A REVIEW OF TESTING PROCEDURES BASED ON THE EMPIRICAL CHARACTERISTIC FUNCTION

Simos G. Meintanis [1]
National and Kapodistrian University of Athens, Department of Economics, 1 Sophocleous Str., and Aristidou Str., 105 59 Athens, Greece;
Unit for Business Mathematics and Informatics, North-West University, Potchefstroom, South Africa
e-mail: *simosmei@econ.uoa.gr*

Key words: Dependent data; Goodness-of-fit; Regression data; Two-sample problem.

Abstract: The empirical characteristic function (ECF) has been in use in statistical inference for nearly fifty years now. We provide an overview of testing procedures based on the ECF within certain statistical models. Specifically our emphasis is on recent developments of ECF procedures for goodness-of-fit testing for parametric families of distributions with structured data, and for the two-sample and the *k*-sample problem.

1. Introduction

Let F denote a distribution function (DF) of a random variable $Y \in \mathbb{R}^d$, $d \geq 1$, and consider the specific parametric class $\mathscr{F}_\Theta = \{\mathscr{F}_\vartheta, \vartheta \in \Theta\}$, of distributions indexed by $\vartheta \in \Theta$, where Θ is an open subset of the Euclidean space of arbitrary dimension $p \geq 1$. A standard statistical decision problem is to test the goodness-of-fit (GOF) null hypothesis,

$$\mathscr{H}_{01} : F \in \mathscr{F}_\Theta, \text{ for some } \vartheta \in \Theta. \tag{1}$$

Another typical problem of interest is testing for the two-sample problem

$$\mathscr{H}_{02} : F_1 \equiv F_2, \tag{2}$$

where F_1 and F_2, denote a couple of unspecified DF's.

These hypotheses have been approached by methods based on the empirical characteristic function (ECF) defined by

$$\varphi_n(t) = \int_{-\infty}^{\infty} e^{iyt} dF_n(y), \tag{3}$$

[1] On sabbatical leave from the University of Athens.

AMS: 62F03, 62F05, 62G10, 62G20, 62H15

where F_n denotes the empirical DF computed from independent copies $Y_1,...,Y_n$ of Y.

The main idea behind the methods utilizing the ECF is that the null hypotheses in (1)-(2) may also be expressed in terms of the characteristic function (CF). For instance, by writing $\varphi(t)$ for the CF corresponding to F, the null hypothesis in (1) may equivalently be written as

$$\mathscr{H}_{01} : \varphi \equiv \varphi_\vartheta, \text{ for some } \vartheta \in \Theta,$$

where $\varphi_\vartheta(t)$ denotes a CF in $\mathscr{F}_\Theta$. Likewise in the Fourier domain the null hypothesis (2) is stated as

$$\mathscr{H}_{02} : \varphi_1 \equiv \varphi_2,$$

where φ_k, stands for the CF corresponding to F_k, $k = 1, 2$.

Then as it is typical in GOF testing, we compare a nonparametric estimator of the CF, which in this review is the ECF $\varphi_n(t)$, with a corresponding parametric estimator of the same quantity reflecting the null hypothesis, which will be denoted by $\varphi_{n0}(t)$. We reject the null hypothesis (1) for large values of a test statistic of the type

$$\Delta_n = \Delta_w(\varphi_n, \varphi_{n0}), \tag{4}$$

where $\Delta_w(\cdot,\cdot)$, denotes some distance function involving a weight function $w(t)$, the role of which will be discussed below. Analogous is the approach for the null hypothesis in (2).

Historically the definition of the ECF first appears in the famous textbook of Cramér (1946), page 342. Several years later the ECF is used by Parzen (1962) as an auxiliary tool in kernel-type density estimation. In each own right however the ECF first appears in the works of Press (1972) for estimation of parameters of stable distributions and Heathcote (1972) for GOF testing, while later work has established the ECF as a main tool for many inferential procedures. Earlier reviews on the subject of testing by the ECF are included in Csörgő (1984), Hušková and Meintanis (2008a, 2008b), and in chapter 3 of Ushakov (1999). In this synopsis we concentrate on recent developments of ECF procedures for GOF testing and for the two-sample problem. Specifically in Section 2, we review GOF tests under various regression settings, while in Section 3 corresponding ECF methods are surveyed which apply with vectorial observations. Section 4 is devoted to approximations of the resulting limit distributions of the ECF test statistics and to the choice of the weight function, while the extension of the methods to data involving dependence is considered in Section 5. Section 6 reviews two-sample and many-sample methods, both in the univariate and multivariate setting, and we conclude with Section 7.

2. GOF tests for parametric families of distributions

In the classical case of independent and identically distributed (i.i.d.) observations, the typical ECF approach in testing (1) is the so-called L2 approach which utilizes the distance

$$\Delta_w(f,g) = \int |g_1(t) - g_2(t)|^2 w(t)dt \tag{5}$$

where the weight function w is assumed to satisfy $w(t) = w(-t),\ t \in \mathbb{R}$. Then the test statistic is defined by $\Delta_{n,w} = n\Delta_w(\varphi_n, \varphi_{n0})$, where φ_n is defined in (3), φ_{n0} is replaced by the estimated CF $\varphi_{\widehat{\vartheta}_n}$

under the null hypothesis (1), and $\widehat{\vartheta}_n$ denotes a consistent estimator of the parameter ϑ. The L2 approach was initiated by Feuerverger and Mureika (1977) for testing symmetry, and by Epps and Pulley (1983, 1986), for testing normality and exponentiality, respectively, and was subsequently followed by many researchers; for more information the reader is referred to the aforementioned earlier reviews. In this connection we also refer to Meintanis and Swanepoel (2007) that present an overview of L2 methods for the null hypothesis in (1) including corresponding asymptotic results and suggestions for a parametric bootstrap resampling procedure in order to circumvent the problem of a highly non trivial asymptotic null distribution of the test statistic.

Another recent development is related to the resampling procedure suggested by Meintanis and Swanepoel (2007). Specifically, Jiménez-Gamero and Kim (2015) suggest to replace the parametric bootstrap procedure by a weighted bootstrap procedure, originally found in Kojadinovic and Yan (2012). The weighted bootstrap, unlike the parametric bootstrap, requires preliminary estimating computations which are not trivial and which are different for different DF's under the null hypothesis (1). On the other hand, these computations are carried out only once with every Monte Carlo whereas with the parametric bootstrap, estimation is performed with each bootstrap sample. Hence the weighted bootstrap is faster to apply and, as reported in Jiménez-Gamero and Kim (2015) the loss of power compared to the parametric bootstrap is small, especially for large sample size n.

In this section we will review certain extensions of the above GOF procedure starting with the case of linear regression. Specifically consider the classical linear model

$$Y = x^T\beta + c\varepsilon, \tag{6}$$

where $x = (1, x_2, ..., x_p)^T \in \mathbb{R}^p$, is a known regressor vector, $\beta \in \mathbb{R}^p$ and $c > 0$ denote unspecified regression and scale parameters, respectively, and ε is the error which has F and φ as its DF and CF, respectively. We wish to test the null hypothesis $\mathscr{H}_{01}$ in (1). In the regression setting however we will further assume that the DF under the null hypothesis $\mathscr{H}_{01}$ is symmetric around zero with unit scale. Moreover we assume that no extra parameters, apart from location and scale, exist in the parametric family $\mathscr{F}_\Theta$, and we will write F_0 and φ_0 for the corresponding DF and CF under the null hypothesis $\mathscr{H}_{01}$ in (1), respectively.

Hušková and Meintanis (2007) and Jiménez-Gamero, Muñoz-García and Pino-Mejías (2005) consider this GOF problem for the errors in the regression setting (6). In particular they apply the test statistic in (5) with g_2 replaced by φ_0 and g_1 replaced by the ECF of the residuals

$$\widehat{\varphi}_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{it(Y_j - x_j^T\widehat{\beta}_n)/\widehat{c}_n}, \ j = 1, ..., n,$$

obtained from a regression fit based on the data $\{x_j, Y_j\}_{j=1}^n$. Here we shall follow the exposition in Hušková and Meintanis (2007). First it is assumed that the estimator $\widehat{\beta}_n := \widehat{\beta}_n\left(\{x_j, Y_j\}_{j=1}^n\right)$ of β, considered is regression equivariant, i.e.,

$$\widehat{\beta}_n\left(\{x_j, Y_j + x_j^T v\}_{j=1}^n\right) = \widehat{\beta}_n\left(\{x_j, Y_j\}_{j=1}^n\right) + v, \tag{7}$$

for each $v \in \mathbb{R}^p$, and that the estimator $\widehat{c}_n = \widehat{c}_n\left(\{x_j, Y_j\}_{j=1}^n\right)$, of c is scale equivariant, i.e.,

$$\widehat{c}_n\left(\{x_j, bY_j\}_{j=1}^n\right) = b\,\widehat{c}_n\left(\{x_j, Y_j\}_{j=1}^n\right), \tag{8}$$

for each $b > 0$.

If the regression estimators satisfy (7) and (8) and under some additional assumptions, it is proved that under the null hypothesis

$$\Delta_{n,w} \xrightarrow{\mathscr{D}} \int \mathscr{Z}^2(t)w(t)dt, \tag{9}$$

where $\{\mathscr{Z}(t),\ t \in \mathbb{R}\}$ is a zero-mean Gaussian process with covariance kernel depending on F_0 and on the type of estimation of the parameters (β, c) but not on the true values of these parameters.

A study the behaviour of $\Delta_{n,w}$ under contiguous alternatives is also carried out by Hušková and Meintanis (2007). Specifically the density of the random errors under this type of alternatives is written as:

$$g_n(x) = (1 + \frac{\kappa}{\sqrt{n}} u(x)) f_0(x), \quad x \in \mathbb{R}, \tag{10}$$

where $f_0(\cdot)$ is the density corresponding to $F_0(\cdot)$, $\kappa \neq 0$, and $u(\cdot)$ is a suitable measurable function. Then it is shown that

$$\Delta_{n,w} \xrightarrow{\mathscr{D}} \int (\mathscr{Z}(t) + \kappa\mu(t))^2 w(t)dt,$$

where $\int \mu^2(t)w(t) \neq 0$. Hence the test procedure is consistent as soon as $\kappa \to \infty$.

The problem of testing (1) for the distribution of the error ε was considered in Hušková and Meintanis (2010) also in the context of the nonparametric regression model

$$Y = m(X) + \sigma(X)\varepsilon, \tag{11}$$

where $m(\cdot)$ and $\sigma(\cdot)$ denote unspecified regression and scale functions, respectively. This test uses the approach in (5) with g_1 replaced by the ECF

$$\widehat{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^{n} e^{it(Y_j - \widehat{m}(X_j))/\widehat{\sigma}(X_j)}, \tag{12}$$

of the nonparametrically computed residuals and g_2 replaced by the CF $\varphi_{\widehat{\vartheta}_n}$ under the null hypothesis. We will not present any technical results here but simply mention that as in the linear regression setting, the limit null distribution of the test statistic depends on the law of the error under the null hypothesis. However it is shown that this distribution does not depend on the density of the regressor X, or the nonparametric functions $m(\cdot)$ and $\sigma(\cdot)$, and even not on the kernel and the bandwidth used in estimating these functions. This result is somewhat surprising but nevertheless in line with the results of Neumeyer, Dette and Nagel (2006) for corresponding classical GOF tests based on the empirical DF. Finally we mention an ECF test for testing GOF of the regression function in (11) developed by Hušková and Meintanis (2009). This test uses the approach in (5) with the ECF as in (12) and g_2 replaced by the ECF of the residuals computed under the null hypothesis $m(X) = m_{\vartheta}(X)$.

3. Multivariate extensions

Versions of the distance measure in (5) for the null hypothesis (1) in the multivariate context ($d > 1$) were considered in Henze and Wagner (1997), Jiménez-Gamero, Alba-Fernández, Muñoz-García and Chalco-Cano (2009) and Meintanis, Allison and Santana (2016), always with special reference to testing for multivariate normality. On the other hand, Arcones (2007) uses certain characterizations of the normal distribution in constructing an L2-type distance analogous to (5). The characterization approach is also followed by Meintanis, Ngatchou-Wandji and Taufer (2015), in order to generalize the L2-type test of Henze and Wagner (1997) from a test of normality to a test for an arbitrary multivariate stable distribution.

On the other hand, Pudelko (2005) employs the supremum (Kolmogorov-Smirnov type) distance

$$\sup_{\mathbf{t}} \|\varphi_n(\mathbf{t}) - \varphi_{\widehat{\vartheta}_n}(\mathbf{t})\| \tag{13}$$

initially suggested by Csörgő (1986) for testing multivariate normality. A supremum distance analogous to (13) is also used by Fang, Li and Liang (1998), with the related notion of the empirical moment generating function.

In the context of testing for the error distribution in linear regression, the corresponding multivariate case was treated by Jiménez-Gamero et al. (2005) by extending the test for normality proposed by Henze and Wagner (1997). The results obtained are similar to those listed in Section 2, since the dimension carries no specific significance in the proofs of Hušková and Meintanis (2007). An exception is that Jiménez-Gamero et al. (2005) also allow the number of regressors p to increase with the sample size at a rate which satisfies $p^2/n = o(1)$.

