Ann. Inst. Statist. Math.
Vol. 45, No. 4, 787–800 (1993)

ON NONPARAMETRIC TESTS FOR SYMMETRY IN R^m

SIGEO AKI

Department of Mathematical Science, Faculty of Engineering Science, Osaka University, Toyonaka 560, Japan

(Received October 2, 1991; revised February 17, 1993)

Abstract. This paper considers the problem for testing symmetry of a distribution in $\boldsymbol{R}^m$ based on the empirical distribution function. Limit theorems which play important roles for investigating asymptotic behavior of such tests are obtained. The limit processes of the theorems are multiparameter Wiener process. Based on the limit theorems, nonparametric tests are proposed whose asymptotic distributions are functionals of a multiparameter standard Wiener process. The tests are compared asymptotically with each other in the sense of Bahadur.

Key words and phrases: Asymptotic distribution, test for symmetry, L_1-norm, L_2-norm, empirical process, central limit theorems, goodness-of-fit tests, multiparameter Wiener process, density estimator, approximate Bahadur efficiency.

1. Introduction

In this paper, we deal with the problem of testing symmetry of a distribution in $\boldsymbol{R}^m$. Throughout the paper, we assume the center of symmetry is known. Hence, without loss of generality, we consider the problem for testing symmetry about $\mathbf{0}$.

In matched pair treatment effect experiments with m measurements on each member of n pairs (each pair has one treated and one control member), the natural null-hypothesis is symmetry about the zero vector. In this case, the difference "treatment-control" of the responses in each pair would be the basic $\boldsymbol{X}$ vectors and under the null-hypothesis of no treatment effect, symmetry holds. Here a "stochastically larger than symmetry" alternative is called for.

In the 1-dimensional case, many statistics have been proposed for the goodness-of-fit test for symmetry. For example, we can mention Butler (1969), Rothman and Woodroofe (1972), Shorack and Wellner ((1986), Section 22), Aki (1987), Csörgő and Heathcote (1987) and Aki and Kashiwagi (1989).

Let $\boldsymbol{X} = (X_1, X_2, \dots, X_m)'$ be a random vector in $\boldsymbol{R}^m$ which is symmetric about $\mathbf{0}$. Suppose that $F_1, F_2, \dots, F_m$ are 1-dimensional cumulative distribution functions which are symmetric about 0. Then the distribution of $(F_1(X_1), F_2(X_2), \dots, F_m(X_m))'$ is symmetric about $(1/2, 1/2, \dots, 1/2)'$. This can be seen as follows:

The symmetry of $\boldsymbol{X}$ and F's imply respectively that

$$(X_1, X_2, \ldots, X_m)' \stackrel{\mathrm{d}}{=} (-X_1, -X_2, \ldots, -X_m)'$$

and for every $x \in \boldsymbol{R}$, $F_i(x) + F_i(-x) = 1$, $i = 1, 2, \ldots, m$. Therefore, for each $x_1, x_2, \ldots, x_m \in \boldsymbol{R}$, it holds that

$$\begin{aligned} &P(1 - F_1(X_1) \le x_1, \ldots, 1 - F_m(X_m) \le x_m) \\ &\quad = P(1 - F_1(-X_1) \le x_1, \ldots, 1 - F_m(-X_m) \le x_m) \\ &\quad = P(F_1(X_1) \le x_1, \ldots, F_m(X_m) \le x_m). \end{aligned}$$

In the above statement, if we take $F_1, F_2, \ldots, F_m$ so as to be continuous and strictly increasing, every continuous and symmetric distribution about $\mathbf{0}$ in $\boldsymbol{R}^m$ is transformed to a continuous and symmetric distribution about $(1/2, 1/2, \ldots, 1/2)'$ in $[0, 1]^m$. Consequently, the problem of investigating the symmetry of a distribution in $\boldsymbol{R}^m$ can be reduced to that of investigating the symmetry of the transformed distribution in $[0, 1]^m$ as far as we know the center of symmetry.

Aki (1987) proposed a limit theorem which plays an important role for deriving asymptotic distributions of tests for symmetry based on the empirical distribution function. In Section 2, we give a limit theorem which can be regarded as the m-dimensional analogue of the limit theorem. The limiting process of the theorem is a (m-parameter) Wiener process w.r.t. a (m-dimensional) distribution function. Next, we consider a transformation of the process to a (m-parameter) standard Wiener process by using the idea of Khmaladze (1988). Further, these results are extended in a general framework on which the central limit theorem for empirical processes indexed by uniformly bounded families of functions is studied recently (cf. Giné and Zinn (1984, 1986)). In Section 3, an integral test is proposed whose asymptotic distribution is the L_2-norm of a (multiparameter) standard Wiener process. The distribution is also investigated. The test is not coordinate free (see Remark 3.1 below). Hence, the test procedure is not invariant under the rotations of the data without extra considerations. Some statisticians may regard it as a defect. However, if we always rotate the data to a following data-dependent coordinate system before transforming them into the unit cube, then we can make the test procedure to be invariant under the rotations. For example, consider the following rotation of the data: Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be independent observations in $\boldsymbol{R}^m$ whose distribution is assumed to be symmetric about $\mathbf{0}$. Let $\boldsymbol{a}_1$ be the unit vector with the same direction of $\sum_{i=1}^n \boldsymbol{X}_i$ and let $\langle\{\boldsymbol{a}_1\}\rangle$ be the linear subspace generated by $\boldsymbol{a}_1$. We denote by $\boldsymbol{X}_1^{(1)}, \ldots, \boldsymbol{X}_n^{(1)}$ the orthogonal projection of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ on $\langle\{\boldsymbol{a}_1\}\rangle^\perp$. Let $\boldsymbol{a}_2$ be the unit vector with the same direction of $\sum_{i=1}^n \boldsymbol{X}_i^{(1)}$. Next, we denote by $\boldsymbol{X}_1^{(2)}, \ldots, \boldsymbol{X}_n^{(2)}$ the orthogonal projection of $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ on $\langle\{\boldsymbol{a}_1, \boldsymbol{a}_2\}\rangle^\perp$. Continuing this procedure, we can define $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ almost surely. Before transforming the data into the unit cube, we always rotate the data so that $\boldsymbol{a}_i$ and $\boldsymbol{e}_i$ have the same direction for every $i = 1, 2, \ldots, m$, where $\boldsymbol{e}_i$ is the unit vector whose i-th component is 1. Then the final data transformed into the unit cube are invariant under any rotation of the original data, since the mean directions are equivariant under rotations of the data. Of course, there may be other methods for the test procedure to be invariant under rotations of the data and we can not show that the method given above is optimal in some sence. In Section 4, three tests including the test proposed in Section 3 are compared asymptotically with each other in the sense of Bahadur (1960). Approximate Bahadur slopes of the statistics are obtained. Approximate Bahadur efficiency has been subject to some criticism. As a first step, however, it is still useful criterion especially for test procedures which have nonnormal limiting distributions like this case. Wieand (1976) studied the relationship between the limiting approximate Bahadur efficiency as the alternative approaches the null hypothesis and the limiting Pitman efficiency as the size of the test tends to zero. Recently, further investigation for efficiency concepts including approximate Bahadur efficiency has been done by Koning (1992).

