
ELSEVIER
Volume 143 Issue 1 March 2008 ISSN 0304-4076
JOURNAL OF
Econometrics
SPECIFICATION TESTING
edited by
MIGUEL A. DELGADO
Editors:
TAKESHI AMEMIYA
A. RONALD GALLANT
JOHN F. GEWEKE
CHENG HSIAO
PETER ROBINSON
ARNOLD ZELLNER
Associate Editors:
YACINE AIT-SAHALIA
BADI H. BALTAGI
R. BANSAL
MICHAEL W. BRANDT
MARCUS J. CHAMBERS
SONGNIAN CHEN
VICTOR CHERNOZHUKOV
MANFRED DEISTLER
MIGUEL A. DELGADO
JEAN-MARIE DUFOUR
SYLVIA FRUHWIRTH-
-SCHNATTER
ERIC GHYSELS
JOHN C. HAM
JAVIER HIDALGO
HAN HONG
YONGMIAO HONG
BO E. HONORÉ
MAXWELL L. KING
YUICHI KITAMURA
GARY M. KOOP
CHUNG-MING KUAN
NAOTO KUNITOMO
KAJAL LAHIRI
QI. LI
TONG LI
ROSA L. MATZKIN
HYUNGSIK R. MOON
FRANZ C. PALM
DALE J. POIRIER
BENEDIKT PÖTSCHER
INGMAR PRUCHA
PETER C. REISS
ERIC RENAULT
FRANK SCHORFHEIDE
ROBIN SICKLES
FALLAW SOWELL
GERARD VAN DEN BERG
HERMAN VAN DIJK
QUANG H. VUONG
EDWARD VYTLACIL
TOM WANSBEEK
TAO ZHA

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Journal of Econometrics 143 (2008) 5–18
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

On distribution-free goodness-of-fit testing of exponentiality

John Haywood, Estate Khmaladze*

School of Mathematics, Statistics and Computer Science, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand

Available online 19 August 2007

Abstract

There is a need to test the hypothesis of exponentiality against a wide variety of alternative hypotheses, across many areas of economics and finance. Local or contiguous alternatives are the closest alternatives against which it is still possible to have some power. Hence goodness-of-fit tests should have some power against all, or a huge majority, of local alternatives. Such tests are often based on nonlinear statistics, with a complicated asymptotic null distribution. Thus a second desirable property of a goodness-of-fit test is that its statistic will be asymptotically distribution free. We suggest a whole class of goodness-of-fit tests with both of these properties, by constructing a new version of empirical process that weakly converges to a standard Brownian motion under the hypothesis of exponentiality. All statistics based on this process will asymptotically behave as statistics from a standard Brownian motion and so will be asymptotically distribution free. We show the form of transformation is especially simple in the case of exponentiality. Surprisingly there are only two asymptotically distribution free versions of empirical process for this problem, and only this one has a convenient limit distribution. Many tests of exponentiality have been suggested based on asymptotically linear functionals from the empirical process. We illustrate none of these can be used as goodness-of-fit tests, contrary to some previous recommendations. Of considerable interest is that a selection of well-known statistics all lead to the same test asymptotically, with negligible asymptotic power against a great majority of local alternatives. Finally, we present an extension of our approach that solves the problem of multiple testing, both for exponentiality and for other, more general hypotheses.

JEL classification: C12; C46; C52

Keywords: Asymptotically linear statistics; Brownian motion; Contiguous alternatives; Distribution free statistics; Goodness-of-fit statistics; Local alternatives; Rate of convergence to Brownian motion; Transformed empirical process

1. Introduction

Testing exponentiality is a long standing problem in statistics. The exponential distribution is needed and used at least as often as the time-homogeneous Poisson process. Applications of the exponential distribution in physics, and more so in survival analysis and reliability theory, are numerous and diverse (e.g., Pham, 2003).

In economics, the assumption of exponentiality has been widely applied. For example, equilibrium search models of the labour market typically use exponential distributions for the durations of unemployment and employment within a particular job (e.g., Van den Berg and Ridder, 1998, who also allow for unobserved

*Corresponding author. Tel.: +64 4 463 5652; fax: +64 4 463 5045.
E-mail addresses: John.Haywood@mcs.vuw.ac.nz (J. Haywood), Estate.Khmaladze@mcs.vuw.ac.nz (E. Khmaladze).

0304-4076/$ - see front matter
doi:10.1016/j.jeconom.2007.08.005

heterogeneity). The exponential distribution of the duration of typical stock market drawdowns (persistent decreases in price over consecutive days) is motivated and discussed by Sornette (2003, Chapter 3), while Granger and Ding (1995) suggest and assess the use of an exponential distribution for the absolute value of mean-corrected stock returns. Evidence has been presented of exponential distributions for wealth and income in the UK and the US (Dragulescu and Yakovenko, 2001a, b), and Stanley et al. (1996) propose and fit a model in which (log) annual growth rates of companies have an exponential distribution. Thus there is a clear need to test the hypothesis of exponentiality against a wide variety of alternative hypotheses, across many areas of economics and finance.

Goodness-of-fit tests are typically based on omnibus statistics, designed to detect any deviation from the hypothesis of interest. Well-known examples include the Kolmogorov–Smirnov statistics (one- and two-sided versions), the Cramér–von Mises statistic and the Anderson–Darling statistic. All these statistics are nonlinear, with complicated asymptotic null distributions. Following estimation of any unknown parameters, those distributions depend on the hypothesised parametric family, the value of the parameter and the estimator used (Durbin, 1973; Khmaladze, 1979). Hence the test statistics are not asymptotically distribution free.

