assertion with tracing the argument of Lemma 2.2 in [3].

Lemma 3.3. *For each $A_i(\gamma)$ and $\varepsilon>0$, there exists a partition of $A_i(\gamma)$, $\{B_l(\gamma\varepsilon)\}$ say, for which*

$$\begin{cases} B_l^0(\gamma\varepsilon) \cap B_k^0(\gamma\varepsilon) = \phi, \ \forall B_l^0(\gamma\varepsilon), B_k^0(\gamma\varepsilon) \in \{B_l^0(\gamma\varepsilon)\}, \\ \text{card } \{B_l(\gamma\varepsilon)\} \leqslant C\varepsilon^{-(d-1)}, \end{cases} \tag{3.7}$$

where

$$B_l(\gamma\varepsilon) = \{a \in A_i(\gamma): \max_{1\leqslant j\leqslant d} |a^{(j)} - a_l^{(j)}| \leqslant \gamma\varepsilon\}$$

for some $a_l \in A_i(\gamma)$, the definition of $B_l^0(\gamma\varepsilon)$ is similar to that of $A_i^0(\gamma)$.

Lemma 3. 4. *For any fixed $a \in S_d$ and fixed $\gamma_1>0$, there exists a sequence*

$$\left\{ t_i^a: i = -\left[\frac{1}{2\gamma_1}\right]+1, \cdots, \left[\frac{1}{2\gamma_1}\right]+1 \right\} \subset R^1$$

for which

$$\begin{cases} P_\theta\{a^T X \leqslant t_0^a\} = \dfrac{1}{2}, \\ P_\theta\{a^T X \leqslant t_{i+1}^a\} - P_\theta\{a^T X \leqslant t_i^a\} \leqslant \gamma_1, \\ P_\theta\{a^T X \leqslant t^a_{-[\frac{1}{2\gamma_1}]+1}\} \leqslant \gamma_1, \\ P_\theta\{a^T X \geqslant t^a_{[\frac{1}{2\gamma_1}]+1}\} \leqslant \gamma_1, \end{cases} \tag{3.8}$$

where $P_\theta\{a^T X \leqslant t\}$ is the marginal distribution of $F(\cdot - \theta)$ at direction a.

Proof Due to the spherical symmetry of F and the continuity of F_1, the conclusion of Lemma 3.4 is an obvious fact, and t_i^a is of the form $t_i - a^T\theta$, where t_i does not depend on θ and a.

Let

$$A_{i,j}(\gamma, \gamma_1) = \{(a, t): a \in A_i(\gamma), t_{2(j-1)} \leqslant t < t_{2j}^a\}, j$$

$$j = -\left[\frac{1}{4\gamma_1}\right]+1, \cdots, \left[\frac{1}{4\gamma_1}\right]+1,$$

and

$$A_{i,-[\frac{1}{4\gamma_1}]-2}(\gamma, \gamma_1) = \{(a, t): a \in A_i(\gamma), t \leqslant t^a_{-[\frac{1}{2\gamma_1}]+1}\},$$

$$A_{i,[\frac{1}{4\gamma_1}]+2}(\gamma, \gamma_1) = \{(a, t): a \in A_i(\gamma) \geqslant t^a_{[\frac{1}{2\gamma_1}]+1}\}.$$

Lemma 3.2 and Lemma 3.4 imply that $\{A_{ij}(\gamma, \gamma_1)\}$ is a partion of $S_d \times R^1$ and that

Lemma 3. 5.

$$\operatorname{card}\{A_{ij}(\gamma,\gamma_1)\leqslant c\gamma^{-(d-1)}\gamma_1^{-1}. \tag{3.9}$$

Lemma 3.6. *For any pair $(a, a_1)\subset S_d$,*

$$\sup_t P_\theta\{\{a^T(X-\theta)\leqslant t\}\Delta\{a_1^T(X-\theta)\leqslant t\}\}\leqslant\|a-a_1\|, \tag{3.10}$$

where the notation "Δ" denotes the symmetric difference of two sets.