2. Asymptotic results

Let G be a continuous distribution function on $[0,1]^m$. A Gaussian process in $D([0,1]^m)$ with mean 0 and covariance function $G(\boldsymbol{x}\wedge\boldsymbol{x}')$ is called a Wiener process w.r.t. G, where $\boldsymbol{x}$ and $\boldsymbol{x}'$ belong to $[0,1]^m$. For the definition and some properties of $D([0,1]^m)$, see, for example, Bickel and Wichura (1971). Let $\boldsymbol{Y}_1, \boldsymbol{Y}_2, \dots, \boldsymbol{Y}_n$ be i.i.d. random vectors with distribution function G. Let $\xi_1, \xi_2, \dots, \xi_n$ be i.i.d. random variables with $E\xi_i = 0$ and $E\xi_i^2 = 1$. We assume that $(\boldsymbol{Y}_1, \boldsymbol{Y}_2, \dots, \boldsymbol{Y}_n)$ and $(\xi_1, \xi_2, \dots, \xi_n)$ are independent. We define a random element of $D([0,1]^m)$ by

$$\begin{aligned} u_n(\boldsymbol{t})&(= u_n(t_1, \dots, t_m)) \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i I([0,t_1] \times \cdots \times [0,t_m]; \boldsymbol{Y}_i), \end{aligned}$$

where $\boldsymbol{t} = (t_1, t_2, \dots, t_m)' \in [0,1]^m$ and $I(A;\cdot)$ denotes the indicator function of the set A.

THEOREM 2.1. *Under the above assumptions, $u_n(\boldsymbol{t})$ converges weakly to a Wiener process w.r.t. G.*

We can easily prove this by checking directly the moment condition of Bickel and Wichura (1971). We omit the proof, since this theorem is essentially a corollary of Theorem 2.4 below. But, we have to note that the definitions of weak convergence are slightly different.

As we explain in Section 3, the distribution function G is unknown and it must be estimated in the problem of testing symmetry. Moreover, for constructing an asymptotically distribution-free test statistic, a transformation of W_G to a standard Wiener process is needed, as Khmaladze (1988) discussed in the situation of usual goodness-of-fit tests in $\boldsymbol{R}^m$.

Now we assume for the distribution function G,

(A.1) G has continuous density g. And there exist constants $0 < c_0 < c_1 < \infty$ such that $c_0 \le g(\boldsymbol{y}) \le c_1$.

Let $g_n(\boldsymbol{y})$ be a density estimator based on $\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_n$. For the estimator $g_n(\boldsymbol{y})$ we assume

(A.2) $E(g_n(\boldsymbol{y}) - g(\boldsymbol{y}))^2$ converges to zero uniformly in $\boldsymbol{y} \in [0,1]^m$ as n tends to infinity.
Define a stochastic process by

$$w_n(\boldsymbol{x}) = \int_{y \le x} \frac{1}{\sqrt{g_n(\boldsymbol{y})}} du_n(\boldsymbol{y}).$$

THEOREM 2.2. *If assumptions* (A.1), (A.2) *and the conditions of Theorem* 2.1 *are satisfied,* w_n *converges weakly to a standard Wiener process* W *in* $D([0,1]^m)$, *as* n *tends to infinity.*

PROOF. From the definition of w_n, $w_n(\boldsymbol{x})$ can be written as

$$\frac{1}{\sqrt{n}} \sum_{Y_i \le x} \frac{1}{\sqrt{g_n(\boldsymbol{Y}_i)}} \xi_i.$$

Let $\tilde{w}_n(\boldsymbol{x}) = (1/\sqrt{n}) \sum_{i=1}^n \xi_i (1/\sqrt{g(\boldsymbol{Y}_i)}) \cdot I([\boldsymbol{0}, \boldsymbol{x}]; \boldsymbol{Y}_i)$. Then we have

$$w_n(\boldsymbol{x}) = \tilde{w}_n(\boldsymbol{x}) + \frac{1}{\sqrt{n}} \sum_{i=1}^n \left(\frac{1}{\sqrt{g_n(\boldsymbol{Y}_i)}} - \frac{1}{\sqrt{g(\boldsymbol{Y}_i)}} \right) \cdot I([\boldsymbol{0}, \boldsymbol{x}]; \boldsymbol{Y}_i) \cdot \xi_i. \tag{2.1}$$

Then, from assumptions (A.1) and (A.2), it is easy to see that the second term of the r.h.s. of (2.1) converges to zero in probability uniformly in $\boldsymbol{x}$. Thus, it suffices to show that $\tilde{w}_n$ converges weakly to a standard Wiener process. But it can be done by checking the moment conditions. This completes the proof. □

Here, we view our problem in a general framework of the central limit theorems for empirical processes indexed by uniformly bounded families of functions.