In this paper we do not advocate any particular distribution-free test. Instead we present a modification of the usual empirical process, from which a variety of distribution-free test statistics can be calculated. Under the hypothesis of exponentiality, the required transformation of the empirical process is particularly straightforward. Any test statistic from this process asymptotically behaves as a statistic from a standard Brownian motion, which makes asymptotic theory convenient and relatively simple (Khmaladze, 1981, 1993). Despite the prevalence of tests for exponentiality, the process described here is one of only two asymptotically distribution free empirical processes for this problem, as far as we know; see our discussion of Angus (1982) and Koul (1977, 1978) in Section 2. Furthermore, our process is the only one with a convenient limit distribution.

A widely used counterpart of goodness-of-fit statistics is asymptotically linear or directional statistics, many of which are score tests. These statistics have a nature very different from omnibus statistics: they lead to asymptotically optimal tests for certain deviations from the hypothetical family, but have very low or no power against the huge majority of other deviations. It might be expected that in a problem like testing exponentiality, there exist many directional statistics able to detect deviations from the hypothesised family in many different directions. It was therefore surprising to discover that a selection of well-known statistics, some of which have been studied and recommended for goodness-of-fit testing of exponentiality, are so closely related that practically they lead to the same test. Thus the lack of power against many alternatives will be shared by all these linear statistics, which clearly questions previous recommendations for goodness-of-fit testing using any of them.

Whenever two or more test statistics are used for the same problem, the issue of multiple testing arises. In principle it is the conditional asymptotic distribution of a statistic that is the correct thing to study in such cases, but in practice marginal asymptotic distributions are commonly used instead. We consider this problem briefly, and show that a transformed version of the empirical process can be constructed which is still distribution free but is also asymptotically independent from any linear statistics already used. This approach obviates the need for any conservative adjustments to test sizes that are commonly used in multiple testing, such as Bonferroni corrections.

In Section 2 we define our problem formally and present the transformed version of empirical process which is distribution free. We demonstrate good finite sample convergence properties of our modified process in Section 3 through convergence properties of several representative test statistics. Section 4 looks at a selection of well-known directional statistics and demonstrates their asymptotic similarity, along with some examples of the use of statistics from our modified empirical process. Finally Section 5 briefly considers a resolution of the multiple testing problem using our transformed empirical process.

2. Distribution-free tests and a transformed version of empirical process

Given a parametric family of distribution functions $\mathscr{P} = \{P_\theta, \theta \in \Theta \subset \mathbb{R}^d\}$ indexed by a finite-dimensional parameter θ, and given a random sample $X_1, \ldots, X_n$ from some unknown distribution function P, suppose we want to test the null hypothesis that $P \in \mathscr{P}$. In the case of exponentiality $P_\theta(x) = 1 - \exp(-\theta x)$ and $\theta > 0$ is one-dimensional. Goodness-of-fit statistics are often chosen as functionals of omnibus nature from the parametric, or estimated, empirical process:

$$v_n(x, \mathscr{P}) = \hat{v}_n(x) = \sqrt{n}\{P_n(x) - P_{\hat{\theta}}(x)\},$$

where $P_n(x)$ denotes an empirical distribution function based on the sample and $P_{\hat{\theta}}$ is the distribution function from $\mathscr{P}$ which fits the sample 'best'. There is no exact definition of goodness-of-fit test statistics, yet such statistics are supposed to be able to detect 'all sorts of deviations' from the hypothesis of interest. One can take the Kolmogorov–Smirnov statistic,

$$\sup_x |\hat{v}_n(x)| = \sqrt{n} \sup_x |P_n(x) - P_{\hat{\theta}}(x)|,$$

as an example of such a statistic.

The empirical process $v_n(x) = \sqrt{n}\{P_n(x) - P(x)\}$, with fixed hypothetical P, converges in distribution to a standard Brownian bridge under the hypothesis, after time transformation $t = P(x)$. In contrast, the estimated empirical process $\hat{v}_n$ converges to a Gaussian process with distribution that very much depends on the parametric family $\mathscr{P}$, on the estimator used and, in general, on the particular values of the parameter; see, e.g., Durbin (1973) and Khmaladze (1979) and references therein. On the other hand, the process $\hat{v}_n$ asymptotically 'preserves all information' available in the sample to discriminate between the hypothesis and any local (contiguous) alternatives. The process w_n we construct below will have this latter property, but will be asymptotically distribution free; for a formal proof of this statement see Khmaladze (1981, 1993).

As the class of alternative distributions we consider the class of all 'local', or better, converging contiguous alternatives, which are defined as the class of all sequences of distributions A_n, for which there exists $P_\theta \in \mathscr{P}$ and the function $h(\cdot, \theta) \in L_2(P_\theta)$ such that

$$\left(\frac{\mathrm{d}A_n}{\mathrm{d}P_\theta}\right)^{1/2}(x) = 1 + \frac{1}{2\sqrt{n}} h(x, \theta) + r_n(x), \quad n \int r_n^2(x)\, \mathrm{d}P_\theta(x) \to 0. \tag{1}$$

The function h describes the 'direction' from which alternatives A_n converge to a particular sub-hypothesis P_θ. Hence h is the score function of the locally most powerful test for testing P_θ against A_n. Thus the statistic of this test is the sum $\sum_{i=1}^n h(X_i, \theta)$; cf. the comments on the Moran and Cox–Oakes statistics in Section 4.