Proof For any orthogonal matrix Q,

$$\|a-a_1\|=\|(a-a_1)^TQ\|. \tag{3.11}$$

We can find an orthogonal matrix Q for which

$$(a-a_1)^TQ=(a^0-a_1^0)^T,$$

where $a^0=(1, 0, \cdots, 0)$ and $a_1^0=(\cos\phi, \sin\phi, 0, \cdots, 0)$, $0\leqslant\phi\leqslant\pi$. So

$$\begin{aligned}\|a-a_1\|&=\|a^0-a_1^0\|=((1-\cos\phi)^2+(\sin\phi)^2)^{\frac{1}{2}}\\&=2\sin\left(\frac{\phi}{2}\right)\geqslant\frac{\phi}{\pi}.\end{aligned} \tag{3.12}$$

On the other hand, Lemma 2.6 in [3] and the spherical symmetry yield

$$\begin{aligned}&\sup_t P_\theta\{\{a^T(X-\theta)\leqslant t\}\Delta\{a^T(X-\theta)\leqslant t\}\}\\&\leqslant P_\theta\{\{a^T(X-\theta)\leqslant 0\}\Delta\{a_1^T(X-\theta)\leqslant 0\}\}\\&=P_\theta\{\{(a^0)^T(X-\theta)\leqslant 0\}\Delta\{(a_1^0)^T(X-\theta)\leqslant 0\}\}\\&=2P_\theta\{\{(a_1^0)^T(X-\theta)\leqslant 0\}\Delta\{(a^0)^T(X-\theta)\leqslant 0\}\}\\&=\frac{\phi}{\pi}.\end{aligned} \tag{3.13}$$

Consequently, it together with (3.12), yields (3.10).

Lemma 3.7. *Under the conditions* (3.1) *and* (3.2), *there is a constant* c *such that*

$$(P\{w(a, t)-f(t-a\theta)a^TZ-(w(a_i, t_{2j-1}^{a_i})-f(t_{2j-1}^{a_i}-a_i^T\theta)a_i^TZ)\}^2)^{\frac{1}{2}}\leqslant\beta \tag{3.14}$$

holds for any pair $(a, t)\in A_{ij}(c\beta^2, c\beta^2)$, where we take $\gamma=c\beta^2$ and $\gamma_1=c\beta^2$.

Proof The left side of (3.14) is less than or equal to

$$\begin{aligned}&2\{(P\{w(a, t)-w(a_i, t_{2j-1}^{a_i})\}^2)^{\frac{1}{2}}+P\{f(t-a^T\theta)a^TX-f(t_{2j-1}^{a_i}-a_i^T\theta)a_i^TX\}^2)^{\frac{1}{2}}\}\\&\triangleq 2(I_1+I_2).\end{aligned} \tag{3.15}$$

By (2.5) we have

$$\begin{aligned}I_1^2&=P_\theta\{a^TX\leqslant t\}(1-P_\theta\{a^TX\leqslant t\})+P_\theta\{a_i^TX\leqslant t_{2j-1}^{a_i}\}(1-P_\theta\{a_i^TX\leqslant t_{2j-1}^{a_i}\})\\&\quad-2P_\theta\{\{a^TX\leqslant t\}\cap\{a_i^TX\leqslant t_{2j-1}^{a_i}\}\}-P_\theta\{a^TX\leqslant t\}P_\theta\{a_i^TX\leqslant t_{2j-1}^{a_i}\}\\&=P_\theta\{a^TX\leqslant t\}\Delta\{a_i^TX\leqslant t_{2j-1}^{a_i}\}\}-(P_\theta\{a^TX\leqslant t\}-P_\theta\{a_i^TX\leqslant t_{2j-1}^{a_i}\})^2\\&\leqslant P_\theta\{\{a^TX\leqslant t\}\Delta\{a_i^TX\leqslant t_{2j-1}^{a_i}\}\}\\&\leqslant P_\theta\{\{a^TX\leqslant t\}\Delta\{a_i^TX\leqslant t\}\}+(P_\theta\{a_i^TX\leqslant t_{2j}^{a_i}\}-P_\theta\{a_i^TX\leqslant t_{2j-1}^{a_i}\})\\&=P_\theta\{\{a^T(X-\theta)\leqslant t-a^T\theta\}\Delta\{a_i^T(X-\theta)\leqslant t-a_i^T\theta\}\\&\quad+(P_\theta\{a_i^TX\leqslant t_{2j}^{a_i}\}-P_\theta\{a_i^TX\leqslant t_{2j-1}^{a_i}\})\\&=I_3+I_4.\end{aligned} \tag{3.16}$$