Let $Z_1, Z_2, \ldots$ be i.i.d. random elements of a Banach space B with norm $\|\cdot\|$. Suppose that $EZ_1 = 0$ and $E\|Z_1\|^2$ is finite. We assume that the central limit theorem holds for $Z_1, Z_2, \ldots$, i.e., there exists a B-valued Gaussian random element γ,

(A.3) $(1/\sqrt{n}) \sum_{i=1}^n Z_i$ converges weakly to γ on B.

Let $\xi_1, \xi_2, \ldots$ be i.i.d. real valued random variables with $E\xi_1 = 0$ and $E\xi_1^2 = 1$ and let $\{\xi_i\}$ and $\{Z_i\}$ be independent. We assume that

(A.4) $\int_0^\infty (P\{|\xi_1| > u\})^{1/2} du < \infty$.

The following theorem is given by Giné and Zinn (1984) (cf. also Giné and Zinn ((1986), Lemma 2.4 and Remark 2.5) and Ledoux and Talagrand (1986))

THEOREM 2.3. *Assume* (A.3) *and* (A.4). *Then* $(1/\sqrt{n}) \sum_{i=1}^n \xi_i Z_i$ *converges weakly to* γ *on* B.

Next, let $(S, \mathcal{S})$ be a measurable space and let P be a probability measure on it. Define the probability space by $(\Omega, \Sigma, \Pr) = (S^{\boldsymbol{N}}, \mathcal{S}^{\boldsymbol{N}}, P^{\boldsymbol{N}})$. Let $X_i : \Omega \to S$ be the projection of Ω into the i-th copy of S and let $P_n = (1/n) \sum_{i=1}^n \delta_{X_i}$ be the empirical measure associated to P. Denote by $\nu_n = \sqrt{n}(P_n - P)$ the normalized empirical process. Let $\mathcal{F}$ be a class of real valued measurable functions on $(S, \mathcal{S})$ and

$$\sup_{f\in\mathcal{F}} |f(s)| < \infty \qquad \text{for every } s \in S.$$

If we define

$$l^\infty(\mathcal{F}) = \{F \mid F : \mathcal{F} \to \boldsymbol{R} \text{ bounded}\},$$

then $l^\infty(\mathcal{F})$ is a Banach space with norm $\|F\| = \sup_{f\in\mathcal{F}} |F(f)|$. But $l^\infty(\mathcal{F})$ is not necessarily separable and we can not use weak convergence theory on complete separable metric spaces.

For a function $h : \Omega \to \boldsymbol{R}$, the upper integral of h is defined as

$$E^*h = \inf\left\{\int g(\omega)d\Pr(\omega); g \text{ is measurable, } g \geq h\right\}.$$

A class of functions $\mathcal{F}$ introduced above is called a P-Donsker class with envelope F_0 if the following conditions are satisfied:

(a) there exists a real valued measurable function F_0 on S such that $\int F_0(s)^2 P(ds) < \infty$ and for every $f \in \mathcal{F}$ and $s \in S$, $|f(s)| \leq F_0(s)$ holds.

(b) the central limit theorem holds for the empirical process indexed by $\mathcal{F}$, i.e., there exists a centered Radon Gaussian measure γ_P on $l^\infty(\mathcal{F})$ such that for every $H : l^\infty(\mathcal{F}) \to \boldsymbol{R}$ bounded and continuous,

$$E^*H(\nu_n) \to \int H d\gamma_P \qquad \text{as} \quad n \to \infty.$$

The Gaussian measure γ_P on $l^\infty(\mathcal{F})$ is the law of Gaussian process G_P indexed by $\mathcal{F}$ whose difference variance is given by

$$E(G_P(f) - G_P(g))^2 = P(f - g)^2 - (P(f - g))^2,$$

where $Pf = \int f dP$. G_P is a generalization of the Brownian bridge (cf. Giné and Zinn (1986)).

Let W_P be the Wiener process indexed by $\mathcal{F}$, i.e. $EW_P(f) = 0$ and $EW_P(f)W_P(g) = P(fg)$ for every f, $g \in \mathcal{F}$. Then we have

THEOREM 2.4. *Let $X_1, X_2, \ldots$ be i.i.d. random elements of $(S, \mathcal{S})$ and let the law of X_1 be P. Let $\xi_1, \xi_2, \ldots$ be i.i.d. real random variables with $E\xi_1 = 0$ and $E\xi_1^2 = 1$. Suppose that $\mathcal{F}$ is a P-Donsker class with envelope F_0. Then $(1/\sqrt{n})\sum_{i=1}^n \xi_i f(X_i)$ converges weakly to $W_P(f)$ in $l^\infty(\mathcal{F})$.*

PROOF. Since $\mathcal{F}$ is a P-Donsker class, the central limit theorem holds, i.e. $(1/\sqrt{n})\sum_{i=1}^n (f(X_i) - Pf)$ converges weakly to a generalized Brownian bridge $G_P(f)$. Then Theorem 2.3 implies that $(1/\sqrt{n})\sum_{i=1}^n \xi_i(f(X_i) - Pf)$ converges weakly to the same limit $G_P(f)$.

Note that $(1/\sqrt{n})\sum_{i=1}^n \xi_i Pf$ converges weakly to NPf, where N is (real) standard normal random variable and independent of G_P.