From a geometric point of view, it is clear that one should not be concerned to detect alternatives which approach any particular P_θ from the direction tangential to $\mathscr{P}$ at P_θ. This direction is given by the score function under the null hypothesis: $\partial \log p(x, \theta)/\partial\theta$, where $\{p(x, \theta), \theta \in \Theta\}$ is the family of density functions corresponding to $\mathscr{P}$ (e.g., see Janssen, 2000). More formally, in (1) the function h must be orthogonal to this direction, or

$$\int h(x, \theta) \frac{\partial \log p(x, \theta)}{\partial \theta}\, \mathrm{d}P_\theta(x) = 0.$$

Indeed, any part of A_n described by the part of h co-linear to $\partial \log p(\cdot, \theta)/\partial\theta$ can be explained as an infinitesimal shift in the value of θ rather than a deviation from the family $\mathscr{P}$.

Although the Kolmogorov–Smirnov test was mentioned explicitly above, we are not suggesting or advocating any particular test. Instead we propose a version of parametric empirical process w_n, different from $\hat{v}_n$. After the time transformation

$$t = P_{\hat{\theta}}(x), \tag{2}$$

the process w_n converges in distribution under the hypothesis to standard Brownian motion and converges to standard Brownian motion plus non-zero shift under any sequence of contiguous alternatives (1). Therefore any statistic that is invariant under the time transformation (2) is asymptotically distribution free, and any statistic of omnibus nature, like the Kolmogorov–Smirnov statistic or the Cramér–von Mises or Anderson–Darling statistics, has asymptotic power against any local alternatives.

The process w_n differs from $\hat{v}_n$ in the way one centres the empirical distribution function P_n. This different centring was first presented in Khmaladze (1981) but for convenience we include a brief description here. Instead of subtracting from P_n its expected value P_θ with estimated parameter, or, in differential form, instead of considering $\hat{v}_n(\mathrm{d}x) = \sqrt{n}\{P_n(\mathrm{d}x) - P_{\hat{\theta}}(\mathrm{d}x)\}$, one can use conditional expected values and consider

$$\sqrt{n}\{P_n(\mathrm{d}x) - E(P_n(\mathrm{d}x)|P_n(y), y \leqslant x, \hat{\theta})\}.$$

Centring of each increment, $P_n(\mathrm{d}x)$, by the conditional expected value renders w_n to be a martingale with respect to the filtration $\{\mathscr{F}_x, x \geqslant 0\}$, where each σ-algebra $\mathscr{F}_x$ is generated by the 'past' of P_n and the maximum likelihood estimator $\hat{\theta}$, i.e., $\mathscr{F}_x = \sigma\{P_n(y), y \leqslant x, \hat{\theta}\}$. It may be difficult to calculate this conditional expectation, but it is easy to calculate the asymptotically equivalent expression $K(\mathrm{d}x, P_n)$, such that

$$w_n(\mathrm{d}x) = \sqrt{n}\{P_n(\mathrm{d}x) - K(\mathrm{d}x, P_n)\} \tag{3}$$

will be a process with uncorrelated increments. As the notation suggests, the centring process $K(x, P_n)$, $x \geqslant 0$, is a modification of $P_{\hat{\theta}}$. It is called the compensator of P_n—it 'compensates' P_n to a process with uncorrelated increments. The form of $K(x, P_n)$ is given by the hypothetical family of distribution functions as follows. Consider the vector function

$$q(x, \theta)^{\mathrm{T}} = \left[1, \frac{\partial \log p(x, \theta)}{\partial \theta}\right], \tag{4}$$

with the first coordinate identically equal to 1 and the second coordinate being the score function $\partial \log p(x, \theta)/\partial \theta$. Using this extended score function construct the matrix

$$C(z, \theta) = \int_z^\infty q(x, \theta) q(x, \theta)^{\mathrm{T}} p(x, \theta)\, \mathrm{d}x, \tag{5}$$

which can be called the incomplete Fisher information matrix. Then the compensator associated with this parametric family is defined (Khmaladze, 1981) as

$$\begin{aligned} K(x, P_n) &= \int_{-\infty}^{x} q(z, \theta)^{\mathrm{T}} C^{-1}(z, \theta) \int_z^\infty q(y, \theta) P_n(\mathrm{d}y) P(\mathrm{d}z) \\ &= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\min(x,y)} q(z, \theta)^{\mathrm{T}} C^{-1}(z, \theta) P(\mathrm{d}z) \right\} q(y, \theta) P_n(\mathrm{d}y). \end{aligned} \tag{6}$$

To see that $K(x, P_n)$ is really a 'modification' of $P_{\hat{\theta}}(x)$, one can check that $K(x, P_\theta) = P_\theta(x)$.

As our first observation in this section we note that the form of $K(x, P_n)$ in the case of exponential distributions is quite straightforward:

$$\begin{aligned} K(x, P_n) &= \hat{\theta} \int_0^\infty \left(2 + \frac{\hat{\theta}}{2} \min(x, y) - \hat{\theta} y\right) \min(x, y) P_n(\mathrm{d}y) \\ &= \hat{\theta} \int_0^x \left(2 - \frac{\hat{\theta}}{2} y\right) y P_n(\mathrm{d}y) + \hat{\theta}\left(2 + \frac{\hat{\theta}}{2} x\right) x\{1 - P_n(x)\} - \hat{\theta}^2 x \int_x^\infty y P_n(\mathrm{d}y) \end{aligned}$$

or

$$K(x, P_n) = \frac{\hat{\theta}}{n} \sum_{i: X_i \leq x} \left(2X_i - \frac{\hat{\theta}}{2} X_i^2\right) + \hat{\theta}\left(2 + \frac{\hat{\theta}}{2} x\right) x\{1 - P_n(x)\} - x \frac{\hat{\theta}^2}{n} \sum_{i: X_i > x} X_i.$$

Secondly we note the rate of convergence of w_n to a standard Brownian motion is good; at least, as good as the convergence of the uniform empirical process to a Brownian bridge in the classical theory. Numerical results that support this claim are presented in Section 3.