Lemma 3.4 implies $I_4\leqslant 2\gamma_1$. For I_3 we can split it into two parts.

$$I_3\leqslant P_\theta\{\{a^T(X-\theta)\leqslant t-a^T\theta\}\Delta\{a_i(X-\theta)\leqslant t-a^T\theta\}\}$$

$$+|P_\theta\{a_i^T(X-\theta)\leqslant t-a^T\theta\}-P_\theta\{a_i^T(X-\theta)\leqslant t-a_i^T\theta\}|$$
$$=I_3'+I_3''. \tag{3.17}$$

Lemma 3.6 and Lemma 3.2 yield

$$I_3'\leqslant\|a-a_i\|\leqslant\gamma. \tag{3.18}$$

Futhermore, invoking the condition (3.2), we have $I_3''\leqslant C\gamma$. Then we have

$$I_1\leqslant c\gamma^{\frac{1}{2}}. \tag{3.20}$$

We now deal with I_2. It is clearly seen that

$$I_2^2=((f(t-a^T\theta)a^T-f(t_{2j-1}^{a_i}-a_i^T\theta)a_i^T)V^{\frac{1}{2}})^2. \tag{3.21}$$

Moreover

$$|(f(t-a^T\theta)a^T-f(t_{2j-1}^{a_i}-a_i^T\theta)a_i^T)V^{\frac{1}{2}}|$$
$$\leqslant|(f(t-a^T\theta)a^T-f(t_{2j-1}^{a_i}-a_i^T\theta)a^T)V^{\frac{1}{2}}|$$
$$+|(f(t_{2j-1}^{a_i}-a_i^T\theta)a^T-f(t_{2j-1}^{a_i}-a_i^T\theta)a_i^T)V^{\frac{1}{2}}|$$
$$=I_5+I_6. \tag{3.22}$$

Applying (3.2) to I_6, (3.3) and (3.13) to I_5, we have

$$I_5\leqslant c\gamma_1,\ I_6\leqslant c\gamma. \tag{3.23}$$

So $I_2\leqslant c(\gamma+\gamma_1)$, which together with (3.15) and (3.20) yields (3.14).

Similar argument leads to the following assertion.

Lemma 3.8. *Choosing $\gamma s^2=c\beta^2 s^2$, $\gamma_1 s^2=c\beta^2 s^2$, we can find apartition of $A_{ij}(c\beta^2, c\beta^2)$, $\{D_k\}$ say, for which*

$$\{D_k\}\leqslant cs^{-2d}. \tag{3.24}$$

Furthermore, there exists an $(a_k, t_k)\in A_{ij}(c\beta^2, c\beta^2)$ corresponding to D_k such that for any $(a, t)\in D_k$,

$$(P\{w(a, t)-f(t-a^T\theta)a^TX-(w(a_k, t_k)-f(t_k-a_k^T\theta)a_k^TX)\}^2)^{\frac{1}{2}}\leqslant s. \tag{3.25}$$

Proof. Using Lemma 3.3 and Lemma 3.6, tracing the arguments of Lemma 3.4 and Lemma 3.7, we can achieve (3.25).

Proof of Theorem 3.1 Lemmas 3.5, 3.7, 3.8 have checked the conditions (2.1) and (2.5) in Adler and Samorodnisky's paper ([1], p. 1340 and p. 1341]. Taking $g(\beta_\lambda)=\beta_\lambda+\frac{2\beta_\lambda}{k^2}$ and $\beta_\lambda=g^{-1}(\lambda^2(1+4d\log h)^{-\frac{1}{2}})$, we see that, from Theorem 2.1 in [1], for appropriate constant c the inequality

$$P\{\sup_{A_{ij}(c\beta^2,c\beta^2)}|w(a, t)-f(t-a^T\theta)a^TX|>\lambda\}$$
$$\leqslant c\lambda^{-1}\exp\left\{-\frac{\lambda^2}{2\sigma^2}\right\}+c\lambda^{-2}\exp\left(-\frac{1}{2}\lambda^4(1+4d\log h)\right) \tag{3.26}$$