4. Limit distribution and the weight function

Meintanis and Swanepoel (2007) show that in the i.i.d. case the test statistic

$$\Delta_{n,w} = n\Delta_w(\varphi_n, \varphi_{\widehat{\vartheta}_n}), \tag{14}$$

for the null hypothesis $\mathscr{H}_{01}$ has the same limit null distribution as the random variable

$$D_w = \sum_{m=1}^{\infty} \lambda_m^{(w)} Z_m^2, \tag{15}$$

where Z_m are i.i.d. standard normally distributed, and the coefficients (or eigenvalues) $\lambda_m = \lambda_m^{(w)}(\vartheta_0, \mathscr{F}_\Theta)$, $m \geq 1$, depend on the weight function $w(t)$ as well as on the family being tested, on the the true value ϑ_0 of the parameter ϑ, but also on the type of estimation used in estimating this parameter.

Numerical approximation of the distribution of D_w in (15) has been attempted by Matsui and Takemura (2008) for the special case of testing GOF to symmetric stable distributions. Even in this special case this is a highly non-trivial numerical task since the coefficients λ_m can only be obtained by solving a complicated integral equation. As expected, analytic solution of this equation, and hence analytic approximation of the distribution of D_w, is in most cases not possible. In fact, the sole exception of an analytically derived approximation of the limit distribution of an L2-type ECF statistic in the multivariate context is for the asymptotic null distribution of the test statistic for multivariate normality of Henze and Wagner (1997) with fixed location and covariance matrix. (Note that this test was initially suggested by Baringhaus and Henze, 1988 and Henze and Zirkler, 1990). This approximation has been derived by Baringhaus (1996) and includes an analytic expression for the eigenvectors.

In order to circumvent the problem of a non-standard asymptotic null distribution, Naito (1996a, 1997) expands on an earlier suggestion of Ahmad (1993) to employ a weighted version of the empirical DF in the Cramér-von Mises statistic, and suggests a modification of the Epps and Pulley (1983) test statistic that is asymptotically normally distributed; see also Wong and Sim (2000).

Another important issue with ECF-based test statistics is the choice of the weight function $w(t)$ figuring in (5). Important aspects for this choice is computational simplicity and level/power properties of the resulting tests. As far as the first aspect is concerned there seems to be a strong bias towards exponential weight functions of the type

$$w_a(t) = e^{-a|t|^b}, \; a > 0, \; b = 1, 2, \tag{16}$$

and their corresponding multivariate versions. In this connection Epps (2005) notes that tests for normality have mostly been implemented with $b = 2$, while for testing for the Cauchy distribution $b = 1$ is utilized in $w_a(t)$ of (16). Clearly this amounts to taking (a specific instance of) the respective population CF as weight function in each case. Epps (2005) provides further grounds and generalizes this type of choice which in many cases leads to a closed formula for the resulting test statistic, while Jiménez-Gamero et al. (2009) carry this approach of hypothesis-specific weight function to the multivariate context.

Besides simplicity, the preference for weight functions such as $w_a(t)$ is grounded on a standard practice of earlier workers whereby ECF procedures are confined to a relatively narrow interval around zero for t. This practice partly rests on the well known fact that tail properties of any distribution are reflected on the behaviour of its CF around the origin. Research papers in this direction include Welsh (1986), Csörgő and Heathcote (1987), and Pourahmadi (1987). Specifically these authors argue that ECF-based inference should not be extended beyond values of the argument t greater than the first root of the real part of the ECF, a suggestion that triggered research towards the stochastic properties of this root; see for example Bräker and Hüsler (1991).

Further study on $w(t)$ was facilitated by the the connection made by Henze and Zirkler (1990) that the choice of $w(t)$ is directly related to the choice of the kernel in L2-type GOF tests of Bowman and Foster (1993) utilizing kernel-based density estimators. This line of research is carried further by Lindsay, Markatou and Ray (2014) in a more general framework whereby distance-based GOF tests are expressed by means of kernels, with ECF tests being a special case. The authors essentially suggest a calculus for kernel/bandwidth choice directed at specific alternatives. The kernel/bandwidth role is also analysed in Sejdinovic, Sriperumbudur, Gretton and Fukumizu (2013) where it shown that distance-based statistics such as those for the two-sample problem in (18) are equivalent with certain statistics encountered in the machine learning literature. In this connection and in the context of parametric weight functions defined in (16), the value of the parameter a is particularly important as it directly relates to the value of the bandwidth in nonparametric density estimation. Hence further work in this line of research has been carried out in order to investigate test-efficiency as a function of a within specific parametric forms such as $w_a(t)$; for more details the reader is referred to Naito (1996b), Epps (1999) and Tenreiro (2009, 2011). Finally, there exists a moment-based interpretation concerning the weight function which is considered in Section 6.

5. Testing with dependent observations

In this section we provide a summary of ECF-based GOF tests under dependence. We begin with tests within a parametric context. Specifically suppose that observations Y_t come from a GARCH model

$$\mathrm{E}(Y_t|Y_s, s<t)=0,\ \ \sigma_t^2=\mathrm{Var}(Y_t|Y_s, s<t)=\beta_0+\sum_{j=1}^{k}\beta_j Y_{t-j}^2+\sum_{j=1}^{\kappa}\gamma_j\sigma_{t-j}^2, \tag{17}$$

where β_j $(0\le j\le k)$ and $\gamma_j(1\le j\le\kappa)$ denote unknown parameters. We are interested in testing $\mathscr{H}_{01}$ of (1) with F standing for the DF of the innovations $\varepsilon_t=Y_t/\sigma_t$ in (17). Jiménez-Gamero (2014) follows the L2 approach in (5) and essentially suggests the test statistic (14) where

$$\widehat{\varphi}_n(t)=\frac{1}{n}\sum_{t=1}^{n}e^{itY_t/\widehat{\sigma}_t},$$

where $\widehat{\sigma}_t$ is a suitable estimator of the volatility parameter. The asymptotic null distribution of the test statistic is derived and a resampling algorithm is suggested by means of which this distribution may be consistently approximating via the bootstrap. It should be mentioned here that the GOF tests proposed by Jiménez-Gamero (2014) were first suggested by Klar, Lindner and Meintanis (2012) who also provided an extensive study of their small sample properties.

More general frameworks for dependence are those of weak dependence and long-range dependence. Leucht (2012) extends the test statistic in (14) from i.i.d. observations to weakly dependent observations. Specifically it is shown that the limit null distribution exists but as in the i.i.d. case is highly non-trivial. However, unlike the case with i.i.d. data, the bootstrap statistic in this context does not mimic the original test statistic, but has to be suitably modified in order to become consistent. On the other hand, the case of long-range dependence is entirely different as far as the limit null distribution is concerned: L2-type statistics for normality testing as in (14) attain trivial limit null distributions. This somewhat surprising fact is proved in Ghosh (2013) where similar test statistics involving the empirical moment generating function are shown to converge in distribution to the χ_1^2 distribution, under the null hypothesis of normality, if the data are long-range dependent.

The favourable feature of a simple limit null distribution seems to be the main reason that Nieto-Reyes, Cuesta-Albertos and Gamboa (2014) deviate from the L2 formulation of (5). These authors suggest a test for normality for dependent data that relies on an earlier χ^2-type method suggested by Epps (1987). Specifically for fixed integer $M>0$, the test is based on a quadratic form of the type $\sum_{m,\ell=1}^{M}v_{m,\ell}\delta_m\delta_\ell$ in which $\delta_m=\varphi_n(t_m)-\varphi_{n0}(t_m),\ m=1,\ldots,M,$ and $v_{m,\ell}$ denotes a specific weighting scheme depending on the spectral density estimate of the process $(\sin(tY_m),\cos(tY_m))$. As expected the test has an asymptotic chi-squared distribution under the null hypothesis. However, test consistency, which is generally true for L2-type procedures as those in (14), is compromised with chi-square type measures. The reason is that the uniqueness property of the CF is only valid if this function is viewed over the entire real line and not at certain isolated values of its argument t. Hence, and in order to restore consistency, Nieto-Reyes et al. (2014) suggest a certain version of the test statistic that makes use of the notion of random projections, which however has the "side-effect" of the test procedure becoming much more complicated to apply in practice.

Returning to more parametric frameworks of dependence, we mention the work by Chan, Chen, Peng and Yu (2009) where GOF tests based on the ECF for the null hypothesis $\mathscr{H}_{01}$ in (1) are constructed in the context of continuous-time Lévy processes. Actually the test utilizes the corresponding increments which are well known to be independent and stationary, i.e. they are i.i.d. The novelty of the procedure lies mainly in that the authors employ the empirical likelihood to estimate the parameters, with the likelihood equations resulting from the CF of the increments under the null hypothesis. Then in a fashion analogous to Einmahl and McKeague (2003) their test statistic is the minimized log-likelihood ratio, properly integrated. The parametric bootstrap is employed in order to actually carry out the test for certain well known processes (Black-Scholes and variance gamma), with estimated parameters. In a similar context Lin, Lee and Guo (2013) use the ECF in order to test goodness-of-fit for the marginal law driven by a continuous time stochastic volatility model and show that the parametric bootstrap consistently estimates the limit null distribution of their test statistic.

6. The two-sample and the k-sample problem

In the classical two-sample problem in (2) the ECF approach utilizes the test statistic

$$\Delta_{n_1,n_2,w} = n_1 \int_{-\infty}^{\infty} |\varphi_{n_1}(t) - \varphi_{n_2}(t)|^2 w(t)dt, \tag{18}$$

which results from (5) with φ_{n_k} being the ECF corresponding to the k^{th} sample $Y_{kj},\ j=1,...,n_k$, $k=1,2$. We will refer to the asymptotics later on but for now we wish to illustrate an interesting interpretation of $\Delta_{n_1,n_2,w}$ in terms of moments. To this end assume that the weight function satisfies $\int t^\rho w(t)dt < \infty$, for each $\rho > 0$, and write for simplicity Δ_w for the test statistic in (18). Then by straightforward algebra we have that

$$\frac{\Delta_w}{n_1} = \int g^2(t)w(t)dt, \tag{19}$$

where $g(t) = R_{n_1}(t) + I_{n_1}(t) - R_{n_2}(t) - I_{n_2}(t)$, with $R_{n_k}(t) = n_k^{-1}\sum_{j=1}^{n_k} \cos tY_{kj}$ and $I_{n_k}(t) = n_k^{-1}\sum_{j=1}^{n_k} \sin tY_{kj}$, being the real and imaginary part of φ_{n_k}, $k=1,2$, respectively. In turn, simple Taylor expansions of the $\sin(\cdot)$ and $\cos(\cdot)$ functions lead to the expansion

$$g(t) = \sum_{m=1}^{M} \frac{\gamma_m t^m}{m!}\mathscr{M}_m + o(t^M),\ t \to 0,\ M = 1,2,..., \tag{20}$$

where $\gamma_m = 1$ or -1, $\mathscr{M}_m = \overline{Y_1^{(m)}} - \overline{Y_2^{(m)}}$ and $\overline{Y_k^{(m)}} = n_k{}^{-1}\sum_{j=1}^{n_k} Y_{kj}^m$, $m=1,...,M$, $k=1,2$. Clearly then moment matching takes place in $g(t)$ between the empirical moments computed from the first sample and the corresponding empirical moments of the same order computed from the second sample. Furthermore by substituting (20) into (19) and by some extra algebra we arrive at

$$\frac{\Delta_w}{n_1} = \sum_{m,\ell=1}^{M} \frac{\gamma_m\gamma_\ell}{m!\ell!}\mathscr{M}_m\mathscr{M}_\ell\, \nu_{m,\ell} + remainder, \tag{21}$$

where $v_{m,\ell} = \int t^{m+\ell} w(t)dt$. Hence the role of the weight function is to regulate the weight according to which each moment equation $\mathscr{M}_m$ enters the value of the test statistic. For parametric weight functions such as those in (16) things are even clear regarding the role of the weight parameter $a > 0$. Specifically if a is large then the rate of decay for $w_a(\cdot)$ is high, and consequently we are led to a test statistic that simply matches the sample moments of only low order. By the same reasoning, choosing a value of a close to zero results in a procedure which takes into account not only lower but also higher order moments. Some caution should be exercised in this respect though, since if $a = 0$ in $w_a(t)$ of (16), then it is straightforward to see that the integral in (18) diverges. Consequently values of a which are too close to zero should be avoided as they lead to numerical instability. Likewise taking a to be too large should also be avoided because such values lead to loss of information. These observations are in agreement with the simulation results in the work of Lindsay et al. (2014) mentioned in Section 4 which show that the bandwidth of the Bowman-Foster test for multivariate normality should be neither too small nor too large. For more discussion on this the reader is referred to Meintanis (2013).

Turning now to the k-sample problem ($k > 2$), an analogous test statistic is given by

$$\Delta_{n,w} = \int_{-\infty}^{\infty} \sum_{j=1}^{k} n_j |\varphi_{n_j}(t) - \varphi_n(t)|^2 w(t)dt, \tag{22}$$

where $\varphi_{n_j}(t)$ is the ECF of sample j resulting from (3) by replacing n by the corresponding sample size n_j, $j = 1, ..., k$, and $\varphi_n(t) = \sum_{j=1}^{k} (n_j/n)\varphi_{n_j}^{(j)}(t)$, with $n = \sum_{j=1}^{k} n_j$. The corresponding asymptotic theory is well documented in several papers so we will not present it here. For asymptotics the interested reader is referred to Meintanis (2005), Hušková and Meintanis (2008b), and Hušková and Meintanis (2008c). In fact the last paper considers the sample homogeneity problem in its full generality of $k \geq 2$ multivariate samples. Instead we will review certain recent developments and variations of the test statistics.