For testing exponentiality there exists one other interesting form of empirical process, which is also asymptotically distribution free. In particular, using the equality $1 - P_\theta(bx) = \{1 - P_\theta(x)\}^b$ more fully than via a single statistic, the latter as used in Deshpande (1983) for example, one can consider the empirical process

$$\alpha_n(x) = -\sqrt{n}(1 - P_n(bx) - \{1 - P_n(x)\}^b)$$

and test if it is asymptotically a zero-mean process. It is easy to see that

$$\alpha_n(x) = \sqrt{n}\{P_n(bx) - P(bx)\} + \sqrt{n}[\{1 - P_n(x)\}^b - \{1 - P(x)\}^b]$$

and hence,

$$\alpha_n(x) = v_n(bx) - b\{1 - P_\theta(x)\}^{b-1} v_n(x) + o_p(1). \tag{7}$$

Following, e.g., Angus (1982) and Nikitin and Tchirina (1996), choose $b = 2$. After the time transformation $t = 1 - P_{\hat{\theta}}(x)$ the process α_n will converge in distribution to the process β,

$$\beta(t) = u(t^2) - 2tu(t), \tag{8}$$

where u is a standard Brownian bridge on $[0, 1]$. Hence we obtain a standard, or distribution free, process as the limit.

A more general version of this last process was studied by Koul (1977, 1978). Starting with the equality $1 - P_\theta(x + y) = \{1 - P_\theta(x)\}\{1 - P_\theta(y)\}$, Koul considered the process

$$\alpha_n(x, y) = -\sqrt{n}[1 - P_n(x + y) - \{1 - P_n(x)\}\{1 - P_n(y)\}].$$

The asymptotic form of Koul's process is

$$\alpha_n(x, y) = v_n(x + y) - \{1 - P(x, \theta)\}v_n(y) - \{1 - P(y, \theta)\}v_n(x) + o_p(1)$$

and therefore, after the time transformation $t = 1 - P_{\hat{\theta}}(x),\ s = 1 - P_{\hat{\theta}}(y)$, it converges in distribution to β^*,

$$\beta^*(t, s) = u(ts) - tu(s) - su(t), \tag{9}$$

which is again a distribution free process in t and s. We note that the process (9) is of a very appealing structure: it can be written as Πu where, for a function $f(t, s), (t, s) \in [0, 1]^2$,

$$\Pi f(t, s) = f(t, s) - tf(1, s) - sf(t, 1) + tsf(1, 1)$$

is the projection of f on the class of functions (on $[0, 1]^2$) equal to zero everywhere on the boundary of the unit square. This does not imply as yet that Πu does not define u uniquely, because in (9) Π is applied to a very narrow class of functions of t and s given by $f(t, s) = u(ts)$. However, the transformation (9) annihilates functions $ct \log t$ where c is constant and therefore, one cannot reconstruct u from β^*.

We also remark that although the process α_n looks like a non-parametric object, unconnected with and free from any estimation of the parameter θ, asymptotically it is actually a one-to-one transformation of $\hat{v}_n$. Indeed, the relationship between $\hat{v}_n$ and v_n is given by

$$\hat{v}_n(x) = v_n(x) - \frac{\mathrm{d}}{\mathrm{d}\theta} P_\theta(x)\sqrt{n}(\hat{\theta} - \theta) + o_p(1) = v_n(x) + \frac{\mathrm{d}}{\mathrm{d}\theta} P_\theta(x) \frac{1}{\theta} \int_0^\infty x v_n(\mathrm{d}x) + o_p(1)$$

if $\hat{\theta} = 1/\bar{X}$. The main term on the right-hand side above is an orthogonal projection of v_n (see Khmaladze, 1979 and Khmaladze and Koul, 2004). At the same time the functions $\lambda \mathrm{d}P_\theta(x)/\mathrm{d}\theta = \lambda x \exp(-\theta x)$ are annihilated by the right-hand side of (7), for any λ independent of x. These are the only functions which are annihilated by this transformation. Therefore in (7) we can replace v_n by $\hat{v}_n$, but this time the relationship between α_n and $\hat{v}_n$ is (asymptotically) one-to-one.

Angus (1982) and Koul (1977) calculated limiting critical values of Kolmogorov–Smirnov statistics from $\alpha_n(x)$ and $\alpha_n(x, y)$, respectively. However, the distribution of (8) and (9) is certainly more complicated than the distribution of a standard Brownian motion.

We note that certain asymptotically distribution free goodness-of-fit statistics can be produced, based on the empirical likelihood principle. This was demonstrated by Einmahl and McKeague (2003), who showed that asymptotically their statistics behave as certain weighted quadratic functionals from a standard Brownian motion or Brownian bridge. That is quite different, however, to the production of an asymptotically distribution free empirical process such as w_n, from which any chosen statistics can be constructed.

3. Assessment of convergence of the new process w_n

It is not entirely clear how to best evaluate the convergence of distribution of a sequence of processes to the limiting process. For w_n we considered several statistics with different behaviour: one-sided and two-sided Kolmogorov–Smirnov statistics,

$$d_n^+ = \sup_{0 \leqslant x < \infty} w_n(x), \quad d_n^- = -\inf_{0 \leqslant x < \infty} w_n(x), \quad d_n = \max(d_n^+, d_n^-) = \sup_{0 \leqslant x < \infty} |w_n(x)|,$$

the Cramér–von Mises statistic,

$$\omega_n^2 = \int_0^\infty w_n^2(x)\,\mathrm{d}P_\theta(x),$$

and the Anderson–Darling statistic,

$$A_n^2 = \int_0^\infty \frac{w_n^2(x)}{P_\theta(x)}\,\mathrm{d}P_\theta(x).$$

We evaluated the distribution functions of all five statistics for finite n and compared these distribution functions with their limits in a simulation study.