Thus $(1/\sqrt{n})\sum_{i=1}^{n}\xi_i f(X_i)$ converges weakly to $G_P(f) + NPf$ whose law is the same as the Wiener process indexed by $\mathcal{F}$. This completes the proof. □

Now we reconsider Theorems 2.1 and 2.2, which treat weak convergence of, respectively,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i I([\mathbf{0},\boldsymbol{t}];\,\boldsymbol{Y}_i) \quad \text{and} \quad \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\xi_i}{\sqrt{g(\boldsymbol{Y}_i)}}I([\mathbf{0},\boldsymbol{t}];\,\boldsymbol{Y}_i).$$

We set $S = [0,1]^m$. Denote by P_G the probability law of $\boldsymbol{Y}_1$. In the former case, we set

$$\mathcal{F} = \{I([\mathbf{0},\boldsymbol{t}];\cdot), \boldsymbol{t} \in S\}.$$

It can be seen that the (nonseparable) Banach space $(l^\infty(\mathcal{F}), \|\cdot\|)$ is regarded as $D([0,1]^m, \|\cdot\|_\infty)$, where $\|\cdot\|_\infty$ is the usual supremum norm. It is well known that the multivariate empirical process

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{I([\mathbf{0},\boldsymbol{t}];\,\boldsymbol{Y}_i) - G(\boldsymbol{t})\}$$

converges weakly to a Brownian bridge $W_G^0(\boldsymbol{t})$ on $D([0,1]^m, \|\cdot\|_\infty)$, where $EW_G^0(\boldsymbol{t}) = 0$ and $EW_G^0(\boldsymbol{s})W_G^0(\boldsymbol{t}) = G(\boldsymbol{s}\wedge\boldsymbol{t}) - G(\boldsymbol{s})G(\boldsymbol{t})$. Hence, $\mathcal{F}$ is P_G-Donsker class with envelope $I([\mathbf{0},\mathbf{1}];\cdot)$. Then Theorem 2.4 implies that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i I([\mathbf{0},\boldsymbol{t}];\,\boldsymbol{Y}_i) \to W_G$$

on $(D([0,1]^m, \|\cdot\|_\infty)$.

In the latter case, we set

$$\mathcal{F} = \left\{\frac{1}{\sqrt{g(\cdot)}}I([\mathbf{0},\boldsymbol{t}];\cdot), \boldsymbol{t} \in S\right\}.$$

We can see that $\mathcal{F}$ is a P_G-Donsker class with envelope $\sqrt{c_0}^{-1}I([\mathbf{0},\mathbf{1}];\cdot)$ by virtue of Assumption (A.1). Define for every $\boldsymbol{t} \in S$

$$f_t(\boldsymbol{y}) = \frac{1}{\sqrt{g(\boldsymbol{y})}}I([\mathbf{0},\boldsymbol{t}];\boldsymbol{y}) \in \mathcal{F}.$$

Then the central limit theorem implies

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(f_t(\boldsymbol{Y}_i) - Ef_t(\boldsymbol{Y}_i)) \to G_P(f_t).$$

Hence, from the proof of Theorem 2.4, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i Ef_t(\boldsymbol{Y}_i) \to N\cdot Ef_t(\boldsymbol{Y}_1) = N\cdot\int_{y\le t}\sqrt{g(\boldsymbol{y})}d\boldsymbol{y}$$

and

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i f_t(\boldsymbol{Y}_i) \to G_P(f_t) + N \cdot Ef_t(\boldsymbol{Y}_1).$$

The covariance of the limit process is the following:

$$\begin{aligned} &E(G_P(f_s) + N \cdot Ef_s(\boldsymbol{Y}_1))(G_P(f_t) + N \cdot Ef_t(\boldsymbol{Y}_1)) \\ &\quad = EG_P(f_s)G_P(f_t) \ + \ Ef_s(\boldsymbol{Y}_1) \cdot Ef_t(\boldsymbol{Y}_1) \\ &\quad = \int \frac{1}{g(\boldsymbol{y})} I([\boldsymbol{0}, \boldsymbol{s} \wedge \boldsymbol{t}]; \boldsymbol{y}) g(\boldsymbol{y}) d\boldsymbol{y} = \boldsymbol{s} \wedge \boldsymbol{t}, \end{aligned}$$

which is the covariance of a multiparameter standard Wiener process.

3. Tests for symmetry

Let $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n$ be i.i.d. random variables with continuous distribution function F on $[0,1]^m$, where $\boldsymbol{X}_i = (X_{i1}, X_{i2}, \ldots, X_{im})'$, $i = 1, 2, \ldots, n$ are column vectors. Define $\boldsymbol{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{im})'$ and ξ_i by

$$Y_{i1} = \begin{cases} 2X_{i1} & \text{if } X_{i1} \le 1/2, \\ 2(1 - X_{i1}) & \text{if } X_{i1} > 1/2, \end{cases}$$
$$Y_{ij} = \begin{cases} X_{ij} & \text{if } X_{i1} \le 1/2, \\ (1 - X_{ij}) & \text{if } X_{i1} > 1/2, \text{ for } j = 2, 3, \ldots, m, \end{cases}$$

and

$$\xi_i = \begin{cases} 1 & \text{if } X_{i1} \le 1/2, \\ -1 & \text{if } X_{i1} > 1/2. \end{cases}$$

Remark 3.1. In the definitions of $\boldsymbol{Y}_i$ and ξ_i, X_{i1} plays a special role. Of course, we can use X_{ij} $(j \neq 1)$ instead of X_{i1} and then the corresponding proposition to Proposition 3.1 below holds.

PROPOSITION 3.1. *If F is symmetric w.r.t. $(1/2, \ldots, 1/2)'$, $\boldsymbol{Y}_i$ and ξ_i are independent, and hence $(\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n)$ and $(\xi_1, \ldots, \xi_n)$ are independent.*