Independently of Hušková and Meintanis (2008c) but in an entirely analogous manner, Alba-Fernández, Jiménez-Gamero and Muñoz-Garcia (2008) consider the ECF test statistic in the special case of the (univariate) two-sample problem ($k = 2$). The authors suggest two resampling schemes for approximating the limit null distribution of the test statistic: The permutation (also employed by Hušková and Meintanis, 2008c) and the bootstrap, and show the consistency of both approximations. In a more numerically oriented work Alba-Fernández, Ibáñez-Pérez and Jiménez-Gamero (2004) consider the test statistic in (18) with

$$w(t) = \begin{cases} 1, & |t| < T, \\ 0, & |t| > T, \end{cases} \tag{23}$$

for some $T > 0$. The novelty of this work lies mainly in the computation of the test statistic: They propose a partition of the interval of integration $[-T, T]$ into N subintervals and use a numerical quadrature technique involving Hermite polynomials in order to compute the integral in the resulting test statistic. The asymptotic null distribution of the test statistic coincides with that of a quadratic form involving zero mean Gaussian variables of order $4N$. Then the authors use the bootstrap in order to approximate this distribution and also show consistency of this approximation. An extension of these numerical quadrature ideas to the bivariate case is carried out by Alba-Fernández, Barrera-

Rosillo, Ibáñez-Pérez and Jiménez-Gamero (2009) which is based on an earlier method of Alba-Fernández, Barrera and Jiménez (2001).

We close this section by mentioning the related works of Ghosh and Beran (2000) for the classical two-sample problem which utilizes the moment generating function, of Baringhaus and Kolbe (2015) which utilizes the Hankel transform for the same problem, and of Jiménez-Gamero, Batsidis and Alba-Fernández (2015) for the problem of model selection. We finally mention the ECF-based methods for the k-sample problem with dependent data suggested by Quessy and Éthier (2012).

7. Conclusion

We review testing procedures based on the empirical characteristic function. Our review is not exhaustive and has basically been confined to goodness-of-fit tests for parametric families of distributions and tests for the two-sample and the k-sample problem. Other related areas in which the ECF has been applied are, testing for symmetry, testing for independence, and change-point detection. As far as the first problem is concerned we refer to Meintanis and Ngatchou-Wandji (2012) that provide a review of symmetry tests with particular emphasis on ECF-based procedures, but also mention the articles of Hušková and Meintanis (2012), Klar et al. (2012), Ngatchou-Wandji and Harel (2013), Laïb, Lemdani and Saïd (2013), Jiménez-Gamero (2014), and Henze, Hlávka and Meintanis (2014), which have appeared since that review.

Along the paper we mention occasionally related work that uses the empirical moment generating function. Note that the empirical Laplace transform is another related notion that applies to non-negative observations. For discrete observations however the natural empirical transform is not the ECF (or its real-valued equivalents) but the empirical probability generating function

$$g_n(t) = \int_{-\infty}^{\infty} t^y dF_n(y), \; 0 \leq t \leq 1.$$

Goodness-of-fit tests based on $g_n(t)$ date back to Kocherlakota and Kocherlakota (1986), while Nakamura and Pérez-Abreu (1993) provide an early review of such procedures. We conclude by mentioning the related works of Epps (1995), Gürtler and Henze (2000), Meintanis and Nikitin (2008), Szűcs (2005), and Novoa-Muñoz and Jiménez-Gamero (2014), all of which centre around goodness-of-fit testing for the Poisson distribution.

Acknowledgement

This research was partially supported by grant Nr.11699 of the Special Account for Research Grants (ELKE) of the National and Kapodistrian University of Athens.

References

AHMAD, I. (1993). Modification of some goodness-of-fit statistics to yield asymptotically normal distributions. *Biometrika*, **80**, 466–472.

ALBA-FERNÁNDEZ, M. V., BARRERA, D., AND JIMÉNEZ, M. D. (2001). A homogeneity test based on empirical characteristic functions. *Computational Statistics*, **16**, 255–270.

ALBA-FERNÁNDEZ, V., BARRERA-ROSILLO, D., IBÁÑEZ-PÉREZ, M. J., AND JIMÉNEZ-GAMERO, M. D. (2009). A homogeneity test for bivariate random variables. *Computational Statistics*, **24**, 513–531.

ALBA-FERNÁNDEZ, V., IBÁÑEZ-PÉREZ, M. J., AND JIMÉNEZ-GAMERO, M. D. (2004). A bootstrap algorithm for the two-sample problem using trigonometric Hermite spline interpolation. *Communications in Nonlinear Science and Numerical Simulation*, **9**, 275–286.

ALBA-FERNÁNDEZ, V., JIMÉNEZ-GAMERO, M. D., AND MUÑOZ-GARCIA, J. (2008). A test for the two-sample problem based on empirical characteristic functions. *Computational Statistics & Data Analysis*, **52**, 3730–3748.

ARCONES, M. A. (2007). Two tests for multivariate normality based on the characteristic function. *Mathematical Methods of Statistics*, **16**, 177–201.

BARINGHAUS, L. (1996). Fibonacci numbers, Lucas numbers and integrals of certain Gaussian processes. *Proceedings of the American Mathematical Society*, **124**, 3857–3884.

BARINGHAUS, L. AND HENZE, N. (1988). A consistent test for multivariate normality based on the empirical characteristic function. *Metrika*, **35**, 339–348.

BARINGHAUS, L. AND KOLBE, D. (2015). Two sample tests based on empirical Hankel transforms. *Statistical Papers*, **56**, 597–617.

BOWMAN, A. W. AND FOSTER, P. J. (1993). Adaptive smoothing and density-based tests of multivariate normality. *Journal of the American Statistical Association*, **88**, 529–537.

BRÄKER, H. U. AND HÜSLER, J. (1991). On the first zero of an empirical characteristic function. *Journal of Applied Probability*, **28**, 593–601.

CHAN, N. H., CHEN, S. X., PENG, L., AND YU, C. L. (2009). Empirical likelihood methods based on characteristic functions with applications to Lévy processes. *Journal of the American Statistical Association*, **104**, 1621–1630.

CRAMÉR, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press.

CSÖRGŐ, S. (1984). Testing by the empirical characteristic function: A survey. *In* MANDL, P. AND HUŠKOVÁ, M. (Editors) *Asymptotic Statistics*. Elsevier, pp. 45–56.

CSÖRGŐ, S. (1986). Test for normality in arbitrary dimension. *Annals of Statistics*, **14**, 708–732.

CSÖRGŐ, S. AND HEATHCOTE, C. R. (1987). Testing for symmetry. *Biometrika*, **74**, 177–184.

EINMAHL, J. H. J. AND MCKEAGUE, I. W. (2003). Empirical likelihood based hypothesis testing. *Bernoulli*, **9**, 267–290.

EPPS, T. W. (1987). Testing that a stationary time series is Gaussian. *Annals of Statistics*, **15**, 1683–1698.

EPPS, T. W. (1995). A test of fit for lattice distributions. *Communications in Statistics - Theory and Methods*, **24**, 1455–1479.

EPPS, T. W. (1999). Limit behavior of the ICF test for normality under Gram-Charlier alternatives. *Statistics & Probability Letters*, **42**, 175–184.

EPPS, T. W. (2005). Tests for location-scale families based on the empirical characteristic function. *Metrika*, **62**, 99–114.

EPPS, T. W. AND PULLEY, L. B. (1983). A test for normality based on empirical characteristic function. *Biometrika*, **70**, 723–726.

EPPS, T. W. AND PULLEY, L. B. (1986). A test for exponentiality vs. monotone-hazard alternatives derived from the empirical characteristic function. *Journal of the Royal Statistical Society, Series B*, **48**, 206–213.

FANG, K. T., LI, R. Z., AND LIANG, J. J. (1998). A multivariate version of Ghosh's t_3-plot to detect non-multinormality. *Computational Statistics & Data Analysis*, **28**, 371–386.

FEUERVERGER, A. AND MUREIKA, R. (1977). The empirical characteristic function and its applications. *Annals of Statistics*, **5**, 88–97.

GHOSH, S. (2013). Normality testing for a long-memory sequence using the empirical moment generating function. *Journal of Statistical Planning and Inference*, **143**, 944–954.

GHOSH, S. AND BERAN, J. (2000). Two-sample t_3 plot: A graphical comparison of two distributions. *Journal of Computational and Graphical Statistics*, **9**, 167–179.

GÜRTLER, N. AND HENZE, N. (2000). Recent and classical goodness-of-fit tests for the Poisson distribution. *Journal of Statistical Planning and Inference*, **90**, 207–225.

HEATHCOTE, C. R. (1972). A test of goodness of fit for symmetric random variables. *Australian & New Zealand Journal of Statistics*, **14**, 172–181.

HENZE, N., HLÁVKA, Z., AND MEINTANIS, S. G. (2014). Testing for spherical symmetry via the empirical characteristic function. *Statistics*, **48**, 1282–1296.

HENZE, N. AND WAGNER, T. (1997). A new approach to the BHEP tests for multivariate normality. *Journal of Multivariate Analysis*, **62**, 1–23.

HENZE, N. AND ZIRKLER, B. (1990). A class of invariant and consistent tests for multivariate normality. *Communications in Statistics – Theory and Methods*, **19**, 3595–3617.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2007). Omnibus tests for the error distribution in the linear regression model. *Statistics*, **41**, 363–376.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2008a). Testing procedures based on the empirical characteristic functions I: Goodness-of-fit, testing for symmetry and independence. *Tatra Mountains Mathematical Publications*, **39**, 225–233.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2008b). Testing procedures based on the empirical characteristic functions II: k-sample problem, change point problem. *Tatra Mountains Mathematical Publications*, **39**, 235–243.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2008c). Tests for the multivariate k-sample problem based on the empirical characteristic function. *Journal of Nonparametric Statistics*, **20**, 263–277.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2009). Goodness-of-fit tests for parametric regression models based on the empirical characteristic function. *Kybernetika*, **45**, 960–971.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2010). Tests for the error distribution in nonparametric possibly heteroscedastic models. *Test*, **19**, 92–117.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2012). Tests for symmetric error distribution in nonparametric regression models. *Communications in Statistics – Simulation and Computation*, **41**, 833–851.

JIMÉNEZ-GAMERO, M. D. (2014). On the empirical characteristic function process of the residuals in GARCH models and applications. *Test*, **23**, 409–432.

JIMÉNEZ-GAMERO, M. D., ALBA-FERNÁNDEZ, V., MUÑOZ-GARCÍA, J., AND CHALCO-CANO, Y. (2009). Goodness-of-fit tests based on empirical characteristic functions. *Computational Statistics & Data Analysis*, **53**, 3957–3971.

JIMÉNEZ-GAMERO, M. D., BATSIDIS, A., AND ALBA-FERNÁNDEZ, M. V. (2015). Fourier methods for model selection. *Annals of the Institute of Statistical Mathematics*. DOI: 10.1007/s10463-014-0491-8.

JIMÉNEZ-GAMERO, M. D. AND KIM, H. M. (2015). Fast goodness-of-fit tests based on the characteristic function. *Computational Statistics & Data Analysis*, **89**, 172–191.

JIMÉNEZ-GAMERO, M. D., MUÑOZ-GARCÍA, J., AND PINO-MEJÍAS, R. (2005). Testing goodness-of-fit for the distribution of errors in multivariate linear models. *Journal of Multivariate Analysis*, **95**, 301–322.

KLAR, B., LINDNER, F., AND MEINTANIS, S. G. (2012). Specification tests for the error distribution in GARCH models. *Computational Statistics & Data Analysis*, **56**, 3587–3598.

KOCHERLAKOTA, S. AND KOCHERLAKOTA, K. (1986). Goodness-of-fit tests for discrete distributions. *Communications in Statistics – Theory and Methods*, **15**, 815–829.

KOJADINOVIC, I. AND YAN, J. (2012). Goodness-of-fit testing based on a weighted bootstrap: A fast large sample alternative to the parametric bootstrap. *Canadian Journal of Statistics*, **40**, 480–500.

LAÏB, N., LEMDANI, M., AND SAÏD, E. O. (2013). A functional conditional symmetry test for a GARCH-SM model: Power asymptotic properties. *Statistics & Risk Modeling*, **30**, 75–103.

LEUCHT, A. (2012). Characteristic function-based goodness-of-fit tests under weak dependence. *Journal of Multivariate Analysis*, **108**, 67–89.

LIN, L. C., LEE, S., AND GUO, M. (2013). Goodness-of-fit tests for stochastic volatility models. *Journal of Multivariate Analysis*, **116**, 473–498.

LINDSAY, B., MARKATOU, M., AND RAY, S. (2014). Kernels, degrees of freedom, and power properties of quadratic distance goodness-of-fit tests. *Journal of the American Statistical Association*, **109**, 395–410.

MATSUI, M. AND TAKEMURA, A. (2008). Goodness-of-fit tests for symmetric stable distributions – Empirical characteristic function approach. *Test*, **17**, 546–566.

MEINTANIS, S. G. (2005). Permutation tests for homogeneity based on the empirical characteristic function. *Journal of Nonparametric Statistics*, **17**, 583–592.

MEINTANIS, S. G. (2013). Comments on: An updated review of goodness-of-fit tests for regression models. *Test*, **22**, 432–436.

MEINTANIS, S. G., ALLISON, J., AND SANTANA, L. (2016). Diagnostic tests for the distribution of random effects in multivariate mixed effects models. *Communications in Statistics – Theory and Methods*, **45** (1), 201–215.

MEINTANIS, S. G. AND NGATCHOU-WANDJI, J. (2012). Recent tests for symmetry with multivariate and structured data: A review. *In* ET AL., J. (Editor) *Non-Parametric Statistical Methods and Related Topics*. World Scientific, pp. 35–73.