The limit distribution function of d_n^+ and d_n^- is $2\Phi(z)-1$, with Φ the standard normal distribution function, which follows from the reflection principle; e.g., Feller (1971, p. 175). The limit distribution of d_n is given, e.g., in Shiryaev (1999, p. 251) and Borodin and Salminen (2002), and was calculated by Shinjikashvili using the numerical method of Khmaladze and Shinjikashvili (2001). For tables of the limit distribution of ω_n^2 see Orlov (1972) or Martynov (1977), while for A_n^2 see Deheuvels and Martynov (2003). In fact, we have now made all these limiting distributions, and others, available on the web site of the School of Mathematics, Statistics and Computer Science, VUW:

http://www.mcs.vuw.ac.nz/~ray/Brownian/

Random samples of sizes 50, 100 and 200 were generated from the exponential distribution with parameter $\theta = 1$, such that $P(x) = 1 - \exp(-x)$. We considered two situations: θ assumed known and θ estimated, using the maximum likelihood estimator, $\hat{\theta} = 1/\bar{X}$. Table 1 reports the empirical sizes of statistics d_n, ω_n^2 and A_n^2 obtained from 50,000 replications, for nominal sizes of 5% and 10%, for both $\theta = 1$ and $\hat{\theta}$.

It is clear from Table 1 that the statistics have accurate empirical sizes, both when θ is assumed known and, more realistically, when it is estimated. Fig. 1 illustrates the good convergence of the entire empirical

Table 1
Convergence properties of representative test statistics from w_n

Test statistic	Size of sample	Nominal size 5%		Nominal size 10%	
		$\theta = 1$	$\theta = \hat{\theta}$	$\theta = 1$	$\theta = \hat{\theta}$
d_n	50	5.56	4.54	9.81	8.90
	100	5.50	4.79	9.89	9.09
	200	5.50	5.09	10.06	9.70
ω_n^2	50	4.95	4.60	9.77	9.62
	100	5.01	4.66	9.92	9.50
	200	5.06	4.87	9.96	9.71
A_n^2	50	4.98	4.79	9.89	9.77
	100	5.13	4.75	10.06	9.65
	200	5.15	4.85	10.05	9.78

Empirical sizes (%) for named test statistics at various sample sizes, with exponential parameter θ either known or estimated. Fifty thousand replications for each sample size.