PROOF. For $t_1, t_2, \ldots, t_m \in [0,1]$, we have that

$$\begin{aligned} &P(Y_{i1} \le t_1, Y_{i2} \le t_2, \ldots, Y_{im} \le t_m) \\ &\quad = P(Y_{i1} \le t_1, Y_{i2} \le t_2, \ldots, Y_{im} \le t_m, X_{i1} \le 1/2) \\ &\qquad + P(Y_{i1} \le t_1, Y_{i2} \le t_2, \ldots, Y_{im} \le t_m, X_{i1} > 1/2) \\ &\quad = P(X_{i1} \le t_1/2, X_{i2} \le t_2, \ldots, X_{im} \le t_m) \\ &\qquad + P(X_{i1} \ge 1 - t_1/2, X_{i2} \ge 1 - t_2, \ldots, X_{im} \ge 1 - t_m) \\ &\quad = 2F(t_1/2, t_2, \ldots, t_m), \end{aligned}$$

where the last equality follows from the fact that cubes

$$[0, t_1/2] \times [0, t_2] \times \cdots \times [0, t_m]$$

and

$$[1-t_1/2,1]\times[1-t_2,1]\times\cdots\times[1-t_m,1]$$

are symmetric w.r.t. $(1/2,\ldots,1/2)'$. It is clear that

$$P(\xi_i=1)=P(X_{i1}\le 1/2)=1/2.$$

Similarly as above, we also have that

$$\begin{aligned}P(Y_{i1}&\le t_1,Y_{i2}\le t_2,\ldots,Y_{im}\le t_m,\xi_i=1)\\&=P(X_{i1}\le t_1/2,X_{i2}\le t_2,\ldots,X_{im}\le t_m)\\&=F(t_1/2,t_2,\ldots,t_m).\end{aligned}$$

Thus, we obtain that $P(\boldsymbol{Y}_i\le \boldsymbol{t},\xi_i=1)=P(\boldsymbol{Y}_i\le\boldsymbol{t})\cdot P(\xi_i=1)$, which completes the proof. □

By using these $\boldsymbol{Y}$'s and ξ's, we set u_n, g_n and w_n as described in Section 2. Then, from Theorem 2.2 and Proposition 3.1, we can conclude that, if F is symmetric w.r.t. $(1/2,\ldots,1/2)'$ and the conditions of Theorem 2.2 are satisfied, the stochastic process w_n converges to a standard Wiener process W in $D([0,1]^m)$, as n tends to infinity.

If we define the L_2-norm of w_n as a test statistic for symmetry w.r.t. $(1/2,\ldots,1/2)'$, the statistic converges weakly to the L_2-norm of a multiparameter standard Wiener process under the null hypothesis from Theorem 2.2 and the continuous mapping theorem. Thus, what we have to investigate for the testing problem is to derive the distribution of the L_2-norm of W. Though few results for functionals of a multiparameter standard Wiener process are known, some properties of functionals of the multiparameter Brownian bridge are well-known. For example, the distribution of the L_2-norm of the two-parameter Brownian bridge was obtained by Blum *et al.* (1961). Cotterill and Csörgő (1985) also investigated the distribution when the number of parameters is greater than two. Since the standard Wiener process is simpler than the Brownian bridge, we can immediately give the results for W corresponding to the above.

As we have currently used, let $W=W(y_1,y_2,\ldots,y_m)$ be a m-parameter standard Wiener process. Then, we can write

$$W(y_1,\ldots,y_m)=2^{m/2}\sum_{n_1=0}^{\infty}\cdots\sum_{n_m=0}^{\infty}\left(\prod_{i=1}^{m}\frac{\sin((n_i+1/2)\pi y_i)}{(n_i+1/2)\pi}\right)\cdot Z_{n_1\cdots n_m},$$

where $Z_{n_1\cdots n_m}$ are i.i.d. sequence having standard normal distribution (see Cotterill and Csörgő (1985)). By using this formula, we have

$$\begin{aligned}T_m&=\int_0^1\cdots\int_0^1 W^2(y_1,\ldots,y_m)dy_1\cdots dy_m\\&=\sum_{n_1=0}^{\infty}\cdots\sum_{n_m=0}^{\infty}\left(\prod_{i=1}^{m}\frac{1}{(n_i+1/2)^2\pi^2}\right)(Z_{n_1\cdots n_m})^2.\end{aligned}$$

Hence, the characteristic function of T_m can be written as

$$\Psi_m(t) = \prod_{n_1=0}^{\infty} \cdots \prod_{n_m=0}^{\infty} \left(1 - \frac{2\sqrt{-1}t}{(\pi^{2m}\prod_{i=1}^{m}(n_i + 1/2)^2)}\right)^{-1/2}.$$

From this formula, we can easily derive that the n-th cumulant of T_m is given by

$$\kappa_n = 2^{(2m+1)n-(m+1)}(2^{2n} - 1)^m B_n^m (n-1)!((2n)!)^{-m},$$

where B_n denotes the Bernoulli number.

Aki (1990) calculated the distribution function of T_2 numerically by Imhof's (1961) method and gave a table of the percentage points of some selected values. Though we can hardly expect the corresponding numerical calculation by the same method for $m > 2$, Martynov's (1975) result would be applicable. Further, we can use at least the Cornish-Fisher asymptotic expansion to calculate the critical values, since we have obtained all cumulants of the distribution. The method for the L_2-norm of the multiparameter Brownian bridge was introduced by Cotterill and Csörgő (1982). In addition to the above methods for numerical calculation, we can note that the bootstrap approximation of the distribution function of the functional of the multiparameter Wiener process is useful.

4. Asymptotic behavior of tests under the alternative hypothesis

In the previous section, an integral test

$$T_n^{(2)} = \sqrt{\int_E w_n^2(\boldsymbol{y})d\boldsymbol{y}}$$

was proposed, where $E = [0,1]^m$, and its asymptotic distribution under the null hypothesis was shown to be the distribution of the L_2-norm of an m-parameter standard Wiener process. Instead of $T_n^{(2)}$, it is very natural to define the following statistics for the same problem,

$$T_n^{(1)} = \int_E |w_n(\boldsymbol{y})|d\boldsymbol{y} \quad \text{or} \quad T_n^{(3)} = \sup_{\boldsymbol{y}\in E} |w_n(\boldsymbol{y})|.$$

Now, we investigate asymptotic behavior of these statistics under the alternative hypothesis. Let $\mathcal{F}$ be the totality of the probability density functions on E which satisfy (A.1). Let $\mathcal{F}_0 = \{f \in \mathcal{F} \mid f \text{ is symmetric about } \boldsymbol{c} = (1/2,\ldots,1/2)\}$ and $\mathcal{F}_1 = \mathcal{F}\setminus\mathcal{F}_0$. In this section, we assume that $\boldsymbol{X_1},\ldots,\boldsymbol{X_n}$ are i.i.d. random vectors with pdf $f \in \mathcal{F}$. Our problem is to test $H_0 : f \in \mathcal{F}_0$ against $H_1 : f \in \mathcal{F}_1$ based on $\boldsymbol{X_1},\ldots,\boldsymbol{X_n}$.

