Statistical Methodology 6 (2009) 230–250

Contents lists available at ScienceDirect

Statistical Methodology

journal homepage: www.elsevier.com/locate/stamet

Testing for symmetry in multivariate distributions

Joseph Ngatchou-Wandji

Département de mathématiques, Université de Caen, Campus 2, BP 5186, 14032 Caen, France
LMNO, Université de Caen, 14032 Caen, France

ARTICLE INFO

Article history:
Received 11 March 2008
Received in revised form
29 August 2008
Accepted 19 September 2008

Keywords:
Contiguity
Empirical characteristic function
Reflected symmetry
Testing hypothesis
Weak convergence

ABSTRACT

Using the empirical characteristic function, a Cramér–von Mises test for reflected symmetry about an unspecified point is derived for multivariate distributions. The test statistic is based on an empirical process for which the weak convergence is established. The null properties of the test are studied as well as its power and local power. Estimators for the unknown symmetric point are previously proposed. Their consistency and asymptotical normality are proved by studying the weak convergence of some multidimensional empirical process. A simulation experiment shows that the estimators of the symmetric point are good, and that the test performs well on the examples tested. The new test is compared to the one derived in [N. Henze, B. Klar, S.G. Meintanis, Invariant tests for symmetry about an unspecified point based on empirical characteristic function, J. Multivariate. Anal. 87 (2003) 275–297].

1. Introduction

In some areas such as economy and finance, testing for symmetry of a distribution may be very important (see, e.g., [1,5,14,6,27,35]). The major part of the considerable literature concerned with this subject is devoted to the univariate case. For a review, see for example [2,4,10–13,15,32,3,33]. The multivariate case has received much attention these last years. In this setting, there are generally two types of symmetry which both match in the univariate case. The most popular one is the elliptical (or spherical) symmetry. Papers dealing with it are, among others, [26,8,28,7,16,29,35,17]. The second type of symmetry is the reflected (or diagonal) symmetry which is involved in this paper. Some relevant works are [18,22,36,24].

The test proposed in [18] is of Cramér–von Mises type. It is based on observations centered at the known symmetric point. The one in [22] is a Kolmogorov–Smirnov with observations centered at an

E-mail address: joseph.ngatchou-wandji@unicaen.fr.

1572-3127/$ – see front matter
doi:10.1016/j.stamet.2008.09.003

estimator of the unknown symmetric point. Both papers suffer the fact that the computation of the tails of the limiting distribution of the test statistic is intractable. Neuhaus and Zhu [36] proposed for these two tests, their so-called permutation counterparts. Although they lead to conditional distribution-free statistics, they need a Monte Carlo procedure for determining the critical value of the test. But it is well known that such procedures can be time consuming and costly. Henze, Klar and Meintanis [24] derived the same type of test with different techniques.

Let X be a d-dimensional random vector defined on some probability space $(\Omega, \mathcal{A}, P)$ with distribution law P_X. Let $\mathcal{S}$ be the collection of all reflected symmetric distribution laws. Formally speaking, testing the reflected symmetry of X is tantamount to testing $H_0 : P_X \in \mathcal{S}$ against $H_1 : P_X \notin \mathcal{S}$. It is clear that $P_X \in \mathcal{S}$ means $P_X = P_\mu$ for some $P_\mu \in \mathcal{S}$, where $\mu \in \mathbb{R}^d$ is the symmetric point. As in [36,24], we propose for this testing problem a Cramér–von Mises type test. In our approach, the observations are centered at a consistent and asymptotical normal estimator of the unknown symmetric point, and the asymptotic distribution of our test statistic under H_0 is that of a weighted sum of independent $\chi^2(1)$ random variables. The derivation of this result makes use of the principal component decomposition of a non-stationary Gaussian process. The resulting test has some advantages over the aforementioned ones. Indeed, theoretically, the assumptions needed are weaker and the convergence results established are more general. Practically, the new test is more flexible and is easier to implement.

In Section 2, we derive two classes of estimators for the symmetric point μ. Their consistency and asymptotical normality are deduced from the weak convergence of some d-dimensional empirical process. In Section 3, we derive our test statistic and study its behavior under the null hypothesis by establishing a functional central limit theorem for some empirical process. In Section 4, the power of the test is investigated under a general alternative H_1. It is next studied under a sequence of local alternatives H_1^n approaching H_0 from a fixed direction. In Section 5, a simulation experiment is presented and discussed. There, a comparison study is done with the test derived in [24]. The proofs of the results are given in the last section.

2. Estimating the symmetric point

In this section, we aim to provide consistent and asymptotical normal estimators which can be more easier to compute than those in Heathcote, Rachev and Cheng [22] and Koutrouvelis [32].

For $t \in \mathbb{R}^d$, denote by t' the transpose of t. Let $t = (t_1, \ldots, t_d)'$, $s = (s_1, \ldots, s_d)' \in \mathbb{R}^d$; $t's$ stands for the Euclidean scalar product of t and s, namely $t's = t_1 s_1 + \cdots + t_d s_d$, and $\|t\| = (t't)^{1/2} = (t_1^2 + \cdots + t_d^2)^{1/2}$ is the Euclidean norm of t. We now recall that a d-dimensional random variable X is reflectively symmetric about μ if the random vectors $X - \mu$ and $\mu - X$ have the same distribution. More precisely, X is reflectively symmetric about μ if and only if

$$E\{\sin[t'(X-\mu)]\} = \int_{\mathbb{R}^d} \sin[t'(x-\mu)]\mathrm{d}P_\mu(x) = 0, \quad t \in \mathbb{R}^d. \tag{1}$$

Denote the characteristic function of X by

$$\varphi(t) = E(\mathrm{e}^{(\mathrm{i}t'X)}) = E([\cos(t'X)]) + \mathrm{i}E([\sin(t'X)]) = \varphi_1(t) + \mathrm{i}\varphi_2(t), \quad t \in \mathbb{R}^d. \tag{2}$$

Recall that, $i^2 = -1$, and that for all $t \in \mathbb{R}^d$, $|\varphi(t)|^2 = \varphi(t)\overline{\varphi(t)} = \varphi_1^2(t) + \varphi_2^2(t)$. Let Ξ be a compact subset of $\mathbb{R}^d$ on which $\varphi_1(t)$ (hence $|\varphi(t)|$) does not vanish. A such compact subset can be for example the closed ball of radius $0 < r < 1$, $B(\mathbf{0}, r) = \{u \in \mathbb{R}^d : \|u\| \leq r\}$. Let $\mathcal{C}[\Xi \to \mathbb{R}^d]$ be the set of continuous functions defined on Ξ with values in $\mathbb{R}^d$ and $\mathcal{C}(\Xi)$ the set of all real-valued continuous functions defined on Ξ, with the usual sup-norm $\|x\|_\infty = \sup_{t\in\Xi} |x(t)|$.

Let $X_1, \ldots, X_n$ be n independent copies of X. Define the following d-dimensional random and non-random functions

$$\beta_n(t) = \frac{\sum_{k=1}^{n}\sum_{j=1}^{n} X_k \cos[t'(X_k - X_j)]}{\sum_{k=1}^{n}\sum_{j=1}^{n} \cos[t'(X_k - X_j)]}, \quad t \in \Xi \tag{3}$$