MEINTANIS, S. G., NGATCHOU-WANDJI, J., AND TAUFER, E. (2015). Goodness-of-fit tests for multivariate stable distributions based on the empirical characteristic function. *Journal of Multivariate Analysis*, **140**, 171–192.

MEINTANIS, S. G. AND NIKITIN, Y. (2008). A class of count modelsand a new consistent test for the Poisson distribution. *Journal of Statistical Planning and Inference*, **138**, 3722–3732.

MEINTANIS, S. G. AND SWANEPOEL, J. W. H. (2007). Bootstrap goodness-of-fit tests with estimated parameters based on empirical transforms. *Statistics & Probability Letters*, **77**, 1004–1013.

NAITO, K. (1996a). Modification of test statistics based on the empirical characteristic function to yield asymptotic normality. *Communications in Statistics – Theory and Methods*, **25**, 105–114.

NAITO, K. (1996b). On weighting the studentized empirical characteristic function for testing normality. *Communications in Statistics – Simulation and Computation*, **25**, 201–213.

NAITO, K. (1997). On the asymptotic normality of the l_2-distance class of statistics with estimated parameters. *Journal of Nonparametric Statistics*, **8**, 199–214.

NAKAMURA, M. AND PÉREZ-ABREU, V. (1993). Empirical probability generating function: An overview. *Insurance: Mathematics and Economics*, **12**, 287–295.

NEUMEYER, N., DETTE, H., AND NAGEL, E. R. (2006). Bootstrap tests for the error distribution in linear and nonparametric regression models. *Australian & New Zealand Journal of Statistics*, **48**, 129–156.

NGATCHOU-WANDJI, J. AND HAREL, M. (2013). A Cramér-von Mises test for symmetry of the error distribution in asymptotically stationary stochastic models. *Statistical Inference for Stochastic Processes*, **16**, 207–236.

NIETO-REYES, A., CUESTA-ALBERTOS, J. A., AND GAMBOA, F. (2014). A random projection based test for Gaussianity for stationary processes. *Computational Statistics & Data Analysis*, **75**, 124–141.

NOVOA-MUÑOZ, F. AND JIMÉNEZ-GAMERO, M. D. (2014). Testing for the bivariate Poisson distribution. *Metrika*, **77**, 771–793.

PARZEN, E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, **33**, 1065–1076.

POURAHMADI, M. (1987). Some sampling properties of empirical characteristic functions viewed as harmonizable stochastic processes. *Journal of Statistical Planning and Inference*, **17**, 345–359.

PRESS, S. J. (1972). Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, **67**, 842–846.

PUDELKO, J. (2005). On an affine invariant and consistent test for multivariate normality. *Probability and Mathematical Statistics*, **25**, 43–54.

QUESSY, J. F. AND ÉTHIER, F. (2012). Cramér-von Mises and characteristic function tests for the two and *k*-sample problems with dependent data. *Computational Statistics & Data Analysis*, **56**, 2097–2111.

SEJDINOVIC, D., SRIPERUMBUDUR, B., GRETTON, A., AND FUKUMIZU, K. (2013). Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *Annals of Statistics*, **41**, 2263–2291.

SZŰCS, G. (2005). Approximations of empirical probability generating processes. *Statistics & Decisions*, **23**, 67–80.

TENREIRO, C. (2009). On the choice of the smoothing parameter for the BHEP goodness-of-fit test. *Computational Statistics & Data Analysis*, **53**, 1038–1053.

TENREIRO, C. (2011). An affine invariant multiple test for assessing multivariate normality. *Computational Statistics & Data Analysis*, **55**, 1980–1992.

USHAKOV, N. (1999). *Selected Topics in Characteristic Functions*. VSP: Utrecht.

WELSH, A. H. (1986). Implementing empirical characteristic functions procedures. *Statistics & Probability Letters*, **4**, 65–67.

WONG, W. K. AND SIM, C. H. (2000). Goodness-of-fit test based on empirical characteristic function. *Journal of Statistical Computation and Simulation*, **65**, 243–269.

Manuscript received, 2015-06-03, revised, 2015-08-31, accepted, 2015-10-31.

COMMENTS: A REVIEW OF TESTING PROCEDURES BASED ON THE EMPIRICAL CHARACTERISTIC FUNCTION

Norbert Henze
Karlsruhe Institute of Technology (KIT)
e-mail: *Henze@kit.edu*

Professor Meintanis is to be congratulated for this nice review paper on tests based on the empirical characteristic function (ECF). There is not much to add, but I want to emphasize one basic point. There are numerous results on the limit behavior of ECF based tests under contiguous alternatives to the null hypothesis, and usually authors derive an almost sure limit for the suitably scaled test statistic under fixed alternatives. At least to my knowledge, however, there is only one result on the limit distribution of the test statistic under *fixed* alternatives, which is due to Gürtler (2000), and which refers to the BHEP tests for multivariate normality. Since the approach taken by Gürtler will probably be useful also for other tests based on the ECF, we outline the reasoning in what follows.

For testing the hypothesis H_0 that an unknown d-variate distribution is some non-degenerate normal distribution, based on independent and identically distributed copies $X_1,\ldots,X_n$ of a d-variate random column vector X satisfying $\mathbb{E}\|X\|^2 < \infty$, the BHEP statistic is

$$T_{n,\beta} = n\int_{\mathbb{R}^d}\left|\Psi_n(t) - \exp\left(-\frac{\|t\|^2}{2}\right)\right|^2 \varphi_\beta(t)\,\mathrm{d}t.$$

Here, $\beta > 0$, $\varphi_\beta(t) = (2\pi\beta^2)^{-1/2}\exp(-t^2/(2\beta^2))$, $\Psi_n(t) = n^{-1}\sum_{k=1}^n \exp(it^\top Y_{n,k})$ is the ECF of the scaled residuals $Y_{n,k} = S_n^{-1/2}(X_k - \overline{X}_n)$, $S_n^{-1/2}$ is the symmetric square root of the inverse of the empirical covariance matrix of $X_1,\ldots,X_n$, and $\overline{X}_n = n^{-1}\sum_{j=1}^n X_j$. Because of affine invariance, we assume that $\mathbb{E}[X] = 0$ and $\mathbb{E}[XX^\top] = \mathrm{I}_d$, the unit d-matrix. Baringhaus and Henze (1988) proved that

$$\frac{T_{n,\beta}}{n} \longrightarrow L_\beta \;=\; \int_{\mathbb{R}^d}\left|\Psi(t) - \exp\left(-\frac{\|t\|^2}{2}\right)\right|^2 \varphi_\beta(t)\,\mathrm{d}t$$

almost surely as $n\to\infty$, where $\Psi(t) = \mathbb{E}[\exp(it^\top X)]$ is the characteristic function of X. Under the condition $\mathbb{E}\|X\|^4 < \infty$ Gürtler (2000) proved that

$$\sqrt{n}\left(\frac{T_{n,\beta}}{n} - L_\beta\right) \xrightarrow{\mathscr{D}} \mathscr{N}(0,\tau_\beta^2), \tag{1}$$

where $\xrightarrow{\mathscr{D}}$ is convergence in distribution and τ_β^2 depends on the distribution of X in an explicit way. To prove (1), Gürtler observed that the left-hand side of (1) equals

$$2\int_{\mathbb{R}^d} Z_n(t)\left[\widetilde{\psi}(t) - \exp\left(-\frac{\|t\|^2}{2}\right)\right]\varphi_\beta(t)\,\mathrm{d}t + \frac{1}{\sqrt{n}}\int_{\mathbb{R}^d} Z_n(t)\varphi_\beta(t)\,\mathrm{d}t, \tag{2}$$

where

$$Z_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left[\cos(t^\top Y_{n,j}) + \sin(t^\top Y_{n,j}) - \widetilde{\Psi}(t) \right], \quad \widetilde{\Psi}(t) = \mathbb{E}[\cos(t^\top X) + \sin(t^\top X)].$$

By working in the Hilbert space $\mathscr{H} = L^2(\mathbb{R}^d, \mathscr{B}^d, \varphi_\beta(t)\mathrm{d}t)$, Gürtler proved the convergence of Z_n to some centered Gaussian element Z of $\mathscr{H}$. From (2) and the continuous mapping theorem, the limit distribution of $\sqrt{n}(T_{n,\beta}/n - L_\beta)$ is then the (centered normal) distribution of the first summand of (2), with Z_n replaced by the limit process Z.

Since a consistent estimator $\tau_{\beta n}$ of τ_β is available, the convergence (1) can be used to approximate the power of the BHEP test. If $q_{\beta,n,1-\alpha}$ denotes the critical value of an (upper rejection region) test based on $T_{n,\beta}$ at nominal level α, (1) yields the approximation

$$\mathbb{P}(T_{n,\beta} > q_{\beta,n,1-\alpha}) \; \approx \; 1 - \Phi\left(\frac{q_{\beta,n,1-\alpha} - nL_\beta}{\sqrt{n}\tau_{\beta n}} \right),$$

where Φ is the distribution function of the standard normal distribution. By considering various alternatives, Gürtler also showed that replacing L_β by its consistent estimator $T_{n,\beta}/n$ in the above expression yields a good approximation to the power of the BHEP test. Moreover,

$$\left[\frac{T_{n,\beta}}{n} - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\tau_{\beta n}}{\sqrt{n}}, \frac{T_{n,\beta}}{n} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\tau_{\beta n}}{\sqrt{n}} \right]$$

is an asymptotic confidence interval for L_β at level $1-\alpha$.

A final important point is that any goodness-of-fit test in the classical sense is not able to 'validate' the null hypothesis H_0. If H_0 is not rejected, the data are not 'in sufficient contradiction to H_0', but nothing more can be concluded. The limit result (1), however, can be used to construct an asymptotic test for the hypothesis $H_\Delta : L_\beta \geq \Delta$ against the alternative $K_\Delta : L_\beta < \Delta$, where $\Delta > 0$ is a given positive number. Suppose that we reject H_Δ if

$$\frac{T_{n,\beta}}{n} \leq \frac{\tau_{\beta n}}{\sqrt{n}} \Phi^{-1}(\alpha) + \Delta.$$

Then for each alternative distribution (function) $F \in H_\Delta$ we have

$$\limsup_{n\to\infty} \mathbb{P}_F(\text{‘rejecting } H_\Delta\text{’}) \leq \alpha$$

with equality if $L_\beta = L_\beta(F) = \Delta$. If such an 'inverse test' rejects H_Δ, there is evidence that the underlying distribution is sufficiently near to a normal distribution, at least with regard to the distance measure L_β.

References

BARINGHAUS, L. AND HENZE, N. (1988). A consistent test for multivariate normality based on the empirical characteristic function. *Metrika*, **35**, 339–348.

GÜRTLER, N. (2000). *Asymptotic Studies on the Class of BHEP Tests for Multivariate Normality with Fixed and Variable Smoothing Parameter (in German)*. Phd thesis, University of Karlsruhe, Germany.

Manuscript received, 2015-07-26, revised, 2015-10-05, accepted, 2015-10-05.

COMMENTS: A REVIEW OF TESTING PROCEDURES BASED ON THE EMPIRICAL CHARACTERISTIC FUNCTION

Marie Hušková
Department of Statistics, Charles University in Prague, Sokolovská 83, 18675 Prague, Czech Republic
e-mail: *huskova@karlin.mff.cuni.cz*

It is a nicely and clearly written paper on statistical procedures based on the empirical characteristic function (ECF). The paper provides an overview of recently developed goodness-of-fit tests based on functionals of the ECF for both simple setups as well as rather advanced situations (with nuisance parameters). The author also touches on procedures based on the ECF for the k-sample problem.

In this discussion I would like to bring attention to another class of procedures based on the ECFs used in *change point analysis*, that is, a class of statistical procedures for the detection of instabilities in various statistical models based on functionals of the ECFs.

Introduction

In change point analysis we typically have observations $Y_1, \ldots, Y_n$ obtained at ordered time points and the basic task is to decide whether the model remains stable during the whole observational period or whether the model changes at some unknown time point (called the change point) or become generally unstable. In case change(s) are detected the further task is also to estimate the time of the change and other parameters of the model in the periods where the model is stable.

Construction of such procedures are usually based on similar principles as in other testing and estimation problems like maximum likelihood ratio, empirical likelihood, robust procedures, rank based procedures, etc. If the location of the change point is known the problem reduces to a general version of the two-sample problem.

There are a number of monographs and survey papers tackling the problem of detection of changes from various points of view, various levels of generality, etc., e.g. Brodsky and Darkhovsky (1993), Basseville and Nikiforov (1993), Csörgő and Horváth (1997), Chen and Gupta (2000). A partial survey of basic procedures till 2000 can be found in, for example, Antoch, Hušková and Jarušková (2000).

In the next two sections we survey test procedures for the detection of changes based on ECF in *off-line* and *on-line* setups.