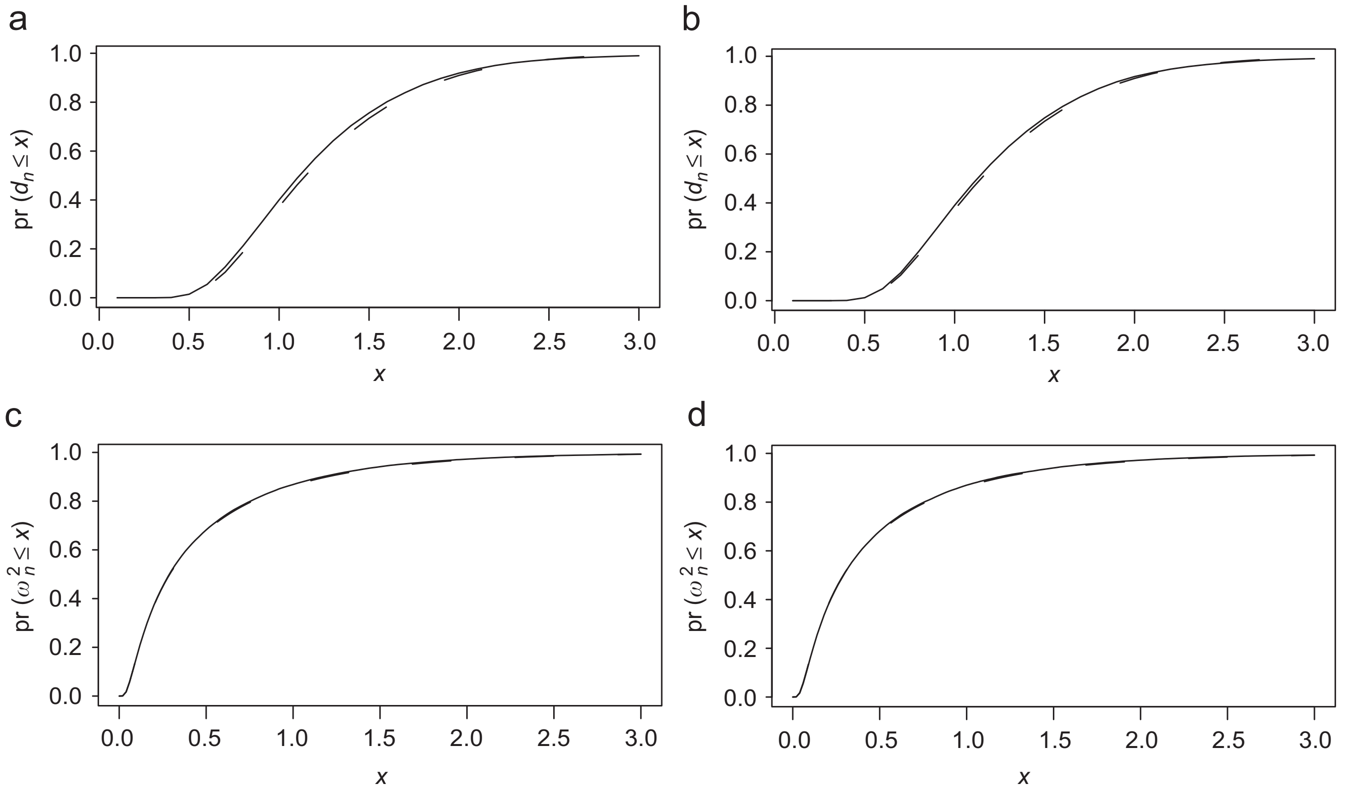

Fig. 1. Comparison of empirical distribution function (solid line) and corresponding limit distribution (dashed line) for two goodness-of-fit test statistics based on the process w_n, calculated in each case from 50,000 standard exponential samples of sizes 50 or 100 with estimated parameters: (a) size 50, d_n; (b) size 100, d_n; (c) size 50, ω_n^2; (d) size 100, ω_n^2.

distribution functions of two of the test statistics to their limits: we present graphical comparisons for each of d_n and ω_n^2, for 50,000 samples of sizes 50 and 100, with $\theta = \hat{\theta}$. Convergence was comparable for the other three test statistics considered, and slightly better with $\theta = 1$ assumed known.

Thus we conclude that the sequence of processes w_n converges in distribution quickly to the limiting Brownian motion. Given this good convergence, the asymptotic theory for any statistic from w_n can be used with confidence; we therefore leave open the choice of particular statistics. However, we note that a chosen statistic need only be calculated once, rather than the more computationally intensive requirements of permutation tests, for example.

We remark that for testing exponentiality, calculation of Kolmogorov–Smirnov statistics d_n^+, d_n^- and d_n is easy and quick, yet exact. In each interval $(X_{(j)}, X_{(j+1)})$, formed by adjacent order statistics, the compensator $K(x, P_n)$ is simply a quadratic function in x that attains its minimum at the point

$$x_j^0 = \frac{\sum_{i:X_i > X_{(j)}} X_i}{n-j} - \frac{2}{\hat{\theta}} = \frac{\sum_{i:X_i > X_{(j)}} X_i}{n-j} - 2\bar{X},$$

while $P_n(x)$ is constant. So the maximum or minimum of $P_n(x) - K(x, P_n)$ on each interval is attained at the end-points or at x_j^0, if $x_j^0 \in (X_{(j)}, X_{(j+1)})$; i.e., $d_n^+ = \max(a_n, b_n)$ with

$$a_n = \max_j \left\{ \left(\frac{j}{n}\right) - K(X_{(j)}, P_n) \right\} \quad \text{and} \quad b_n = \max_{j:x_j^0 \in (X_{(j)}, X_{(j+1)})} \left\{ \left(\frac{j}{n}\right) - K(x_j^0, P_n) \right\},$$

while

$$d_n^- = -\min_j \left\{ \left(\frac{j-1}{n}\right) - K(X_{(j)}, P_n) \right\}.$$

To close this section, we note that a simulation study of local power would require a vast number of alternatives, the choice of which is somewhat arbitrary, so we do not pursue that here. We suspect, however, that the findings from an extensive simulation study using local alternatives would be quite different to those from a study with fixed alternatives, such as Henze and Meintanis (2005), for example.

4. Comments on some well-known statistics

A widely used counterpart of goodness-of-fit statistics is formed by asymptotically linear statistics of the form

$$\int g(x,\hat{\theta})\hat{v}_n(\mathrm{d}x) + \mathrm{o_p}(1) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left\{g(X_i,\hat{\theta}) - \int g(x,\hat{\theta})P_{\hat{\theta}}(\mathrm{d}x)\right\} + \mathrm{o_p}(1) \tag{10}$$

for some deterministic function g, usually with certain square integrability requirements. These statistics have a nature very different from omnibus statistics: they lead to asymptotically optimal tests for certain deviations from $\mathscr{P}$, but have very low or no power against the huge majority of other deviations from the hypothetical family.

Indeed, under the hypothesis, i.e., under any $P_\theta \in \mathscr{P}$, asymptotically linear statistics (10) converge to a Gaussian random variable with zero-mean and variance $\int g^2(x,\theta)P_{\hat{\theta}}(\mathrm{d}x)$. However, under any sequence of contiguous alternatives A_n satisfying (1), these statistics converge to the same Gaussian random variable plus the shift

$$\int h(x,\theta)g(x,\theta)\,\mathrm{d}P_\theta(x), \tag{11}$$

e.g., see Khmaladze (1979), Janssen (2000). Clearly, (11) is the inner product between functions $h(\cdot,\theta)$ and $g(\cdot,\theta)$ in $L_2(P_\theta)$. Although for $h(x,\theta) = cg(x,\theta)$, with c constant, the shift (11) can be quite large, for $h(\cdot,\theta) \perp g(\cdot,\theta)$ this shift is simply zero. Hence any asymptotically linear statistic under all alternatives orthogonal to $g(\cdot,\theta)$ will have the same limiting distribution as under the hypothesis.

However, tests based on several asymptotically linear statistics, such as $\int g_j(x)\hat{v}_n(\mathrm{d}x), j = 1,\ldots,k$, can be considered and perceived as goodness-of-fit tests. The chi-square test is probably the most prominent such example. If each linear statistic, with particular kernel g_j, is sensitive to particular alternatives, their combination may be sensitive to 'sufficiently many' different alternatives, and therefore will be of 'sufficiently' omnibus nature.

