Journal of Statistical Planning and Inference 146 (2014) 122–132

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

The probability weighted characteristic function and goodness-of-fit testing

Simos G. Meintanis [a,b,*,1], Jan Swanepoel [b], James Allison [b]

[a] Department of Economics, National and Kapodistrian University of Athens, Athens, Greece
[b] Unit for Business Mathematics and Informatics, North-West University, Potchefstroom, South Africa

ARTICLE INFO

Article history:
Received 29 May 2013
Received in revised form
18 September 2013
Accepted 24 September 2013
Available online 30 September 2013

Keywords:
Goodness-of-fit test
Characteristic function
Empirical characteristic function
Exponential distribution
Normal distribution

ABSTRACT

We introduce the notion of the probability weighted characteristic function (PWCF) as a generalization of the characteristic function. Some of the properties of the PWCF and of its empirical counterpart are studied and the potential use of these quantities in goodness-of-fit testing is examined in detail. The corresponding limiting null distributions and consistency results for location-scale models are derived and finite-sample comparisons are presented. Also, the notion of the PWCF is extended to arbitrary dimension.

1. Introduction

For the past four decades, the empirical characteristic function (ECF) has been a valuable tool for statistical inference. A list of recent publications in estimation, goodness-of-fit testing and other areas of application includes Ghosh (2013), Balakrishnan et al. (2013), Meintanis and Hlávka (2010), Tenreiro (2011, 2009, 2007), Ngatchou-Wandji (2009a,b), Jiménez-Gamero et al. (2009), Matsui and Takemura (2008, 2005), Beran and Ghosh (2006), Epps (2005), Besbeas and Morgan (2004, 2001), Henze et al. (2003), and Gürtler and Henze (2000), among others. For a review of the early work based on the ECF the reader is referred to Csörgő (1984), while more recent synopses are included in Hušková and Meintanis (2008a,b) and Braun et al. (2008). The aim of the current work is to introduce a generalization of the ECF and to explore some of its potential applications. In this connection, and in order to motivate the discussion, note that the characteristic function (CF) and the ECF can be used to derive the theoretical and empirical moments, respectively, of the underlying distribution. Recently however there is a line of research in which theoretical and sample moments are replaced by weighted counterparts, the so-called probability weighted moments; see for instance Furrer and Naveau (2007), Diebolt et al. (2007) and Diebolt et al. (2008). The probability weighted CF (PWCF) and its empirical counterpart, the probability weighted ECF (PWECF), come as generalizations of the corresponding notions of the CF and the ECF, where the weights are generalized versions of those used to compute the empirical probability weighted moments.

* Corresponding author at: Department of Economics, National and Kapodistrian University of Athens, Athens, Greece. Tel.: +30 210 3689814.
E-mail address: simosmei@econ.uoa.gr (S.G. Meintanis).
[1] On sabbatical leave from the University of Athens.

0378-3758/$ - see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jspi.2013.09.011

To be specific let $X_1,\ldots,X_n$ be i.i.d. observations on an arbitrary random variable X, with strictly increasing continuous distribution function (DF), $F(x)=\mathrm{P}(X\le x)$ and recall that the CF of X is defined by $\varphi(t)=\mathbb{E}(e^{itX})$. In turn, the ECF is defined as

$$\varphi_n(t)=\frac{1}{n}\sum_{j=1}^{n}\exp(itX_j). \tag{1.1}$$

In the past, standard methods of estimation and testing via the ECF have utilized the L2-type distance

$$\int_{-\infty}^{\infty}|\varphi_n(t)-\varphi(t)|^2 w_\lambda(t)\,dt, \tag{1.2}$$