Let $\boldsymbol{X_1},\ldots,\boldsymbol{X_n}$ be an i.i.d. sample with density $f \in \mathcal{F}$. We denote by $g(\boldsymbol{y})$ the pdf of $\boldsymbol{Y_1},\ldots,\boldsymbol{Y_n}$. Let F be the distribution function of f and $\alpha = F(1/2,1,\ldots,1)$. Define

$$g_1(\boldsymbol{x}) = \frac{f(x_1/2, x_2,\ldots,x_m)}{2\alpha} \quad \text{and} \quad g_2(\boldsymbol{x}) = \frac{f(1-x_1/2, 1-x_2,\ldots,1-x_m)}{2(1-\alpha)}.$$

Then we see that the conditional pdf of $\boldsymbol{Y}_i$ given that $\xi_i = 1$ ($\xi_i = -1$) is g_1 (resp. g_2) and $g = \alpha g_1 + (1-\alpha)g_2$. We set

$$\phi(f, \boldsymbol{x}) = \int_{\boldsymbol{y} \leq \boldsymbol{x}} \frac{\alpha g_1(\boldsymbol{y}) - (1-\alpha)g_2(\boldsymbol{y})}{\sqrt{g(\boldsymbol{y})}} d\boldsymbol{y}.$$

We also assume that the density estimator $g_n(\boldsymbol{y})$ satisfies $\sup_{\boldsymbol{y} \in E} |g_n(\boldsymbol{y}) - g(\boldsymbol{y})| \to 0$ in probability. Since $f \in \mathcal{F}$, the density $g(\boldsymbol{y})$ becomes uniformly continuous on E under the null and alternative hypotheses. If we use a kernel density estimator with bounded kernel with compact support and fold back $g_n(\boldsymbol{y})I(E^c; \boldsymbol{y})$ symmetrically with respect to the boundary, the above property follows from Theorems of Devroye and Wagner (1980).

Definition 4.1. (Bahadur (1960)) Let $\{P_\theta, \theta \in \Theta\}$ be a set of probability measures on a measurable space $(S, \mathcal{S})$. Let Θ_0 be some subset of Θ and let H be the hypothesis that $\theta \in \Theta_0$. Suppose we are given a sequence of real-valued statistics $\{T_n\}$ defined on $(S, \mathcal{S})$ based on a sample of size n. We say that $\{T_n\}$ is a standard sequence if the following three conditions are satisfied:

(i) There exists a continuous probability distribution function G such that, for each $\theta \in \Theta_0$,

$$\lim_{n \to \infty} P(T_n < x) = G(x), \qquad \text{for every } x.$$

(ii) There exists a constant a, $0 < a < \infty$, such that

$$\log(1 - G(x)) = -\frac{ax^2}{2}[1 + o(1)],$$

where $o(1) \to 0$ as $x \to \infty$.

(iii) There exists a real valued function $b(\theta)$ on $\Theta \setminus \Theta_0$, with $0 < b(\theta) < \infty$, such that, for each $\theta \in \Theta \setminus \Theta_0$,

$$\lim_{n \to \infty} P\left(\left|\frac{T_n}{\sqrt{n}} - b(\theta)\right| > x\right) = 0 \qquad \text{for every } x.$$

For every standard sequence $\{t_n\}$, $c(\theta) = ab^2(\theta)$ is called approximate Bahadur slope. For two standard sequences $\{t_n^{(1)}\}$ and $\{t_n^{(2)}\}$, the approximate Bahadur efficiency of $\{t_n^{(1)}\}$ to $\{t_n^{(2)}\}$ is defined by the ratio of the approximate Bahadur slopes, $c_1(\theta)/c_2(\theta)$.

Theorem 4.1. *$\{T_n^{(1)}\}$, $\{T_n^{(2)}\}$ and $\{T_n^{(3)}\}$ are standard sequences in the sense of Bahadur for each integer $m \geq 1$.*

For proving the theorem, we need to prove the following lemmas.

Lemma 4.1. *Let X be a $C(E)$-valued Gaussian random element with $E(X(\boldsymbol{t})) = 0$ and $E(X(\boldsymbol{s})X(\boldsymbol{t})) = R(\boldsymbol{s}, \boldsymbol{t})$ for $\boldsymbol{s}$ and $\boldsymbol{t} \in E$. Suppose that $R(\boldsymbol{s}, \boldsymbol{t}) \geq 0$ for each $\boldsymbol{s}$ and $\boldsymbol{t} \in E$. Then it holds that*

$$\lim_{t \to \infty} \frac{1}{t^2} \log P\left(\int_E |X(\boldsymbol{s})| d\boldsymbol{s} > t\right) = -\frac{1}{2 \iint_{E \times E} R(\boldsymbol{s}, \boldsymbol{t}) d\boldsymbol{s} d\boldsymbol{t}}.$$

Remark 4.1. If $X(\boldsymbol{t}) = W(\boldsymbol{t})$, then $R(\boldsymbol{s}, \boldsymbol{t}) = \boldsymbol{s} \wedge \boldsymbol{t}$ and hence it holds that

$$\lim_{t\to\infty} \frac{1}{t^2} \log P\left(\int_E |W(\boldsymbol{s})| d\boldsymbol{s} > t\right) = -\frac{3^m}{2}.$$

Aki and Kashiwagi (1989) proved the result of Lemma 4.1 for $m = 1$. Since the extension to the multivariate case is straightforward, we omit the proof.