holds for $\lambda>1$, where $\sigma^2=\sup_{a,t}P(w(a, t)-f(t-a^T\theta)a^TZ)^2$. Then from Lemma 3.5

$$P\{\sup_{a,t}|(w(a, t)-f(t-a^T\theta)a^TZ|>\lambda\}$$
$$\leqslant\sum_{i,j}P\{\sup_{A_{ij}}|w(a,t)-f(t-a^T\theta)a^TZ|>\lambda\}$$

$$\leqslant c\lambda^{2d}\lambda^{-1}\exp\left(-\frac{\lambda^2}{2\sigma^2}\right)=c\lambda^{2d-1}\exp\left(-\frac{\lambda^2}{2\sigma^2}\right). \qquad (3.27)$$

Remark 3.1. In Theorem 3.1, we impose the conditions on density function f. Luckily, these conditions are mild, and we can see that almost all spherical distributions possess properties (3.2) and (3.3).

§ 4. The Case of Unknown Dispersion Matrix

Suppose that $Q_\Sigma(\cdot)$ is an elliptically symmetric probability measure whose density function is of the form $f(x^T\Sigma x)$; we further assume that Σ is the positive definite matrix. The elliptical symmetry implies that $\Sigma^{\frac{1}{2}}X$ is distributed spherically symmetrically. When Σ is unknown, we need to estimate it so as to construct the estimate as (1.1). By the same reason as that in Section 1, we can construct projection pursuit type Kolmogorov statistics.

$$D'_n=\sup_{t\in R^1}\sup_{a\in S_a}\sqrt{n}\left|P_nI(a^TX\leqslant t)-F_1\left(\frac{t}{\sqrt{a^T\hat{\Sigma}_na}}\right)\right|, \qquad (4.1)$$

where F_1 is one-dimensional margined distribution of $\Sigma^{\frac{1}{2}}X$, $a^T\hat{\Sigma}_na=n^{-1}\sum_{j=1}^n(a^TX_j)^2$. We know that F_1 does not depend on the projective direction a, and that $a^T\ \Sigma_na$ is $\sqrt{n}$-consistent estimate for the variance of a^TX since $PX=0$.

If the density function of F_1, f say, are bounded and continuous uniformly, taking $L_a(X)=(a^TX)^2$, along with the similar argument in Section 2 as long as we pay attention to the uniform consistence of $a^T\Sigma_na$ with respect to a, we then have

$$D_n\to D'=\sup_{a,t}\left|w(a,\ t)-\frac{f\left(\frac{t}{\sqrt{a^T\Sigma a}}\right)tw_1(a)}{2(a^T\Sigma a)^{\frac{3}{2}}}\right|, \qquad (4.2)$$

where $w(a,\ t)$ is denoted as (2.5), $w_1(a)$ is also a zero mean Gaussian process whose sample paths continue uniformly and boundedly with respect to Euclidean norm $\|\cdot\|$, and covariance function is of analogous form as that of (2.5)

$$R(a,\ a_1)=P(a^TX)^2(a_1^TX)^2-(a^T\Sigma a)(a_1^T\Sigma a_1). \qquad (4.3)$$

We obtain the following result.

Theorem 4.1. *Assume that the conditions are fulfilled.*

i) *$f(t)$ continues uniformly and boundedly;*

ii) *there is a constant B such that for any $\varepsilon>0$ and any pair $(t,\ t_1)$, if*

$$|F_1(t)-F_1(t_1)|<\varepsilon$$

then

$$|f(t)t-f(t_1)t_1|<B\varepsilon; \qquad (4.5)$$

iii) *if for any $a>a_1>0$ for which $a-a_1<\varepsilon$, there exists a constant c such that*

$$\sup_t\left|f\left(\frac{t}{a}\right)\frac{t}{a}-f\left(\frac{t}{a_1}\right)\frac{t_1}{a}\right|<c\varepsilon, \qquad (4.6)$$

then we have for $\lambda>1$

$$P\{D'>\lambda\}\leqslant c\lambda^{2(d-1)+1}\exp\left\{-\frac{\lambda^2}{2\sigma^2}\right\}, \tag{4.7}$$

where

$$\sigma^2=\sup_{a,t}P\left(w(a,t)+\frac{f\left(\frac{t}{\sqrt{a^T\Sigma a}}\right)tw_1(a)}{2(a^T\Sigma a)^{\frac{3}{2}}}\right)^2.$$