Off-line procedures

Let $Y_1,\ldots,Y_n$ be independent random variables, let Y_j have a distribution function $F_j, j=1,\ldots,n$, and consider the testing problem

$$H_0: F_1 = \ldots = F_n \tag{1}$$

against

$$H_1: F_1 = \ldots = F_{k_0} \neq F_{k_0+1} = \ldots = F_n \qquad for \quad k_0 < n, \tag{2}$$

where k_0 is an unknown *change point*, F_1 and F_n are also unknown. Motivated by the two-sample tests based on ECFs (see Section 6 of the overview paper) Hušková and Meintanis (2006a) introduced the following class of test statistics:

$$T_{n,\gamma}(w) = \max_{1\le k<n} \Big(\frac{k(n-k)}{n^2}\Big)^{1+\gamma} \int_{-\infty}^{\infty} n|\widehat{\varphi}_k(t) - \widehat{\varphi}_k^0(t)|^2 w(t)dt, \tag{3}$$

where $w(\cdot)$ is a nonnegative weight function, $\widehat{\varphi}_k(.)$ and $\widehat{\varphi}_k^0(.)$ are empirical characteristic functions based on $Y_1,\ldots,Y_k$ and $Y_{k+1},\ldots,Y_n$, respectively, i.e.,

$$\widehat{\varphi}_k(t) = \frac{1}{k}\sum_{j=1}^{k} \exp\{itY_j\}, \quad \widehat{\varphi}_k^0(t) = \frac{1}{n-k}\sum_{j=k+1}^{n} \exp\{itY_j\}, \quad k=1,\ldots,n, \tag{4}$$

and γ is a positive tuning constant. The weight function w plays the same role as in the overview paper. Large values of the test statistics indicate that the null hypothesis is violated. Limit properties can be formulated in two ways, either through function of degenerate U-statistics or as a functional of a Gaussian processes. Particularly, the limit behavior ($n\to\infty$) under the null hypothesis of $T_{n,\gamma}(w)$ is the same as that of

$$\max_{s\in(0,1)} (s(1-s))^{1+\gamma} \int V^2(s,t)w(t)dt,$$

where $\{V(s,t); t\in R^1, s\in(0,1)\}$ is a zero-mean Gaussian process with the covariance structure

$$\begin{aligned} cov(V(s_1,t_1),V(s_2,t_2)) =& cov(\cos(t_1Y_1)+\sin(t_1Y_1), \cos(t_2Y_2)+\sin(t_2Y_2)) \\ &\times \frac{1}{\max(s_1,s_2)(1-\min(s_1,s_2))} \end{aligned}$$

when the weight function $w(\cdot)$ and the characteristic function satisfy some mild conditions. The limit distribution is not useful in getting an approximation for critical values, but these can be obtained through the bootstrap (with or without replacement). However, while the bootstrap approximation is asymptotically correct under the null hypothesis or local alternatives, this is not true for fixed alternatives, but the resulting test is consistent. Details including theorems, proofs, simulations and discussion are found in Hušková and Meintanis (2006a).

In addition, assuming that the distribution functions $F_1,\ldots,F_n$ are continuous, Hušková and Meintanis (2006b) proposed a rank based version of the above statistics. Denote by $R_1,\ldots,R_n$ the ranks related to $Y_1,\ldots,Y_n$. The respective test statistic is defined as follows

$$T_{n,\gamma}(w,R) = \max_{1\le k<n} \Big(\frac{k(n-k)}{n^2}\Big)^{1+\gamma} n\int_{-\infty}^{\infty} |\widehat{\varphi}_k(t,R) - \widehat{\varphi}_k^0(t,R)|^2 w(t)dt, \tag{5}$$

where $R = (R_1, \ldots, R_n)$, $w(\cdot)$ is a nonnegative weight function, $\widehat{\varphi}_k(.,R)$ and $\widehat{\varphi}_k^0(.,R)$ are empirical characteristic functions based on $R_1, \ldots, R_k$ and $R_{k+1}, \ldots, R_n$, respectively, i.e.,

$$\widehat{\varphi}_k(t,R) = \frac{1}{k}\sum_{j=1}^{k} \exp\{itR_j/n\}, \quad \widehat{\varphi}_k^0(t,R) = \frac{1}{n-k}\sum_{j=k+1}^{n} \exp\{itR_j/n\}, \quad k = 1, \ldots, n-1, \tag{6}$$

and γ is a positive constant.

The advantage of the respective test procedure is that under the null hypothesis it is distribution free and therefore the approximation for critical values can be obtained by simulations, which is not the case for $T_{n,\gamma}(w)$. Limit properties under the null as well as under alternatives are studied in the paper by Hušková and Meintanis (2006b), where the results of a simulation study is also presented. In this paper procedures based on empirical distribution functions (Kolmogorov-Smirnov type) are also discussed, including comparisons via simulations.

The above described testing procedures are suitable when there are not finite dimensional nuisance parameters. For example, if the problem is to test for a change in the distribution of the error terms, then the regression parameters are nuisance ones. The above procedures can be modified to replace the observations in (6) by residuals similarly to what was done in Section 2 of the survey paper.

On-line procedures

This section is devoted to the on-line procedures for detection of change based on ECF. We explain the basic formulation in the following simple basic setup. The observations $Y_1, \ldots, Y_n, \ldots$ arrive sequentially, Y_i has the continuous distribution function F_i, $i = 1, \ldots,$ and a training sample of size m with no change is available such that the first m observations have the same distribution function F_0, i.e.,

$$F_1 = \ldots = F_m = F_0, \tag{7}$$

where F_0 is unknown. We are interested in testing the null hypothesis:

$$H_0 : F_i = F_0, \quad \forall i > m, \tag{8}$$

against the alternative:

$$\begin{aligned} H_1 &: \text{there exists } 1 \le k^* < \infty \text{ such that } F_i = F_0,\ 1 \le i < m + k^* \\ &\quad \text{and } F_i = F^0 \neq F_0,\ k^* + m \le i < \infty. \end{aligned}$$

The number k^* is also called the *change point*.

The related test procedures are usually described by the stopping rule:

$$\tau_{m,T} = \inf\{1 \le k < mT + 1 : Q(m,k) \ge c\,q(k/m)\}, \tag{9}$$

with $\inf\emptyset := +\infty$, where the detector $Q(m,k)$ is a statistic calculated from observations $Y_1, \ldots, Y_{m+k}$, $k = 1, \ldots,$ c is a tuning constant, and

$$q_\gamma(t) = (t/(1+t))^{1+\gamma}, \quad t > 0,$$

is a boundary function with $0 < \gamma \le 1$. Finally, $0 < N < \infty$ is such that $\lfloor (N+1)m \rfloor + 1$ is the upper bound of the maximum number of observations that can be conducted during a certain period.

The corresponding decision rule is usually formulated as follows: the null hypothesis is rejected as soon as for some $1 \le k < mN+1$

$$Q(m,k) \ge c_\alpha \, q_\gamma(k/m)$$

where c_α is determined in such a way that the asymptotic significance level of the test is equal to $\alpha(\in (0,1))$, i.e., under the null hypothesis:

$$\lim_{m\to\infty} P_{H_0}\big(\max_{1\le k<N+1} (Q(m,k)/q(k/m)) \ge c_\alpha \big) = \alpha.$$

Moreover, it is required that at least for fixed alternatives the test is consistent, i.e., a false null hypothesis is rejected with probability tending to 1, as $m \to \infty$. Notice that in this formulation the number of observations until a decision is made is a random variable.

The choice of the detectors $Q(m,k), k = 1,\ldots,$ based on observations $Y_1,\ldots,Y_{m+k}$ is crucial. We take here ECFs based ones:

$$Q(m,k) = \frac{1}{m} \int_{-\infty}^{\infty} |\widehat{\varphi}_{m,m+k}(u) - \widehat{\varphi}_{p,m}(u)|^2 w(u)du, \tag{10}$$

where

$$\widehat{\varphi}_{j_1,j_2}(u) = \frac{1}{j_2 - j_1} \sum_{t=j_1+1}^{j_2} \exp\{iuY_t\}$$

is the ECF based on $Y_{j_1+1},\ldots,Y_{j_2}$.

It can be shown that if the observations $Y_1,\ldots,Y_n,\ldots,$ are i.i.d. random variables and under some additional quite mild assumptions, as $m \to \infty$ (i.e., the number of training data tends to tend to ∞), then

$$\max_{1\le k<mT} \frac{Q(m,k)}{q_\gamma(k/m)}$$

has the same limit distribution as

$$\sup_{0<s<T} s^{\gamma-1}(1+s)^{-(1+\gamma)} \int Z^2(t,s)w(t)dt,$$

where $\{Z(t,s); t \in R^1, 0 < s < T\}$ is a Gaussian process with zero mean and dependence structure depending on the unknown characteristic function. The limit distribution can therefore not be used to approximate the critical value c_α. However, the bootstrap based on training data provides asymptotically correct approximations for critical values. Also, it can be shown that under quite mild conditions the test is consistent for a quite wide spectrum of alternatives. More details are available in Hlávka and Hušková (2012).

Next we consider a more general setup, namely, we assume $\{Y_j,\ j = p+1,\ldots,\}$ is an AR(p) process defined by the equation

$$Y_j = \beta^T Y_{j-1} + \varepsilon_j, \quad j > p, \tag{11}$$

where $Y_{j-1} = (Y_{j-1}, \ldots, Y_{j-p})^T$, and $\beta = (\beta_1, ..., \beta_p)^T$ is a vector of unknown regression parameters. The error terms $\{\varepsilon_j, j = p+1, \ldots, \}$ are independent, each having a corresponding distribution function $F_j, j = p+1, \ldots,$ with mean zero and finite variance.

We are interested in a change in distributional aspect of the errors ε_j, i.e. we wish to test the hypothesis

$$H_0 : F_j = F_0, j = p+1, m+2, \ldots \quad \text{vs.}$$
$$H_1 : F_j = F_0, j \leq m+k_0; F_j = F^0 \neq F_0, j > m+k_0,$$

where F_0, F^0, and the time of a change $k_0 \geq 1$, are considered unknown. Under the null hypothesis the AR process is assumed to be stationary i.e., the characteristic polynomial $P(z) = 1 - \beta_1 z - \ldots - \beta_p z^p$, is assumed to satisfy $P(z) \neq 0, \forall |z| \leq 1$. Notice that β is a nuisance parameter.

Since the error terms are unobserved, typically one computes the residuals

$$\widehat{\varepsilon}_j = Y_j - \widehat{\beta}_m^T \mathbf{Y}_{j-1}^T, \tag{12}$$

where $\widehat{\beta}_m := \widehat{\beta}_T(Y_1, \ldots, Y_m)$ is an estimator of β, based only on the training data set $Y_1, \ldots, Y_m$, and satisfying

$$\sqrt{m}(\widehat{\beta}_m - \beta) = O_P(1), \text{ as } m \to \infty. \tag{13}$$

Based on the estimated residuals we now replace in the ECF in (10) $\widehat{\varphi}_{j_1,j_2}(u)$ by a different ECF

$$\tilde{\varphi}_{j_1,j_2}(u) = \frac{1}{j_2 - j_1} \sum_{t=j_1+1}^{j_2} \exp\{iu\widehat{\varepsilon}_t\}.$$

With this replacement we proceed as in the previous simple situation (even the limit behaviour as $m \to \infty$ is the same). Interestingly, as soon as (13) holds true the limit behaviour does not depend on the chosen estimator β_m. It turns out, however, that the monitoring schemes developed for these distributional changes are also able to detect changes in the regression coefficient.

Further results including discussion, modifications, proofs, simulations, etc. on these type of procedures can be found in Hlávka, Hušková, Kirch and Meintanis (2012, 2015). These also include references for other procedures with the focus on the procedures based on empirical distribution functions of the estimated residuals and comparison with ours.

Detection of changes in time series of counts

Finally, we briefly mention procedures for detection of changes in time series of counts. Most of the above procedures can also be used in this situation, but it appears that in this situation it is more natural to base the procedures on the empirical moment generating functions as mentioned at the end of the survey paper. Hudecová, Hušková and Meintanis (2015a, b, c) have recently developed and studied such procedures for changes in integer autoregression (INAR) Poisson autoregression (PAR). Theoretical results are accompanied by simulations and applications to real data sets.

Acknowledgement

The work was partially supported by the grant GAČR 15-096635.

References

ANTOCH, J., HUŠKOVÁ, M., AND JARUŠKOVÁ, D. (2000). Change point analysis. Lecture Notes from the 5th IASC Summer School. 1–76.

BASSEVILLE, M. AND NIKIFOROV, I. V. (1993). *Detection of Abrupt Changes: Theory and Applications*. Prentice-Hall, Inc.

BRODSKY, B. S. AND DARKHOVSKY, D. E. (1993). *Nonparametric Methods in Change-point Problems*. Kluwer Academic: Dordrecht.

CHEN, J. AND GUPTA, A. K. (2000). *Parametric Statistical Change Point Analysis*. Birkhäuser Boston, Inc.: Boston, MA.

CSÖRGŐ, M. AND HORVÁTH, L. (1997). *Limit Theorems in Change-Point Analysis*. Wiley: Chichester.

HLÁVKA, Z. AND HUŠKOVÁ, M. (2012). Nonparametric sequential monitoring. *Sequential Analysis*, **31**, 278–296.

HLÁVKA, Z., HUŠKOVÁ, M., KIRCH, C., AND MEINTANIS, S. G. (2012). Monitoring changes in the error distribution of autoregressive models based on Fourier methods. *Test*, **21**, 605–634.

HLÁVKA, Z., HUŠKOVÁ, M., KIRCH, C., AND MEINTANIS, S. G. (2015). Bootstrap procedures for on-line monitoring of changes in autoregressive models. *Communications in Statistics – Simulation and Computation*. DOI: 10.1080/03610918.2014.904346.

HUDECOVÁ, Š., HUŠKOVÁ, M., AND MEINTANIS, S. G. (2015a). Tests for structural changes in time series of counts. *Submitted*.

HUDECOVÁ, Š., HUŠKOVÁ, M., AND MEINTANIS, S. G. (2015b). Detection of changes in INAR models. *In* STELAND, A., RAFAJOWICZ, E., AND SZAJOWSKI, K. (Editors) *Stochastic Models, Statistics and Their Applications*, volume 122. Springer Proceedings in Mathematics & Statistics, pp. 11–18.