It is remarkable though that many papers explicitly or implicitly devoted to goodness-of-fit testing of exponentiality are actually based on a single asymptotically linear statistic. Quite unexpectedly, we also realised that some well-known statistics lead asymptotically to basically the same test and no combination of these statistics can be used for goodness-of-fit testing; at least, not for local alternatives.

Before we consider examples of such usage below, we need to make one observation. Suppose $g_\perp(\cdot,\theta)$ is the part of $g(\cdot,\theta)$ orthogonal to the hypothetical score function $\partial \log p(x,\theta)/\partial\theta$. Since all functions $h(\cdot,\theta)$ are orthogonal to it we have

$$\int h(x,\theta)g(x,\theta)\,\mathrm{d}P_\theta(x) = \int h(x,\theta)g_\perp(x,\theta)\,\mathrm{d}P_\theta(x).$$

Therefore it will be inefficient to use the asymptotically linear statistics with $g(\cdot,\theta) \neq g_\perp(\cdot,\theta)$: the part of it equal to $\int\{g(x,\hat{\theta}) - g_\perp(x,\hat{\theta})\}\hat{v}_n(\mathrm{d}x)$ will contribute to its asymptotic variance but not to its asymptotic shift.

Asymptotically linear statistics, centred as in (10), are asymptotically equivalent to those with the kernel $g_\perp$. Indeed, using the asymptotic representation for $\hat{\theta} - \theta$ (see, e.g., (19) in Section 5), it is not difficult to obtain

$$\int g(x,\hat{\theta})\hat{v}_n(\mathrm{d}x) = \int g_\perp(x,\theta)v_n(\mathrm{d}x) + \mathrm{o_p}(1) = \int g_\perp(x,\theta)\hat{v}_n(\mathrm{d}x) + \mathrm{o_p}(1). \tag{12}$$

Now let us consider several well-known statistics, which are all scale invariant. Hence in what follows we set $\theta = 1$ and we also denote by P the exponential distribution function P_θ, $P_\theta(x) = 1 - \exp(-\theta x)$, with $\theta = 1$.

The papers of Deshpande (1983) and Bandyopadhyay and Basu (1989) are based on testing whether $1 - P(bx) = \{1 - P(x)\}^b$, and the test statistic can be written as the integral

$$\frac{1}{n(n-1)} \sum_{i \neq j} \mathbb{I}\{X_j > bX_i\} = \int_0^\infty \{1 - P_n(bx+)\} P_n(\mathrm{d}x), \tag{13}$$

where $\mathbb{I}\{\cdot\}$ denotes the indicator function. For any fixed b, (13) is asymptotically linear:

$$\sqrt{n}\left\{\int_0^\infty \{1 - P_n(bx+)\} P_n(\mathrm{d}x) - \int_0^\infty \{1 - P(bx)\}\,\mathrm{d}P(x)\right\}$$
$$= \int_0^\infty \{\exp(-bx) - \exp(-x/b)\} v_n(\mathrm{d}x) + \mathrm{o_p}(1) = \int_0^\infty g_{\perp,D}(x) v_n(\mathrm{d}x) + \mathrm{o_p}(1), \tag{14}$$

where

$$g_{\perp,D}(x) = \exp(-bx) - \exp(-x/b) - (1-b)/(1+b).$$

Deshpande (1983) and Bandyopadhyay and Basu (1989) each suggest different choices of b, but it is one fixed value of b in both cases. The right-hand side of (14) may be studied as a process in b. Based on the combination of the Laplace transform of the empirical process $\hat{v}_n$, this process in b will converge to a zero-mean Gaussian process under the hypothesis, and will have non-zero shift under all contiguous alternatives A_n. Basically such an approach was taken and extended in Baringhaus and Henze (1991, 2000) and Henze (1993), where quadratic functionals from not only the Laplace transform but also from the Fourier transform of $\hat{v}_n$ were studied. Epps and Pulley (1986) previously investigated 'smooth' functionals from the empirical characteristic function.

A statistic known as the Gini index,

$$G_n = \frac{\sum_{i \neq j} |X_i - X_j|}{2n(n-1)\bar{X}},$$

received a thorough treatment in Gail and Gastwirth (1978). G_n can be rewritten as

$$G_n = \frac{n}{(n-1)} \frac{1}{2\bar{X}} \iint |x-y|\,\mathrm{d}P_n(x)\,\mathrm{d}P_n(y).$$

After normalisation, it becomes evident that G_n is also an asymptotically linear statistic:

$$\frac{n}{(n-1)\bar{X}} \frac{\sqrt{n}}{2} \iint |x-y| \{\mathrm{d}P_n(x)\mathrm{d}P_n(y) - \mathrm{d}P(x)\mathrm{d}P(y)\} + \frac{\sqrt{n}}{2}\left(\frac{1}{\bar{X}} - 1\right) \iint |x-y|\,\mathrm{d}P(x)\,\mathrm{d}P(y)$$
$$= \int \left\{\int |x-y| P(\mathrm{d}y)\right\} v_n(\mathrm{d}x) - \sqrt{n}\left(\frac{\bar{X}-1}{2}\right) + \mathrm{o_p}(1)$$
$$= \int \{x - 1 + 2\exp(-x)\} v_n(\mathrm{d}x) - \int \frac{x-1}{2} v_n(\mathrm{d}x) + \mathrm{o_p}(1)$$
$$= \int g_{\perp,G}(x) v_n(\mathrm{d}x) + \mathrm{o_p}(1), \tag{15}$$

where

$$g_{\perp,G}(x) = \frac{x-1}{2} - 1 + 2\exp(-x).$$

Here we made use of $\bar{X} \to 1$, $\int |x-y| \exp(-y)\,\mathrm{d}y = x - 1 + 2\exp(-x)$ and $\iint |x-y| \exp(-x)\exp(-y)\,\mathrm{d}x\,\mathrm{d}y = 1$. It is obvious that G_n is asymptotically normal, but Gail and Gastwirth (1978) also show the exact distribution of G_n and demonstrate that convergence to the normal distribution is quick.