LEMMA 4.2. *If $\boldsymbol{X_1}, \boldsymbol{X_2}, \dots, \boldsymbol{X_n}$ are i.i.d. r.v.'s with pdf $f \in \mathcal{F}$, then*

$$\sup_{\boldsymbol{x}\in E} \left| \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sqrt{g(\boldsymbol{Y}_i)}} I([\boldsymbol{0}, \boldsymbol{x}]; \boldsymbol{Y}_i)\xi_i - \phi(f, \boldsymbol{x}) \right| \to 0 \qquad \textit{almost surely.}$$

PROOF. Since $\phi(f, \boldsymbol{x})$ is the mean function of $(1/n)\sum_{i=1}^{n}(1/\sqrt{g(\boldsymbol{Y}_i)})I([\boldsymbol{0}, \boldsymbol{x}]; \boldsymbol{Y}_i)\xi_i$, this statement is a special case of a uniform law of large numbers. Noting that

$$\left| \frac{1}{\sqrt{g(\boldsymbol{Y}_i)}} I([\boldsymbol{0}, \boldsymbol{x}]; \boldsymbol{Y}_i)\xi_i \right| \le \frac{1}{\sqrt{g(\boldsymbol{Y}_i)}} \equiv F_i(\omega) \qquad \text{(say)}.$$

Since $EF_i^2(\omega) = 1$, $\sum_{i=1}^{\infty}(EF_i^2(\omega)/i^2) < \infty$ holds. Then, Theorem 8.3 of Pollard (1990) implies the result. This completes the proof. □

LEMMA 4.3. *For every $m \in \boldsymbol{N}$, it holds that*

$$\lim_{t\to\infty} \frac{1}{t^2} \log P\left(\sup_{\boldsymbol{x}\in E} |W(\boldsymbol{x})| > t\right) = -\frac{1}{2}.$$

PROOF. From Theorem 5.3 of Adler (1990), there exists a finite positive constant C such that for all $t > 0$,

$$\begin{aligned} C^{-1}t^m(\log t)^{-m/2}(1 - \Phi(t)) &\le P\left(\sup_{\boldsymbol{x}\in E} |W(\boldsymbol{x})| > t\right) \\ &\le Ct^m(1 - \Phi(t)), \end{aligned}$$

where $\Phi(t)$ denotes the cumulative distribution function of the standard normal distribution.

Then, by using the well-known inequality, for $w > 0$,

$$\frac{1}{\sqrt{2\pi}} \left\{ \frac{1}{w} - \frac{1}{w^3} \right\} \exp\left(-\frac{w^2}{2}\right) \le \Phi(-w) \le \frac{1}{\sqrt{2\pi}} \frac{1}{w} \exp\left(-\frac{w^2}{2}\right),$$

we have the desired result. This completes the proof. □

PROOF OF THEOREM 4.1. From Theorem 2.2 and the continuous mapping theorem, $\{T_n^{(1)}\}$, $\{T_n^{(2)}\}$ and $\{T_n^{(3)}\}$ converge in law to the distributions under the null hypothesis with distribution functions G_1, G_2 and G_3, respectively, where

$$G_1(x) = P\left(\int_E |W(\boldsymbol{y})|d\boldsymbol{y} \le x\right),$$

$$G_2(x) = P\left(\sqrt{\int_E W^2(\boldsymbol{y})d\boldsymbol{y}} \le x\right)$$

and

$$G_3(x) = P\left(\sup_{\boldsymbol{y}\in E} |W(\boldsymbol{y})| \le x\right).$$

Lemmas 4.1 and 4.3 imply that the sequence $\{T_n^{(1)}\}$ and $\{T_n^{(3)}\}$ satisfy the second condition of Definition 4.1 with $a = 3^m$ and $a = 1$, respectively. Since the distribution of G_2 is a quadratic form of i.i.d. normal random variables (see Section 3), we have

$$\log(1 - G_2(x)) = \log P\left(\int_E W^2(\boldsymbol{y})d\boldsymbol{y} > x^2\right) = -\frac{x^2}{2\lambda_1}(1 + o(1)),$$

from the result of Zolotarev (1961), where $\lambda_1 = (2/\pi)^{2m}$. Thus, G_2 satisfies the second condition of Definition 4.1 with $a = (\pi/2)^{2m}$.

Now we check the third condition of Definition 4.1. First, we show that $\{T_n^{(3)}\}$ satisfies the condition with $b(f) = \sup_{\boldsymbol{x}\in E} |\phi(f, \boldsymbol{x})|$. For $f \in \mathcal{F}$, it is seen that

$$\begin{aligned}
&\left|\sup_{\boldsymbol{x}\in E}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{g_n(\boldsymbol{Y}_i)}}I([\boldsymbol{0},\boldsymbol{x}];\,\boldsymbol{Y}_i)\xi_i\right| - \sup_{\boldsymbol{x}\in E}|\phi(f,\boldsymbol{x})|\right| \\
&\le \left|\sup_{\boldsymbol{x}\in E}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{g_n(\boldsymbol{Y}_i)}}I([\boldsymbol{0},\boldsymbol{x}];\,\boldsymbol{Y}_i)\xi_i\right|\right. \\
&\qquad \left. - \sup_{\boldsymbol{x}\in E}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{g(\boldsymbol{Y}_i)}}I([\boldsymbol{0},\boldsymbol{x}];\,\boldsymbol{Y}_i)\xi_i\right|\right| \\
&\qquad + \left|\sup_{\boldsymbol{x}\in E}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{g(\boldsymbol{Y}_i)}}I([\boldsymbol{0},\boldsymbol{x}];\,\boldsymbol{Y}_i)\xi_i\right| - \sup_{\boldsymbol{x}\in E}|\phi(f,\boldsymbol{x})|\right|.
\end{aligned}$$