HUDECOVÁ, Š., HUŠKOVÁ, M., AND MEINTANIS, S. G. (2015c). Detection of changes in INARCH models. *Submitted*.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2006a). Change point analysis based on empirical characteristic functions. *Metrika*, **63**, 145–168.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2006b). Change point analysis based on empirical characteristic functions of ranks. *Sequential Analysis*, **25**, 421–436.

Manuscript received, 2015-08-13, revised, 2015-10-12, accepted, 2015-10-12.

COMMENTS: A REVIEW OF TESTING PROCEDURES BASED ON THE CHARACTERISTIC FUNCTION

J.W.H. Swanepoel
North-West University, Potchefstroom, South Africa
and
J.S. Allison [1]
North-West University, Potchefstroom, South Africa
e-mail: *james.allison@nwu.ac.za*

First of all, we would like to congratulate the author for writing this interesting and timely review paper on testing procedures based on the empirical characteristic function. The exposition of the ideas and methods is necessarily brief, given the wide scope of the paper. So, we strongly encourage the author to write a monograph on the topic, where main procedures can be developed in detail. Although it is really hard to add something interesting to this review, we would like to comment on three issues. The author provides a very limited discussion on the first, but did not consider the second and third in his paper.

Bootstrap critical values

In spite of the good statistical properties of tests based on the empirical characteristic function (they are usually consistent against fixed alternatives and able to detect local alternatives converging to the null hypothesis at the rate $n^{-1/2}$), these tests possess certain computational difficulties from a practical point of view. A main problem is the calculation of the critical values of the test, because the exact null distribution of the test statistic is unknown. In most cases the asymptotic null distribution does not provide a useful approximation. Moreover, large-sample critical values are extremely complicated (if not impossible) to compute.

Meintanis and Swanepoel (2007) discussed a parametric bootstrap (PB) procedure, which can be applied to approximate the critical values consistently under quite general circumstances. The validity of the PB procedure was shown analytically by the authors. Although very easy to implement, the PB procedure can become computationally expensive as the sample size, the number of parameters or the dimension of the data increase. This problem is not specific to goodness-of-fit test based on the empirical characteristic function, the same problem arises when one instead consider a test based on comparing the empirical cumulative distribution function (cdf) and a parametric estimator of the cdf under the null hypothesis.

[1] Corresponding author.

In their paper "Fast goodness-of-fit tests based on the characteristic function", Jiménez-Gamero and Kim (2015) proposed to estimate the null distribution of the test statistic consistently through a weighted bootstrap (WB), in the sense of Burke (2000), Horváth, Kokoszka and Steineback (2000) and Kojadinovic and Yan (2012). In their numerical examples carried out, the estimated type I errors were close to the prescribed nominal values. The general consensus reached by all the above-mentioned authors is that for small samples, the PB usually appears more powerful, and, since it typically has an acceptable computational cost in that case, it is the recommended approach. The asymptotic properties of the PB and WB procedures are similar. However, for larger samples, the use of the PB procedure can become very tedious in practice and the WB procedure appears as a natural alternative, which is less computational intensive. The following major disadvantages of the WB, which makes it difficult to apply in practice, should be mentioned:

(a) To perform the WB, so-called *multiplier* i.i.d. random variables $\zeta_1, \zeta_2, \ldots, \zeta_n$ with expectation zero and variance one, which are independent of the sample $X_1, X_2, \ldots, X_n$, are generated for *each* bootstrap replication. The question here, which has not yet been addressed in the literature satisfactorily, is from which distribution should these multipliers be drawn? What is a best choice?

(b) The WB procedure assumes that $\hat{\theta}_n$ is an estimator of θ that is asymptotically linear with influence function $l(x;\theta)$, with θ being the true parameter value when the null hypothesis is true.

Although this condition is satisfied by some commonly used estimators such as quasi maximum likelihood estimators, method of moment estimators and minimum integrated squared error estimators, it is generally a difficult (if not impossible) condition to verify, especially for a more complicated $\hat{\theta}_n$. In addition, the function $l(x;\theta)$ is usually unknown and has to be estimated, which can also be an impossible task. Contrary to this, the PB procedure does not require such an assumption on $\hat{\theta}_n$.

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

Recall that the characteristic function (CF) of a random variable X is given by $\varphi(t) = E(e^{itX})$, and the empirical CF (ECF) is defined as

$$\varphi_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{itX_j},$$

where $X_1, X_2, \ldots, X_n$ are i.i.d. observations on X. Clearly the ECF is an unweighted average and puts the same weight on observations near the centre of the distribution and on observations at the tails of the distribution. In the past, standard methods of estimation and testing via the ECF have utilized the L2-type distance

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

which apart from $\varphi(t)$ and $\varphi_n(t)$ 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. 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. Some discussion on how to estimate λ by the bootstrap, for a given known functional form of $w_\lambda(\cdot)$, in a goodness-of-fit type setting for a location-scale parametric family of distributions, will be discussed in Section 3 below.

In order to alleviate the problem of choosing $w_\lambda(\cdot)$, Meintanis, Swanepoel and Allison (2014) introduced the notion of the probability weighted characteristic function (PWCF) as a generalization of the usual characteristic function. The authors then suggested 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 parameter λ. Specifically for $\lambda \geq 0$, they defined the PWCF as

$$\chi(t;\lambda) = \int_{-\infty}^{\infty} W(x;\lambda t) e^{itx} dF(x),$$

where F is the cdf of X and the probability weight is given by

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

The PWCF has some interesting properties which often relate to the corresponding properties of the classical CF. Among others, the PWCF characterizes the distribution of a random variable uniquely.

In their paper Meintanis et al. (2014) were primarily interested in applying the PWCF to a goodness-of-fit type setting by testing that F belongs to a parametric family of distributions $\{F_\theta : \theta \in \Theta\}$, where Θ is an open subset of $\mathbb{R}^p$, $p \geq 1$. They defined the empirical PWCF as

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

where the estimated probability weight is given by

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

and $\hat{\theta}_n = \hat{\theta}_n(X_1,\ldots,X_n)$ is a consistent estimator of θ. The empirical PWCF assigns more weight to observations near the median of the distribution, while observations away from the center receive progressively less weight. It was shown that empirical PWCF-based test statistics are convenient from a computational point of view, and that, under general conditions, leads to consistent procedures with a well defined asymptotic null distribution. Moreover, Monte Carlo results indicated that the new tests compared favorably with more established procedures.

We close this section by noting that there exist several alternative aspects of the PWCF that deserve further attention:

(i) The study of the empirical PWCF as a stochastic process and proving, e.g., its weak convergence.

(ii) Introduction of a nonparametric version of the PWCF. In this regard a natural idea would be to replace in the definition of $W(x;\beta)$ the cdf F by, e.g., the empirical cdf or an appropriate kernel estimator of F.

(iii) Application of the PWCF in semi-parametric and nonparametric inference problems, such as testing for symmetry and testing for independence.

A data-dependent choice of the tuning parameter λ

In a multitude of goodness-of-fit tests based on the empirical CF an unknown tuning parameter λ appears, as in the case of the weight functions $w_\lambda(t)$ and $W(x;\lambda t)$ discussed above. To apply these tests in practice, a specific choice of λ would be required. Researchers typically approach this problem by evaluating the power performance of their tests across a grid of values of λ and then suggest a compromise choice by selecting a value for λ that fares well for the majority of the alternatives considered. However, even though these fixed values for λ are suggested in the literature, there is a general agreement that the development of a data-based choice of λ is required for practical implementation. Further motivation for this choice is that the powers of the tests are often observed to fluctuate wildly for different alternatives with a fixed choice of the tuning parameter λ.

Allison and Santana (2015) proposed a data-dependent choice for λ, which appears in many goodness-of-fit test statistics, in an attempt to move away from a fixed choice of the parameter. Their method is based on the bootstrap and is applicable to a class of distributions for which the null distribution of the test statistic is independent of unknown parameters. The new method was investigated by means of a Monte Carlo study and the authors found that the Monte Carlo power of these tests, using the data dependent choice of λ, compared favorably to the maximum achievable power for the tests calculated over a grid of values of the tuning parameter λ.

In view of the above discussion, a challenging research problem would be to modify the ideas and methods by Allison and Santana (2015) in order to apply them to other related areas in which the CF has been applied, e.g., testing for symmetry, testing for independence, and change-point detection.

Final remark

This review by Meintanis will contribute much to the research activity of people concerned with testing procedures based on the empirical CF. It provides a quick and clear insight into the main ideas, problems and existing results in the area. We would be interested to hear the thoughts of the author on the above raised issues, and on potential approaches based on the literature surveyed.

Acknowledgements

The first author would like to thank the National Research Foundation of South Africa for financial support.

References

ALLISON, J. S. AND SANTANA, L. (2015). On a data-dependent choice of the tuning parameter appearing in certain goodness-of-fit tests. *Journal of Statistical Computation and Simulation*, **85** (16), 3276–3288.

BURKE, M. D. (2000). Multivariate tests-of-fit and uniform confidence bands using a weighted bootstrap. *Statistics & Probability Letters*, **46**, 13–20.

HORVÁTH, L., KOKOSZKA, P., AND STEINEBACK, J. (2000). Approximation for weighted bootstrap processes with an application. *Statistics & Probability Letters*, **48**, 59–70.

JIMÉNEZ-GAMERO, M. D. AND KIM, H.-M. (2015). Fast goodness-of-fit tests based on the characteristic function. *Computational Statistics & Data Analysis*, **89**, 172–191.

KOJADINOVIC, I. AND YAN, J. (2012). Goodness-of-fit testing based on a weighted bootstrap: A fast large-sample alternative to the parametric bootstrap. *Canadian Journal of Statistics*, **40**, 480–500.

MEINTANIS, S. G. AND SWANEPOEL, J. W. H. (2007). Bootstrap goodness-of-fit tests with estimated parameters based on the empirical tranform. *Statistics & Probability Letters*, **77**, 1004–1013.

MEINTANIS, S. G., SWANEPOEL, J. W. H., AND ALLISON, J. S. (2014). The probability weighted characteristic function and goodness–of–fit testing. *Journal of Statistical Planning and Inference*, **146**, 122–132.

Manuscript received, 2015-06-17, revised, 2015-10-05, accepted, 2015-10-05.

COMMENTS: A REVIEW OF TESTING PROCEDURES BASED ON THE EMPIRICAL CHARACTERISTIC FUNCTION

Emanuele Taufer
Department of Economics and Management, University of Trento, Via Inama 5, 38100 Trento - Italy
e-mail: *emanuele.taufer@unitn.it*

The paper nicely reviews testing procedures based on the empirical characteristic function. The discussion touches univariate and multivariate tests, the cases of i.i.d. and dependent observations; the two and k-sample problem. Various technicalities, such as the choice of weights and convergence under the null and under local alternatives are also discussed. In the review several different sources are connected and unfolded in an illuminating way.

In particular the discussion relating literature concerning the asymptotic distributions of the test statistics and the choice of the weights is quite interesting. In my opinion, although goodness-of-fit tests based on the characteristic function have excellent properties and straightforward extensions to the multivariate setting, often, an intractable asymptotic null distribution and the necessity to choose a weight parameter make them hard to apply in practice.

Some solutions in this direction have been proposed by Fernández, Jiménez-Gamero and Castillo Gutiérrez (2014) and Jiménez-Gamero and Hyoung-Moon (2015) by introducing an approximation of the test statistics and a weighted bootstrap as already nicely described by prof. Meintanis.

On the other hand, Henze and Zirkler (1990) and Meintanis, Ngatchou-Wandji and Taufer (2015) have shown, by simulations, the very good agreement, in several cases, of the percentiles of the null distribution with those of a log-normal distribution with parameters determined from the mean and variance of the null distribution of the characteristic function-based test statistic. The approximation already works well for sample sizes as small as 20.

This relationship could be worth of further investigation given the considerable simplification it could bring to application of goodness-of-fit tests based on the characteristic function.

Finally I'd like to congratulate prof. Meintanis for his excellent review of the topic and I would ask him to give a brief overlook about possible extensions and future research directions in the area.

References

FERNÁNDEZ, M. V. A., JIMÉNEZ-GAMERO, M. D., AND CASTILLO GUTIÉRREZ, S. (2014). Approximating a class of goodness-of-fit test statistics. *Mathematics and Computers in Simulation*, **102**, 24–38.

HENZE, N. AND ZIRKLER, B. (1990). A class of invariant consistent tests for multivariate normality. *Communications in Statistics - Theory and Methods*, **19**, 3595–3617.

JIMÉNEZ-GAMERO, M. D. AND HYOUNG-MOON, K. (2015). Fast goodness-of-fit tests based on the characteristic function. *Computational Statistics and Data Analysis*, **89**, 172–191.

MEINTANIS, S. G., NGATCHOU-WANDJI, J., AND TAUFER, E. (2015). Goodness-of-fit tests for multivariate stable distributions based on the empirical characteristic function. *Journal of Multivariate Analysis*, **140**, 171–192.

Manuscript received, 2015-07-31, revised, 2015-10-27, accepted, 2015-10-27.

COMMENTS: A REVIEW OF TESTING PROCEDURES BASED ON THE EMPIRICAL CHARACTERISTIC FUNCTION

Nikolai Ushakov
Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491, Trondheim, Norway
e-mail: *ushakov@math.ntnu.no*

Simos Meintanis presents an interesting and useful survey of testing procedures based on the Empirical Characteristic Function (ECF). The ECF was originally introduced for parameter estimation of a stable law, i.e. in the situation where a description in terms of characteristic functions is much simpler than that in terms of distribution functions. It turned out, however, that there was a wide range of statistical problems where the ECF approach was a competitive alternative for other methods. Hypothesis testing, in particular, goodness-of-fit testing and testing in the two-sample problem, is among these problems, and the survey under discussion gives a lot of helpful information.