Proof The proving procedure is almost the same as that of Theorem 3.1. We easily know that (3.10) continues to hold by noting that $\frac{a^TX}{\sqrt{a^T\Sigma a}}$ is distributed with the spherically symmetric distribution. Applying Lemmas 3.2—3.5, we can obtain (4.7) as long as we show that

$$\left(P\left\{\left(w(a,t)-\frac{f\left(\frac{t}{\sqrt{a^T\Sigma a}}\right)tw_1(a)}{2(a^T\Sigma a)^{\frac{1}{2}}}\right)-\left(w(a_i,t_{2j-1}^{a_i})-f\left(\frac{ta_{2j-1}^{a_i}}{\sqrt{a_i^T\Sigma a_i}}\right)\frac{t_{2j-1}^{a_i}w_1(a_i)}{2(a_i^T\Sigma a_i)^{\frac{3}{2}}}\right)\right\}\right)^{\frac{1}{2}}$$
$$\leqslant\beta \tag{4.8}$$

holds for $\gamma=c\beta^2$ and $\gamma_1=c\beta^2$. But this is relatively easy. Similar to Lemma 3.7

$$(P\{(w(a,t)-w(a_i,t_{2j-1}^{a_i}))^2\})^{\frac{1}{2}}\leqslant c\beta, \tag{4.9}$$

and similar to (3.16)

$$\left(P\left(f\left(\frac{t}{\sqrt{a^T\Sigma a}}\right)\frac{tw_1(a)}{2(a^T\Sigma a)^{\frac{3}{2}}}-f\left(\frac{t_{2j-1}^{a_i}}{\sqrt{a_i^T\Sigma a_i}}\right)\frac{t_{2j-1}^{a_i}w_1(a_i)}{2(a_i^T\Sigma a_i)^{\frac{3}{2}}}\right)^2\right)^{\frac{1}{2}}$$
$$\leqslant\left(P\left(f\left(\frac{t}{\sqrt{a^T\Sigma a}}\right)\frac{t(a^TX)^2}{2(a^T\Sigma a)^{\frac{3}{2}}}-f\left(\frac{t_{2j-1}^{a_i}}{\sqrt{a_i^T\Sigma a_i}}\right)\frac{t_{2j-1}^{a_i}(a_i^TX)^2}{2(a_i^T\Sigma a_i)^{\frac{3}{2}}}\right)^2\right)^{\frac{1}{2}}$$
$$\leqslant\left(P\left(f\left(\frac{t_{2j-1}^{a_i}}{\sqrt{a^T\Sigma a}}\right)\frac{t_{2j-1}^{a_i}}{2(a^T\Sigma a)}\frac{(a^TX)^2}{2(a^T\Sigma a)^2}-f\left(\frac{t}{\sqrt{a^T\Sigma a}}\right)\left(\frac{t}{2(a^T\Sigma a)}\right)\frac{(a^TX)^2}{2(a^T\Sigma a)^2}\right)^2\right)^{\frac{1}{2}}$$
$$+\left(P\left(f\left(\frac{t_{2j-1}^{a_i}}{\sqrt{a^T\Sigma a}}\right)\frac{t_{2j-1}^{a_i}}{2(a^T\Sigma a)}\frac{(a^TX)^2}{2(a^T\Sigma a)^2}\right.\right.$$
$$\left.\left.-f\left(\frac{t_{2j-1}^{a_i}}{\sqrt{a_i^T\Sigma a_i}}\right)\left(\frac{t_{2j-1}^{a_i}}{2(a_i^T\Sigma a_i)}\right)\frac{(a^TX)^2}{2(a^T\Sigma a)^2}\right)^2\right)^{\frac{1}{2}}$$
$$+\left(P\left(f\left(\frac{t_{2j-1}^{a_i}}{\sqrt{a_i^T\Sigma a_i}}\right)\frac{t_{2j-1}^{a_i}}{2(a_i^T\Sigma a_i)}\frac{(a_i^TX)^2}{2(a_i^T\Sigma a_i)^2}\right.\right.$$
$$\left.\left.-f\left(\frac{t_{2j-1}^{a_i}}{\sqrt{a_i^T\Sigma a_i}}\right)\left(\frac{t_{2j-1}^{a_i}}{2(a_i^T\Sigma a_i)}\right)\frac{(a^TX)^2}{2(a^T\Sigma a)^2}\right)^2\right)^{\frac{1}{2}}=I_8+I_9+I_{10}. \tag{4.10}$$

