Journal of Multivariate Analysis 108 (2012) 67–89

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

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

Characteristic function-based hypothesis tests under weak dependence

Anne Leucht

Friedrich-Schiller-Universität Jena, Institut für Stochastik, Ernst-Abbe-Platz 2, D-07743 Jena, Germany

ARTICLE INFO

Article history:
Received 17 May 2011
Available online 18 February 2012

AMS 2010 subject classifications:
62G10
62E20
62F40

Keywords:
Bootstrap
Empirical characteristic function
Goodness-of-fit
Symmetry test
Time series
V-statistics

ABSTRACT

In this article we propose two consistent hypothesis tests of L_2-type for weakly dependent observations based on the empirical characteristic function. We consider a symmetry test and a goodness-of-fit test for the marginal distribution of a time series. The asymptotic behaviour under the null as well as fixed and certain local alternatives is investigated. Since the limit distributions of the test statistics depend on unknown parameters in a complicated way, we suggest to apply certain parametric bootstrap methods in order to determine critical values of the tests.

1. Introduction and motivation

Time series models are well-established tools to describe various real-life phenomena, e.g. in finance and biometrics. In the present paper we provide consistent testing procedures that allow for checking adequacy of a certain model with respect to symmetry and parametric structure of the marginal distribution of a time series.

Let $X_1, \ldots, X_n$ be $\mathbb{R}^d$-valued observations of a stationary weakly dependent process with marginal distribution P_X. We consider problems of the following structure:

$$\mathcal{H}_0 : g_X(\cdot) = f_X(\cdot, \theta_0) \quad \text{for some } \theta_0 \in \Theta \subseteq \mathbb{R}^p \text{ against}$$
$$\mathcal{H}_1 : g_X(\cdot) \neq f_X(\cdot, \theta) \quad \forall \theta \in \Theta.$$

Besides supremum-type tests, L_2-tests are the most convenient ones in mathematical statistics. We apply a test statistic of L_2-type, $\widehat{T}_n = n \int_{\mathbb{R}^d} |\psi(c_n)(t) - f_X(t, \widehat{\theta}_n)|^2 w(t)dt$. Here, $\widehat{\theta}_n$ is a consistent estimator of θ_0 and c_n denotes the empirical characteristic function, i.e. $c_n(t) = n^{-1} \sum_{k=1}^n e^{it'X_k}$. The function ψ is chosen such that $\psi(c_n)$ forms an appropriate nonparametric estimate of $g_X = \psi(c_X)$, where c_X is the characteristic function of X_1. We specify the functions ψ and f_X for the concrete test problems in Sections 2 and 3. There are several approaches in order to derive the asymptotic null distribution of $\widehat{T}_n$. Under certain regularity conditions the latter statistic can be approximated by an L_2-statistic $T_n = \int_{\mathbb{R}^d} |n^{-1/2} \sum_{k=1}^n h(X_k, t, \theta_0)|^2 w(t)dt$ with a certain continuous function h and with fixed parameter θ_0. The limit of latter quantity can be deduced invoking empirical process theory. Alternatively, $n^{-1/2} \sum_{k=1}^n h(X_k, t, \theta_0),\ t \in \mathbb{R}^d$, can be viewed as an element of the Hilbert space $L_2(\mathbb{R}^d, \mathcal{B}^d, w(t)dt)$. Therefore, the asymptotics of T_n can also be obtained invoking a

E-mail address: anne.leucht@uni-jena.de.

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2012.02.003

central limit theorem (CLT) for random variables with values in a Hilbert space and the continuous mapping theorem. In contrast, the approximating statistic can also be understood as a degenerate V-statistic. In the present paper, we employ a recent result on degenerate V-statistics by Leucht [25] that allows for the derivation of the limit distributions under easily verifiable assumptions. In order to study the behaviour of our test statistics under local alternatives, this approach is carried over to a special kind of triangular schemes of random variables; cf. Section 4.2.