In this discussion I review some results of two problems concerning statistical testing based on the ECF. The first problem is testing independence, where the ECF turns out to be a very powerful tool, and where a number of remarkable results have been obtained during the last decade. The second problem concerns goodness-of-fit testing based on the ECF and is connected with the fact that the data are always given in the discretized form and this often must be taken into account.

Testing independence

The problem of testing independence is among those testing problems where the ECF is successfully used and gives many effective possibilities. Let $X_j = (X_{j1},...,X_{jd})$, $1 \le j \le n$, be i.i.d. d-dimensional random variables. Denote the CF of X_j by $\varphi(\mathbf{t})$, $\mathbf{t} = (t_1,...,t_d)$ and the marginal CF of its component $X_{j,k}$ by $\varphi_k(t_k)$, $k = 1,...,d$. The hypothesis of independence of the components can be formulated as

$$H_0:\ \varphi(\mathbf{t}) = \prod_{k=1}^{d} \varphi_k(t_k) \text{ for all } \mathbf{t} = (t_1,...,t_d).$$

Denote the ECF of the sample $X_1,...,X_n$ by $\hat{\varphi}(\mathbf{t})$, $\mathbf{t} \in \mathrm{R}^d$ and the ECF of the sample $X_{1k},...,X_{nk}$ by $\hat{\varphi}_k(t_k)$, $1 \le k \le d$. It is natural to use test statistics based on the difference between $\hat{\varphi}(\mathbf{t})$ and the product $\prod_{k=1}^{d} \hat{\varphi}_k(t_k)$.

The problem of testing independence was among the first applications of the ECF. In this connection one can refer to De Silva and Griffiths (1980), Csörgő and Hall (1982), Csörgő (1985), Feuerverger (1993), Kankainen (1995), Kankainen and Ushakov (1998).

The test, proposed by Csörgő (1985), was based on the test statistic

$$S_n = \sqrt{n}\left(\hat{\varphi}(\mathbf{t}^{(n)}) - \prod_{k=1}^{d} \hat{\varphi}_k(t_k^{(n)})\right)$$

where $\mathbf{t}^{(n)} = (t_1^{(n)},...,t_d^{(n)})$ were random points selected in a special manner. The hypothesis of independence is rejected for large deviations of the test statistic from 0. The limit distribution of this test statistic was obtained under some mild conditions.

Note that in the general case, the considered test is not consistent. Consistent tests were proposed by Feuerverger (1993), Kankainen (1995), Kankainen and Ushakov (1998). In the partial case $d = 2$, two test statistics were proposed by Feuerverger (1993). The first one is

$$T_n^{(1)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{|\tilde{\varphi}(t_1,t_2) - \tilde{\varphi}_1(t_1)\tilde{\varphi}_2(t_2)|^2}{(1-e^{-t_1^2})(1-e^{-t_2^2})} w(t_1,t_2)dt_1dt_2,$$

where $w(t_1,t_2)$ is a weight function and $\tilde{\varphi}$ is a slightly modified version of $\hat{\varphi}$. The second statistic is

$$T_n^{(2)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\hat{\varphi}(t_1,t_2) - \hat{\varphi}_1(t_1)\hat{\varphi}_2(t_2)|^2 w(t_1,t_2)dt_1dt_2.$$

For arbitrary d, tests based on the test statistic of the form

$$\int_{\mathrm{R}^d} |\hat{\varphi}(\mathbf{t}) - \prod_{k=1}^{d} \hat{\varphi}_k(t_k)|^2 w(\mathbf{t})d\mathbf{t}$$

were studied in details by Kankainen (1995). See also Hušková and Meintanis (2008) for a brief review of early papers concerning application of the ECF to testing independence.

Bilodeau and Lafaye de Micheaux (2005) used the ECF for testing independence (or serial independence) between marginal vectors each of which was normally distributed. These marginal subvectors are not assumed to be jointly multinormal, therefore independence cannot be tested parametrically using covariances. The authors propose a test which is based on the test statistic of the Cramer-von Mises type. The test is consistent to detect any form of dependence.

Meintanis and Iliopoulos (2008) develop a class of tests for testing multivariate independence and study their small-sample performance using simulation. The test statistics employ the familiar equation between the joint characteristic function and the product of component characteristic functions, and may be written in a closed form convenient for computer implementation. Simulations on a distribution-free version of the new test statistic show that the proposed method compares well to standard methods of testing independence via the empirical distribution function. The methods are applied to multivariate observations incorporating data from several major stock-market indices.

An essential advance in the use of the ECF for testing independence has been made in a series of articles by G. J. Székely and M. L. Rizzo. Székely, Rizzo and Bakirov (2007) defined the distance covariance and the distance correlation between two random vectors and used them for testing independence. Distance covariance and correlation characterize independence, so they equal zero if and only if the random vectors are independent. They are applicable to random vectors of arbitrary and not necessarily equal dimension, but first we consider for simplicity the univariate case. Let X and Y be random variables with characteristic functions $\varphi_X(t)$, $\varphi_Y(s)$ and the joint characteristic function $\varphi_{X,Y}(t,s)$. The distance covariance $\mathscr{V}(X,Y)$ between X and Y is defined as the square root of the integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{|\varphi_{X,Y}(t,s) - \varphi_X(t)\varphi_Y(s)|^2}{|ts|^2} dtds.$$

The distance variance is defined as $\mathscr{V}(X) = \mathscr{V}(X,X)$. The distance correlation is

$$\mathscr{R}(X,Y) = \frac{\mathscr{V}(X,Y)}{\sqrt{\mathscr{V}(X)\mathscr{V}(Y)}}.$$

Suppose now that there is a random sample $(X_1,Y_1),...,(X_n,Y_n)$ — independent copies of the random vector (X,Y). The empirical distance correlation (or the distance correlation statistic) of this sample $\mathscr{R}_n = \mathscr{R}_n(X_1,Y_1,...X_n,Y_n)$ is defined in the same way as the theoretical distance correlation, where characteristic functions are replaced by the corresponding empirical characteristic functions. The test, proposed by Székely et al. (2007), rejects the hypothesis of independence of X and Y for large values of $\mathscr{R}_n(X,Y)$. Asymptotic critical values are obtained. The test is consistent against all dependent alternatives. A Monte Carlo study shows that the proposed test may be more powerful than the likelihood ratio test. In the normal case, they are quite close in power.

The method was further developed in Székely and Rizzo (2009a), where the notion of covariance with respect to a stochastic process was introduced. It was shown that the distance covariance coincided with the covariance with respect to the Brownian motion. This gives additional computational possibilities and simplifications.

Consider now the multivariate case. For random vectors X, Y of arbitrary dimensions p, q, the distance covariance and correlation are defined and used in a similar way as in the univariate case, just the denominator $|ts|^2$ under the integral sign is replaced by $|t|_p^{1+p}|s|_q^{1+q}$, where $|\cdot|_d$ is the Euclidean distance in R^d. The extension of the distance correlation to the problem of testing independence of random vectors in arbitrary high, not necessarily equal dimensions was developed by Székely and Rizzo (2013). Suppose that there is a sample $(X_1,Y_1),...,(X_n,Y_n)$ — independent copies of the $(p+q)$-dimensional random vector (X,Y). The authors define the modified distance correlation statistic $\mathscr{R}_n^*$ — a certain modification of $\mathscr{R}_n$. The main result is that as p, q tend to infinity, under the independence hypothesis,

$$\mathscr{T}_n = \sqrt{\nu - 1} \cdot \frac{\mathscr{R}_n^*}{\sqrt{1 - (\mathscr{R}_n^*)^2}}$$

converges in distribution to Student t with $\nu - 1$ degrees of freedom, where

$$\nu = \frac{n(n-3)}{2}.$$

For $n \geq 10$ this limit is approximately standard normal. Using this, the authors obtain a distance correlation t-test for independence of random vectors in arbitrary high dimension. The test is unbiased for every sample size greater than three and all significance levels. The developed method is applied to testing independence of two time series.

A further development of the concept of the distance covariance and distance correlation was made in Székely and Rizzo (2014). Here partial distance covariance and partial distance correlation are introduced as measures of dependence of two random vectors X, Y, controlling for a third random vector Z, where X, Y and Z are in arbitrary, not necessarily equal dimensions. A test for zero partial distance correlation is proposed.

Binning

It also should be mentioned the problem of binning in goodness-of-fit tests based on ECF. Binning is a special technique for reduction of computational expenses, which consists of replacing the original observations by the prebinned data: each observed data value is distributed (with some weights, possibly negative) among grid points on an equally spaced mesh. Binning is an effective tool in situations when direct operation with the original sample leads to high computational expenses, that is typical in particular for bootstrap procedures. Meintanis and Ushakov (2004) studied binned goodness-of-fit tests based on the ECF. Let $X_1,...,X_n$ be i.i.d. random variables drawn from a location-scale family with parameters α, β and the characteristic function $\varphi(t;\alpha,\beta)$. Given a bin width $\delta > 0$ and bin origin x_0, the binned ECF is defined as

$$\hat{\varphi}_n(t) = \frac{1}{n}\sum_{k=-\infty}^{\infty} N_k e^{i(x_0+k\delta)t},$$

where N_k is the number of observations in the interval $[x_0+k\delta-\delta/2, x_0+k\delta+\delta/2]$.

Let $Y_1,...,Y_n$ be normalized data

$$Y_j = \frac{X_j-\hat{\alpha}}{\hat{\beta}},$$

where $\hat{\alpha}$, $\hat{\beta}$ are consistent estimators of α, β. Denote the binned ECF of the normalized data by $\hat{\psi}_n(t)$. Meintanis and Ushakov (2004) used two types of binned test statistics

$$D_n^\delta = \sqrt{n}\sup_{|t|\le\tau}|\hat{\psi}_n(t)-\varphi(t;0,1)|,$$

$$T_n^\delta = n\int_{-\infty}^{\infty}|\hat{\psi}_n(t)-\varphi(t;0,1)|w(t)dt.$$

It was shown (under some conditions on τ and $w(t)$) that the limiting null distributions of these statistics coincided with those of the ordinary unbinned statistics of the same form. If $\delta = O(n^{-1/2})$, then the consistency of the unbinned tests imply consistency of the corresponding binned tests. Using a simulation experiment, Meintanis and Ushakov (2004) demonstrated that the binned versions of the tests did not essentially lose power compared with their counterparts.

References

BILODEAU, M. AND LAFAYE DE MICHEAUX, P. (2005). A multivariate empirical characteristic function test of independence with normal marginals. *Journal of Multivariate Analysis*, **95**, 345–369.

CSÖRGŐ, S. (1985). Testing for independence by the empirical characteristic function. *Journal of Multivariate Analysis*, **16**, 290–299.

CSÖRGŐ, S. AND HALL, P. (1982). Estimable versions of Griffiths' measure of association. *Australian & New Zealand Journal of Statistics*, **24**, 296–308.

DE SILVA, B. M. AND GRIFFITHS, C. R. (1980). A test of independence for bivariate symmetric stable distributions. *Australian & New Zealand Journal of Statistics*, **22**, 172–177.

FEUERVERGER, A. (1993). A consistent test for bivariate dependence. *International Statistical Review*, **61**, 419–433.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2008). Testing procedures based on the empirical characteristic function I: goodness-of-fit, testing for symmetry and independence. *Tatra Mountains Mathematical Publications*, **39**, 225–233.

KANKAINEN, A. (1995). *Consistent Testing of Total Independence Based on the Empirical Characteristic Fucntion*. Phd thesis, University of Jyväskylä.

KANKAINEN, A. AND USHAKOV, N. G. (1998). A consistent modification of a test for independence based on the empirical characteristic function. *Journal of Mathematical Sciences* (New York) , **89**, 1486–1493.

MEINTANIS, S. AND USHAKOV, N. G. (2004). Binned goodness-of-fit tests based on the empirical characteristic function. *Statistics & Probability Letters*, **69**, 305–314.

MEINTANIS, S. G. AND ILIOPOULOS, G. (2008). Fourier methods for testing multivariate independence. *Computational Statistics & Data Analysis*, **52**, 1884–1895.

SZÉKELY, G. J. AND RIZZO, M. L. (2009a). Brownian distance covariance. *Annals of Applied Statistics*, **3**, 1236–1265.

SZÉKELY, G. J. AND RIZZO, M. L. (2009b). Rejoinder: Brownian distance covariance. *Annals of Applied Statistics*, **3**, 1302–1308.

SZÉKELY, G. J. AND RIZZO, M. L. (2013). The distance correlation t-test of independence in high dimensions. *Journal of Multivariate Analysis*, **117**, 193–213.

SZÉKELY, G. J. AND RIZZO, M. L. (2014). Partial distance correlation with methods for dissimilarities. *Annals of Statistics*, **42**, 2382–2412.

SZÉKELY, G. J., RIZZO, M. L., AND BAKIROV, N. K. (2007). Measuring and testing dependence by correlation of distances. *Annals of Statistics*, **35**, 2769–2794.

Manuscript received, 2015-07-31, revised, 2015-10-06, accepted, 2015-10-06.

REJOINDER ON: A REVIEW OF TESTING PROCEDURES BASED ON THE EMPIRICAL CHARACTERISTIC FUNCTION

Simos G. Meintanis [1]
National and Kapodistrian University of Athens, Department of Economics, 1 Sophocleous Str., and Aristidou Str., 105 59 Athens, Greece;
Unit for Business Mathematics and Informatics, North-West University, Potchefstroom, South Africa
e-mail: *simosmei@econ.uoa.gr*

I would like to thank the Editor, Prof. Sarah Radloff, for giving me the opportunity to write this review article. A further word of thanks is addressed to all discussants. Some of their comments nicely complement this review on subjects that were not addressed and hence require a round of extra thanks. In what follows, and in alphabetic order, I will try to address the comments of the discussants.