which apart from the CF and the ECF employs a parametric weight function $w_\lambda(\cdot)$ indexed by the parameter λ and satisfying some integrability conditions. There has been considerable discussion on $w_\lambda(\cdot)$. In the literature one can find $w_\lambda(t)=e^{-\lambda|t|^\beta}$, $\beta=1,2$, as standard choices. Both values $\beta=2$ and $\beta=1$ correspond to kernel-based choices, with the weight function being a multiple in the first case (resp. second case) of the standard normal density (resp. standard Laplace density) as kernel with bandwidth equal to $1/(\lambda\sqrt{2})$ (resp. $1/\lambda$). In fact Tenreiro (2007) made the connection of a test statistic as in (1.2) with an affine invariant version of the celebrated Bickel–Rosenblatt test, where the weight function $w_\lambda(t)$ is chosen on the basis of the kernel employed in estimating the underlying density. In a follow-up paper Tenreiro (2009) studied the specific case of testing normality and takes the zero-mean normal density as a weight function. The author suggested criteria for choosing λ on the basis of approximate Bahadur slopes and specific deviations from the null hypothesis. Along similar lines Epps (2005) made a connection between the test statistic in (1.2) with $w_\lambda(t)=|\varphi(t)|^2$ and the Anderson–Darling test. However, the proper choice of the weight function (both in terms of the function itself as well as the choice of the value of λ) is a difficult problem which affords a solution only in the case of highly structured models, and even then this solution is computationally demanding. One aim of this work is to alleviate this problem by introducing an alternative to the CF. In this alternative we suggest a statistically meaningful way of choosing the weight function in L2-type procedures thereby reducing the aforementioned problem to one of only choosing the value of the weight parameter λ. Specifically for $\lambda\ge 0$, we suggest the PWCF defined as

$$\chi(t;\lambda):=\mathbb{E}[W(X;\lambda t)e^{itX}]=\int_{-\infty}^{\infty}W(x;\lambda t)e^{itx}\,dF(x), \tag{1.3}$$

where the probability weight is given by

$$W(x;\beta)=[F(x)(1-F(x))]^{|\beta|},\quad \beta\in\mathbb{R},\ x\in\mathbb{R}. \tag{1.4}$$

Throughout the paper we will assume that F belongs to a parametric family of distributions

$$\mathcal{F}_\Theta=\{F_\vartheta:\vartheta\in\Theta\},$$

where Θ is an open subset of $\mathbb{R}^p$, $p\ge 1$. The reason for this assumption is that we will be using the PWCF exclusively in a goodness-of-fit type setting. Some characterizing properties of the PWCF are discussed in Section 3 below which will enable one to prove the consistency of tests based on the PWECF against fixed alternatives.

The PWECF is defined formally as

$$\chi_n(t;\lambda):=\frac{1}{n}\sum_{j=1}^{n}\widehat{W}(X_j;\lambda t)\exp(itX_j),\quad t\in\mathbb{R}, \tag{1.5}$$

where the estimated probability weight is given by

$$\widehat{W}(x;\beta)=[F_{\hat\vartheta_n}(x)(1-F_{\hat\vartheta_n}(x))]^{|\beta|},\quad \beta\in\mathbb{R},\ x\in\mathbb{R}, \tag{1.6}$$

and $\hat\vartheta_n:=\hat\vartheta_n(X_1,\ldots,X_n)$ is a consistent estimator of ϑ. Clearly, under some mild conditions placed on F and $\hat\vartheta_n$, by the Law of Large Numbers and for fixed t we have,

$$\chi_n(t;\lambda)\xrightarrow{\mathrm{P}}\chi(t;\lambda).$$

The fact that some type of weighting may lead to improvement of CF-based procedures goes back to Markatou et al. (1995), where these authors use $W(x;\beta):=W(x;\beta_n)=\mathbf{1}_{\{|x|\le\beta_n\}}$ for an appropriately chosen constant $\beta_n>0$, satisfying $\lim_{n\to\infty}\beta_n=\infty$. One could think of more than one ways for weighting the ECF, some based on the density, others on the DF, or a combination thereof. A density-based procedure would perhaps have the advantage of putting more weight at the most frequently occurring observations. On the other hand a DF-based weighting, such as that in (1.4), puts more weight at the center of the distribution. It is often the case that this center carries high frequency, but it is also possible that under this scheme few observations will receive maximum weight, for example, when sampling from certain multimodal densities. Nevertheless, we opted to base our weighting on the DF as this quantity most accurately reflects the tail properties which are of course one of the major features of an underlying distribution. In particular, the weight in (1.4) implies that for fixed $\beta\in\mathbb{R}$, $W(x;\beta)$ puts most of the weight near the median of F, at which point receives maximum weight of size $(1/4)^{|\beta|}$. As we move away from the center, either toward the left or toward the right tail of F, the corresponding weights decrease progressively, and approach zero as X reaches the left or the right boundary of the domain of definition of F. Also for fixed x maximum weight is assigned at $\beta=0$, while with increasing $|\beta|$, $W(x;\beta)$ decreases and tends to zero as $|\beta|\to\infty$. Therefore the type of weighting