In Section 2, we extend the symmetry test which was initially proposed by Feuerverger and Mureika [17] for independent and identically distributed (i.i.d.) data and a known centre of symmetry to the case of weakly dependent observations with unknown location parameter. Answering the question whether a distribution is symmetric or not is interesting for several reasons. First, there is a pure statistical interest since presence or absence of symmetry is important for deciding what parameter to estimate. If the underlying distribution is symmetric, the point of symmetry is the only reasonable location parameter, whereas in the non-symmetric case there is no longer only one measure of location; cf. [1]. Moreover, robust estimators of location, e.g. trimmed means, and robust tests for location parameters assume the underlying observations to arise from a symmetric distribution; see for instance [34]. Consequently it is important to check this assumption before applying those methods. Rejecting the hypothesis of symmetry also has substantial impact on model selection. In this case, fitting data to nonlinear AR(p) processes with skew symmetric regression functions and symmetric innovations is inappropriate; cf. [30]. Finally, symmetry plays a central role in analysing and modelling business circles since time reversibility of a process $(X_t)_{t\in\mathbb{Z}}$, i.e. $P_{X_{t_1},\ldots,X_{t_k}} = P_{X_{t_k},\ldots,X_{t_1}},\ t_1,\ldots,t_k \in \mathbb{Z}, k \in \mathbb{N}$, implies symmetry of the distributions of the differences $Y_{t,k} = X_t - X_{t-k},\ k \in \mathbb{N}$, about the origin. For a detailed discussion of this topic see [31,10].

A great variety of symmetry tests for i.i.d. random variables are available in the literature. Detailed lists of references are given by Lee [24] as well as by Henze et al. [20]. Also in the context of time series numerous tests of symmetry have been employed. There are several moment-based tests, e.g. by Ramsey and Rothman [31] or Bai and Ng [3]. However, both approaches are inconsistent against alternatives whose third moments vanish but still are not symmetric. Lee [24] built an M test for the composite hypothesis of symmetry around an unknown location parameter and derived asymptotic normality under the null and certain local alternatives. However, a consistency result is not stated. Fan and Ullah [16] considered a consistent L_2-test for the simple hypothesis based on kernel density estimates with vanishing bandwidth under the assumption that the sample arises from a certain absolute regular process. The test for the simple hypothesis by Chen et al. [10] employs the fact that a distribution is symmetric if and only if the imaginary part of its characteristic function $\Im c(t) = \mathbb{E}\sin(t'X)$ vanishes. No moment restrictions on the marginal distributions of the underlying data are required. They investigated the asymptotics under the null and certain local alternatives but they did not derive consistency of their test. We also establish a characteristic function-based test for symmetry that is asymptotically unbiased against Pitman local alternatives. In contrast to [10], we consider the composite hypothesis of symmetry around an unknown centre and obtain consistency of our approach.

In Section 3, a goodness-of-fit test for the marginal distribution of a time series is constructed. So far, the normality assumption is still dominating the literature in many fields, however, not for reasons of empirical evidence but for theoretical simplicity. Within the last decades the literature on non-Gaussian time series has addressed the topic of finding distributions of the innovations corresponding to specified marginals of a certain model. Surveys of those results are given by Jose et al. [23] as well as by Block et al. [7].

While there is a great variety of tests concerning the problem whether the distribution of a sample belongs to a parametric class of distributions if the underlying observations are i.i.d., the number of consistent tests is limited in the dependent case. Recently, Ignaccolo [21] and Munk et al. [26] developed results for α-mixing variables generalizing Neyman's smooth test for a simple null hypothesis. Tests for the composite hypothesis based on the L_2-difference between a smoothed version of the parametric density estimate and a nonparametric estimator were considered by Fan and Ullah [16] as well as by Neumann and Paparoditis [28]. The bandwidths involved in the estimators were supposed to be asymptotically vanishing. A drawback resulting from the latter assumption is the loss of power against Pitman local alternatives compared to approaches with fixed bandwidths; see [18]. Motivated by the fact that tests based on kernel density estimators with decreasing smoothing parameter are very sensitive to the choice of the bandwidth, Fan [15] established a characteristic-function based test for the composite hypothesis in the i.i.d. case that evaluates the difference between the characteristic function and its fixed-kernel estimate, the empirical characteristic function. To the author's best knowledge this approach has not been considered if the observations are dependent. Here, we extend the L_2-test of [15] to the time series setting. This type of test is useful in particular when the density has a complicated structure while the characteristic function is of simple form. A typical example of such a distribution is the normal inverse Gaussian distribution that is widely used in mathematical finance, see e.g. [4].

For the above-named tests, we derive their limit distributions under the respective null hypotheses and under certain local alternatives. It turns out that already in the case of i.i.d. observations the asymptotic distributions depend on unknown parameters in a complicated way. The bootstrap offers a convenient way to determine critical values of the tests. However, a naive application of these resampling methods fails. We obtain bootstrap consistency after a modification of the test statistics on the bootstrap side. It is well-known that model-based bootstrap counterparts of the underlying observations can often not proved to be mixing even though the original process satisfies some mixing condition. However, other concepts of weak dependence can be verified, see e.g. [5] or [14]. Throughout the paper, the underlying process is assumed to meet the following definition, which is due to [12].