The first term of the r.h.s. of the above inequality is not greater than

$$\begin{aligned}
&\sup_{\boldsymbol{x}\in E}\left|\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\sqrt{g_n(\boldsymbol{Y}_i)}} - \frac{1}{\sqrt{g(\boldsymbol{Y}_i)}}\right)I([\boldsymbol{0},\boldsymbol{x}];\,\boldsymbol{Y}_i)\xi_i\right| \\
&\qquad \le C\cdot\sup_{\boldsymbol{y}\in E}|g_n(\boldsymbol{y}) - g(\boldsymbol{y})|,
\end{aligned}$$

where C is a constant depending only on f. By the assumption in this section, we see that the first term of the above inequality converges in probability to zero. The second term of the r.h.s. of the inequality is not greater than

$$\sup_{\boldsymbol{x}\in E}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{g(\boldsymbol{Y}_i)}}I([\boldsymbol{0},\boldsymbol{x}];\,\boldsymbol{Y}_i)\xi_i - \phi(f,\boldsymbol{x})\right|.$$

But, this converges to zero almost surely by Lemma 4.2. It is easy to see that $b(f) \geq 0$ and that $b(f) = 0$ if and only if $f \in \mathcal{F}_0$. Secondly, we show that $\{T_n^{(1)}\}$ satisfies the third condition of Definition 4.1 with $b(f) = \int_E |\phi(f, \boldsymbol{x})| d\boldsymbol{x}$. Note that

$$\left| \int_E \left| \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sqrt{g_n(\boldsymbol{Y}_i)}} I([\boldsymbol{0}, \boldsymbol{x}];\, \boldsymbol{Y}_i)\xi_i \right| d\boldsymbol{x} - \int_E |\phi(f, \boldsymbol{x})| d\boldsymbol{x} \right|$$
$$\leq \sup_{\boldsymbol{x} \in E} \left| \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sqrt{g_n(\boldsymbol{Y}_i)}} I([\boldsymbol{0}, \boldsymbol{x}];\, \boldsymbol{Y}_i)\xi_i - \phi(f, \boldsymbol{x}) \right|.$$

Then, we can see it from the above argument. Similarly, we see that $\{T_n^{(2)}\}$ satisfies the third condition of Definition 4.1 with $b(f) = \sqrt{\int_E \phi^2(f, \boldsymbol{x}) d\boldsymbol{x}}$. This completes the proof. □

COROLLARY 4.1. *The approximate Bahadur slopes of $\{T_n^{(1)}\}$, $\{T_n^{(2)}\}$ and $\{T_n^{(3)}\}$ are given by $3^m \left(\int_E |\phi(f, \boldsymbol{x})| d\boldsymbol{x}\right)^2$, $(\pi/2)^{2m} \int_E \phi^2(f, \boldsymbol{x}) d\boldsymbol{x}$ and $(\sup_{\boldsymbol{x} \in E} |\phi(f, \boldsymbol{x})|)^2$, respectively.*

Acknowledgements

I would like to express my appreciation to the referees for their helpful comments. The general approach in Section 2 including Theorems 2.3 and 2.4, which was suggested by one of the referees, has much improved the earlier version of the manuscript.

REFERENCES

Adler, R. J. (1990). *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*, IMS Lecture Notes-Monograph Series, Vol. 12, Hayward, California.

Aki, S. (1987). On nonparametric tests for symmetry, *Ann. Inst. Statist. Math.*, **39**, 457–472.

Aki, S. (1990). On a test based on a sequential empirical process (unpublished manuscript).

Aki, S. and Kashiwagi, N. (1989). Asymptotic properties of some goodness-of-fit tests based on the L_1-norm, *Ann. Inst. Statist. Math.*, **41**, 753–764.

Bahadur, R. R. (1960). Stochastic comparison of tests, *Ann. Math. Statist.*, **31**, 276–295.

Bickel, P. J. and Wichura, M. J. (1971). Convergence criteria for multiparameter stochastic process and some applications, *Ann. Math. Statist.*, **42**, 1656–1670.

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

Butler, C. (1969). A test for symmetry using the sample distribution function, *Ann. Math. Statist.*, **40**, 2209–2210.

Cotterill, D. S. and Csörgő, M. (1982). On the limiting distribution of and critical values for the multivariate Cramér-von Mises statistic, *Ann. Statist.*, **10**, 233–244.

Cotterill, D. S. and Csörgő, M. (1985). On the limiting distribution of and critical values for the Hoeffding, Blum, Kiefer, Rosenblatt independence criterion, *Statist. Decisions*, **3**, 1–48.

Csörgő, S. and Heathcote, C. R. (1987). Testing for symmetry, *Biometrika*, **74**, 177–184.

Devroye, L. P. and Wagner, T. J. (1980). The strong uniform consistency of kernel density estimates, *Multivariate Analysis*, *V* (ed. P. R. Krishnaiah), 59–77, North-Holland, New York.

Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes, *Ann. Probab.*, **12**, 929–989.

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

Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables, *Biometrika*, **48**, 419–426.

Khmaladze, E. V. (1988). An innovation approach to goodness-of-fit tests in R^m, *Ann. Statist.*, **16**, 1503–1516.

Koning, A. J. (1992). Approximation of stochastic integrals with applications to goodness-of-fit tests, *Ann. Statist.*, **20**, 428–454.

Ledoux, P. M. and Talagrand, M. (1986). Conditions d'intégrablité pour les multiplicateurs dans le tlc Banachique, *Ann. Probab.*, **14**, 916–921.

Martynov, G. V. (1975). Computation of distribution functions of quadratic forms of normally distributed random variables, *Theory Probab. Appl.*, **20**, 782–793.

Pollard, D. (1990). *Empirical Processes: Theory and Applications*, NSF-CBMS Regional Conference Series in Probability and Statistics, **2**, IMS, Hayward, California.

Rothman, E. D. and Woodroofe, M. (1972). A Cramér-von Mises type statistic for testing symmetry, *Ann. Math. Statist.*, **43**, 2035–2038.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics,* Wiley, New York.

Wieand, H. S. (1976). A condition under which the Pitman and Bahadur approaches to efficiency coincide, *Ann. Statist.*, **4**, 1003–1011.

Zolotarev, V. M. (1961). Concerning a certain probability problem, *Theory Probab. Appl.*, **6**, 201–204.