Since $\frac{a^TX}{\sqrt{a^T\Sigma a}}$ is distributed with F_1, by (4.5) we bound $I_8\leqslant c\beta$. And $I_9<c\beta$ due to (4.6). Finally, since $\inf_{a\in S_d}\sqrt{a^T\Sigma a}>0$ follows from the positive definiteness of Σ, utilizing (4.4) we easily verify

$$I_{10}\leqslant c\beta. \tag{4.11}$$

(4.8) is showed.

Remark 4.1. We have to point out that there are also many elliptically symmetric probability measure satisfying conditions (4.4)—(4.6), for example, multivariable normal distribution, the distribution whose one-dimensional density function is of the form $f(t)=\frac{c}{t^{\alpha}}$, $\alpha\geqslant 3$j and so on.

In view of Theorem 3.1 and Theorem 4.1, we know that one can construct an estimate for which both location parameter and dispersion matrix are unknown as follows.

$$D_n''=\sqrt{n}\sup_{a,t}\left|P_nI(a^TX\leqslant t)-F\left(\frac{t-a^T\theta_n}{\sqrt{a^T\Sigma_n a}}\right)\right|. \tag{4.12}$$

Similarly

$$D_n''\to\sup_{a,t}\left|w(a,\ t)+f\left(\frac{t-a^T\theta}{\sqrt{a^T\Sigma a}}\right)a^TN(0,\ v)+f\left(\frac{t-a^T\theta}{\sqrt{a^T\Sigma a}}\right)\frac{tw_1(a)}{2(a^T\Sigma a)^{\frac{3}{2}}}\right|=D''$$

as long as the conditions in Theorem 3.1 and Theorem 4.1 are fulfilled. And the following result is obtained.

Theorem 4. 2. *Suppose $P(\cdot,\ \theta,\ \Sigma)$ is an elliptically symmetric probability measure whose location parameter and dispersion matrix are unknown. θ_n and $a^T\Sigma_n a$ are, respectively, the estimates of θ and $a^T\Sigma a$, defined as in Theorem 3.1 and Theorem 4.1, F_1 is defined as (4.1). If the conditions (3.2), (3.3) and (4.4)—(4.6) are all fulfilled, then for $\lambda>1$*

$$P\{D''>\lambda\}\leqslant c\lambda^{2(d-1)+1}\exp\left\{-\frac{\lambda^2}{2\sigma^2}\right\}, \tag{4.13}$$

where

$$\sigma^2=\sup_{a,t}P\left\{w(a,\ t)+f\left(\frac{t-a^T\theta}{\sqrt{a^T\Sigma a}}\right)a^TN(0,\ v)+f\left(\frac{t-a^T\theta}{\sqrt{a^T\Sigma a}}\right)\frac{tw_1(a)}{2(a^T\Sigma a)^{\frac{3}{2}}}\right\}^2.$$

References

[1] Adler, R. J, & Samorodnisky, G., Tail behavior for the suprema of Gaussian process with applications to empirical processes, *Ann. Probab*,, 15 (1987), 1339—1351.

[2] Pollard, D., Convergence of stochastic processes, Springer-Verlag, New York, (1984).

[3] Zhu Lixing, Tail probability bounds for the suprema of P-bridges with applications to empirical processes (to appear in *Kexue Tongbao* (*Chinese Science Buletin*), Vol. 35.

[4] Cai, Y, H., PP tpye statistics for testing normality (manuscript, 1988).