1. The BHEP test and 'precise' hypothesis

I wish to thank Norbert Henze for bringing up the normal limit in (1) to our attention. Although I am aware of Nora Gürtler's thesis, I am not familiar with the details of its content (it is in German anyway). In this connection, I just wish to point out that such a limit also appears in Meintanis and Swanepoel (2007) and Henze and Meintanis (2010). In fact, and apart from the nice power approximation noted by Prof. Henze, Henze and Meintanis (2010) use L_β as a descriptive measure of power and compare the relative performance of an ECF-based test for exponentiality against two fixed alternative distributions, say F_1 and F_2, by considering the corresponding curves $L_{F_1,\beta}$ and $L_{F_2,\beta}$, over an interval of β-values.

The second comment of Prof. Henze is of an even wider interest. It refers to what Berger and Delampady (1987) call the principle of *precise hypothesis* about a parameter. In the more familiar current context of GOF testing it may also be found in Dette and Munk (1998) where this principle is applied to an L_2-type distance procedure with an asymptotic normal law; see also Borovkov (1998), §49 & §55, and Lavergne (2014). I am not aware of work related to this approach based on the ECF. In this regard, it would be interesting to consider all aspects involved with such a decision rule, i.e., dependence of power on the values of Δ and β, and performance with different distributions under test, as well as its applicability for alternative testing criteria, besides those based on the ECF.

[1] On sabbatical leave from the University of Athens.

2. Change-point detection

I am pleased for the discussion of Marie Hušková which complements this review by considering ECF-methods for the related problem with change-point detection. Certainly there is room for much more work in this context. For instance one may combine the two-sample test of Hušková and Meintanis (2008) to the change-point procedures of Hušková and Meintanis (2006a, 2006b), in order to construct a detection method for the two-sample change-point problem, a problem which is relevant in the context of several applied circumstances; see for instance Hlávka and Hušková (2015). I just wish to add that the interesting aspect of decoupling location from scale changes is addressed by Steland and Rafajlowicz (2014), and the work of Matteson and James (2014) for retrospective multiple change-point analysis with multivariate observations.

Another issue of wide interest is testing the so-called *martingale difference hypothesis* (MDH). The standard formulation of the MDH is

$$\mathbb{E}(Y_t|\mathbb{I}_{t-1}) = 0,\ t = 1,\ldots, \tag{1}$$

where $\mathbb{I}_t$ denotes the information set available at time t, and Y_t represents first differences of a process which under this hypothesis forms a martingale sequence. We will not go into detail here but simply mention that there exists a Fourier formulation of the MDH initially put forward by Bierens (1982) which closely relates to ECF-based methods. Following this approach Hlávka, Hušková, Kirch and Meintanis (2014) consider test procedures for the MDH as well as several change-point problems associated with the MDH.

Finally, in the context of time series of counts, it should be pointed-out that natural precursors of the procedures mentioned in the discussion of Prof. Hušková are the GOF procedures of Meintanis and Karlis (2014) and Hudecová, Hušková and Meintanis (2015), both of which are based on the empirical probability generating function.

3. Resampling and the probability weighted ECF

It should be said from the start that I always favoured the parametric bootstrap (PB) for actually performing ECF-based tests involving unknown parameters. So my discussion on the issue of weighted bootstrap (WB) vs. PB resampling raised by Jan Swanepoel and James Allison might be a bit biased. One reason for this bias is that with the power of present day computers, CPU time is often not a problem. Moreover there is the new suggestion of the warp-speed method by Giacomini et al. (2013) which essentially suggests ‘one bootstrap for each Monte Carlo’ and hence drastically reduces the computational cost of calculating the estimator for each bootstrap iteration. While the current theory of the warp-speed method is restricted to the estimation of a single parameter, Dalla et al. (2016) and Francq et al. (2015), applied the warp-speed method to much more complicated situations and to a good effect, which implies wider applicability of the procedure. Regarding the comment on estimation of the parameter involved, the issue with the WB as I see it is not so much the asymptotic linear representation, which is a common assumption in the PB also, but the fact that with the WB one actually needs to compute $l(x;\theta)$ and then consistently estimate it, which is not the case in the PB. With respect to this, Jiménez-Gamero and Kim (2015) carry out this computation for several standard estimators, but the problem still remains with other estimators particularly in more complex situations. The effect of the choice of the standardized variables $\zeta_j,\ j = 1,...,n$, on the properties of the WB implemented test also needs to be addressed, and to that end Jiménez-Gamero and Kim (2015) are not very helpful as they do not mention, at least to my reading, which sequence they actually employed in their Monte Catlo study. In a related context however, Ghoudi and Rémillard (2014) settle for standard normal ζ_j and mention nothing about other choices. One might guess that the impact on test performance is minimal, but then this should somehow also be verified.

At this point I wish to thank Profs. Swanepoel and Allison for bringing up the idea of the probability weighted ECF (PWECF). As they mention, potential statistical applications of the PWECF extend far beyond the original strict parametric context of Meintanis, Swanepoel and Allison (2014) In this direction, Meintanis and Stupfler (2015), Meintanis and Ushakov (2016), and Meintanis, Allison and Santana (2016), present several results for the PWECF, and apply this tool in a variety of semiparametric, nonparametric problems, and parametric problems with complex data structures, including some of those mentioned by Swanepoel and Allison.

4. Testing for independence and grouped data

Nikolai Ushakov provides a thorough review of testing procedures for independence, an area where along with testing for symmetry ECF methods enjoy a definite competitive advantage over other more standard procedures. The review of Prof. Ushakov is pretty much complete. I just wish to add the work of Hlávka, Hušková and Meintanis (2011) on testing independence in the context of nonparametric regression, and the upcoming paper of Leucht, Strauch and Meintanis (2015) which applies the PWECF for testing independence of several random vectors. The issue raised with binned or groupd data is also of interest in some applications. I am not aware of any progress in this direction involving ECF-based procedures. The sole exception appears to be the extention of the BHEP test for normality to rank set sampling data investigated by Balakrishnan et al. (2013).

5. Outlook and conclusion

I guess Prof. Emanuele Taufer will find much information regarding possible extensions and future research directions in my responses to the other discussants. In general, the ECF may be applied to several other statistical problems, besides GOF. Estimation for instance is another broad area of application of the ECF, and I refer to Braun et al. (2008), Xu and Knight (2011), Wang et al. (2012), Potgieter and Genton (2013), Kotchoni (2014), Francq and Meintanis (2015), the PhD thesis of Mofei Jia (2014), and references therein, for a variety of ECF-based estimation methods under various data generating mechanisms.

In closing, I would like to express the hope that this review will help researchers to take an interest in this type of methodology, and thus stimulate further developments in this area either in a similar context as that considered herein or in any other direction.

Acknowledgement

This research was partially supported by grant Nr.11699 of the Special Account for Research Grants (ELKE) of the National and Kapodistrian University of Athens.

References

BALAKRISHNAN, N., BRITO, M. R., AND QUIROZ, A. J. (2013). On the goodness-of-fit procedure for normality based on the empirical characteristic function for rank set sampling data. *Metrika*, **76**, 161–177.

BERGER, J. O. AND DELAMPADY, M. (1987). Testing precise hypotheses. (with discussion). *Statistical Science*, **2**, 317–352.

BIERENS, H. J. (1982). Consistent model specification tests. *Journal of Econometrics*, **20**, 105–134.

BOROVKOV, A. A. (1998). *Mathematical Statistics*. Gordon & Breach.

BRAUN, M. P., MEINTANIS, S. G., AND MELAS, V. B. (2008). Optimal design approach for GMM estimation of parameters based on empirical transforms. *International Statistical Review*, **76**, 387–400.

DALLA, V., BASSIAKOS, Y., AND MEINTANIS, S. G. (2016). Characteristic function-based estimation for GARCH models with heavy-tailed innovations. *Communications in Statistics – Simulation and Computation*. DOI: 10.1080/03610918.2015.1060332.

DETTE, H. AND MUNK, A. (1998). Validation of linear regression models. *Annals of Statistics*, **26**, 778–800.

FRANCQ, C. AND MEINTANIS, S. G. (2015). Fourier-type estimation of the power GARCH model with stable Paretian innovations. *Metrika*. DOI: 10.1007/s00184-015-0560-x.

FRANCQ, C., MEINTANIS, S. G., AND JIMÉNEZ-GAMERO, M. D. (2015). Tests for sphericity in multivariate GARCH models. Forthcoming.

GHOUDI, K. AND RÉMILLARD, B. (2014). Comparison of specification tests for GARCH models. *Computational Statistics & Data Analysis*, **76**, 297–300.

GIACOMINI, R., POLITIS, D. N., AND WHITE, H. (2013). A warp-speed method for conducting monte carlo experiments involving bootstrap estimators. *Econometric Theory*, **29**, 567–589.

HENZE, N. AND MEINTANIS, S. G. (2010). A characterization and a class of omnibus tests for the exponential distribution based on the empirical characteristic function. *Journal of Mathematical Sciences*, **167**, 588–595.

HLÁVKA, Z. AND HUŠKOVÁ, M. (2015). Two-sample gradual change-point analysis. *PROBASTAT 2015, Book of Abstracts*, p. 27.

HLÁVKA, Z., HUŠKOVÁ, M., KIRCH, C., AND MEINTANIS, S. G. (2014). Fourier-type tests involving martingale difference processes. *Econometric Reviews*. DOI: 10.1080/07474938.2014.977074.

HLÁVKA, Z., HUŠKOVÁ, M., AND MEINTANIS, S. G. (2011). Tests for independence in nonparametric heteroscedastic regression models. *Journal of Multivariate Analysis*, **102**, 816–827.

HUDECOVÁ, S., HUŠKOVÁ, M., AND MEINTANIS, S. G. (2015). Tests for time series of counts based on the probability-generating function. *Statistics*, **49**, 316–337.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2006a). Change point analysis based on empirical characteristic functions. *Metrika*, **63**, 145–168.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2006b). Change point analysis based on empirical characteristic functions of ranks. *Sequential Analysis*, **25**, 421–436.

HUŠKOVÁ, M. AND MEINTANIS, S. G. (2008). Tests for the multivariate *k*-sample problem based on the empirical characteristic function. *Journal of Nonparametric Statistics*, **20**, 263–277.

JIA, M. (2014). *Heavy-tailed Phenomena and Tail Index Inference*. PhD thesis, University of Trento.

JIMÉNEZ-GAMERO, M. D. AND KIM, H. M. (2015). Fast goodness-of-fit tests based on the characteristic function. *Computational Statistics & Data Analysis*, **89**, 172–191.

KOTCHONI, R. (2014). The indirect continuous GMM estimation. *Computational Statistics & Data Analysis*, **76**, 464–488.

LAVERGNE, P. (2014). Model equivalence tests in a parametric framework. *Journal of Econometrics*, **178**, 414–425.

LEUCHT, A., STRAUCH, C., AND MEINTANIS, S. G. (2015). Tests for independence based on the multivariate probability weighted characteristic function. Forthcoming.

MATTESON, D. S. AND JAMES, N. A. (2014). A nonparametric approach for multiple change point analysis of multivariate data. *Journal of the American Statistical Association*, **109**, 334–345.

MEINTANIS, S. G., ALLISON, J., AND SANTANA, L. (2016). Goodness-of-fit tests for semiparametric and parametric hypotheses based on the probability weighted characteristic function. *Statistical Papers*. Forthcoming.

MEINTANIS, S. G. AND KARLIS, D. (2014). Validation tests for the innovation distribution in INAR time series models. *Computational Statistics*, **29**, 1221–1241.

MEINTANIS, S. G. AND STUPFLER, G. (2015). Transformation to symmetry based on the probability weighted characteristic function. *Kybernetika*, **51**, 571–587.

MEINTANIS, S. G. AND SWANEPOEL, J. W. H. (2007). Bootstrap goodness-of-fit tests with estimated parameters based on empirical transforms. *Statistics & Probability Letters*, **77**, 1004–1013.

MEINTANIS, S. G., SWANEPOEL, J. W. H., AND ALLISON, J. S. (2014). The probability weighted characteristic function and goodness-of-fit testing. *Journal of Statistical Planning and Inference*, **146**, 122–132.

MEINTANIS, S. G. AND USHAKOV, N. (2016). Nonparametric empirical probability weighted characteristic function and applications. *Statistics & Probability Letters*, **108**, 52–61.

POTGIETER, C. J. AND GENTON, M. G. (2013). Characteristic function based semiparametric inference for skew-symmetric models. *Scandinavian Journal of Statistics*, **40**, 471–490.

STELAND, A. AND RAFAJLOWICZ, E. (2014). Decoupling change-point detection based on characteristic functions: Methodology, asymptotics, subsampling and application. *Journal of Statistical Planning and Inference*, **145**, 49–73.

WANG, D., MIECZNIKOWSKI, J. C., AND HUSTON, A. D. (2012). Direct density estimation of *l*-estimates via characteristic functions with applications. *Journal of Statistical Planning and Inference*, **142**, 567–578.

XU, D. AND KNIGHT, J. (2011). Continuous empirical characteristic function estimation of mixtures of normal parameters. *Econometric Reviews*, **30**, 25–50.

Manuscript received, 2015-09-31, revised, 2015-10-06, accepted, 2015-10-06.