The so-called Moran statistic, introduced in Moran (1951), has the form

$$M_n = \frac{1}{n} \sum_{i=1}^n \log \frac{X_i}{\bar{X}}.$$

This is the score test statistic for testing $\alpha = 1$ assuming an underlying Gamma density $a(x) = \theta^{-\alpha}\Gamma^{-1}(\alpha)x^{\alpha-1}\exp(-x/\theta)$. Indeed, the score function in this case is $\partial \log a(x)/\partial\alpha = \log(x/\theta) - \Gamma'(1)/\Gamma(1)$. Recently Tchirina (2003) studied large deviations of a test based on the Moran statistic and some of its local properties. For the normalised form of this statistic with $\theta = 1$, we obtain

$$\sqrt{n}\left\{\int \log\frac{x}{\bar{X}}\,\mathrm{d}P_n(x) - \int \log x\,\mathrm{d}P(x)\right\} = \int g_{\perp,M}(x)v_n(\mathrm{d}x) + \mathrm{o_p}(1), \tag{16}$$

where

$$g_{\perp,M}(x) = \log x - \gamma - (x-1)$$

and $\gamma = \Gamma'(1)/\Gamma(1) = \int \log x \exp(-x)\mathrm{d}x$, so that $-\gamma$ is Euler's constant. This statistic is suggested as an "omnibus" goodness-of-fit test by Stephens (1986), although he adds the caveat that its use can be "risky", due to problems if the data include ties or zeros. For an omnibus test, Stephens in fact restricts attention to alternatives with monotone failure rate.

One more statistic for testing exponentiality was suggested in Cox and Oakes (1984),

$$T_n = n^{-1}\sum_{i=1}^{n}\left(1 - \frac{X_i}{\bar{X}}\right)\log\frac{X_i}{\bar{X}}. \tag{17}$$

Its representation as an asymptotically linear statistic when $\theta = 1$ is

$$\int g_{\perp,\mathrm{CO}}(x)v_n(\mathrm{d}x) + \mathrm{o_p}(1) \quad \text{where } g_{\perp,\mathrm{CO}}(x) = (1-x)(\log x + 1 - \gamma). \tag{18}$$

Statistic (17) was introduced explicitly as the score test statistic for testing $\alpha = 1$ assuming a Weibull distribution $A(x) = 1 - \exp\{(-x/\theta)^{\alpha}\}$. Indeed, the function $(1 - x/\theta)\log(x/\theta) + 1$ is the score function of the Weibull density $a(x) = (\alpha x^{\alpha-1}/\theta^{\alpha})\exp\{(-x/\theta)^{\alpha}\}$ with respect to parameter α at $\alpha = 1$. Later, however, the Cox–Oakes statistic has been studied and recommended for goodness-of-fit testing (e.g., Ascher, 1990; Henze and Meintanis, 2005).

The four asymptotically linear statistics above appear different, each capable of detecting deviations from exponentiality in different directions. Thus it might seem important to compare their asymptotic behaviour against various alternatives. Yet for local alternatives all four statistics are extremely similar. Though formally the corresponding kernel functions $g_{\perp}$ are different they all have similar graphs, as shown in Fig. 2, and all four statistics are very highly correlated. In particular, the correlations between statistic (18) and (14), (15), (16) are $0.909, -0.936, 0.971$; between statistic (16) and (14), (15) are $0.909, -0.833$; and between statistic (15) and (14) -0.865. So while these statistics are optimal for different alternatives, asymptotically they are very similar and it is not important to study their relative power, at least for local alternatives. Note, following Bandyopadhyay and Basu (1989), we use a value of $b = 0.44$ in Deshpande's (1983) statistic (13).

Furthermore, as noted above, no single asymptotically linear statistic can provide a goodness-of-fit test with asymptotic power against every sequence of contiguous alternatives (1). To illustrate this we give an explicit example of a local alternative for which these four statistics have very low power, while the power of an omnibus test, calculated on the transformed empirical process w_n, is essentially higher. Perhaps unexpectedly, for such an example we make use of the uniform distribution on $[0, 1]$, against which the Cox–Oakes statistic, for example, was reported to have high power (Ascher, 1990; Henze and Meintanis, 2005). We convert this uniform distribution into a local alternative by using the mixture

$$a_n(x) = \left(1 - \frac{c}{\sqrt{n}}\right)p(x) + \frac{c}{\sqrt{n}}q(x) = p(x) + \frac{c}{\sqrt{n}}\{q(x) - p(x)\},$$

where c is a scalar, $p(x)$ is the exponential density with parameter $\theta = 1$ and $q(x)$ is the uniform density on $[0, 1]$. Denote $\|g_{\perp,\mathrm{CO}}\| = (\int g^2_{\perp,\mathrm{CO}}(x)\mathrm{e}^{-x}\,\mathrm{d}x)^{1/2}$. The value of the integral

$$\frac{1}{\|g_{\perp,\mathrm{CO}}\|}\int g_{\perp,\mathrm{CO}}(x)\{q(x) - p(x)\}\,\mathrm{d}x = \frac{1}{\|g_{\perp,\mathrm{CO}}\|}\int_0^1 g_{\perp,\mathrm{CO}}(x)\,\mathrm{d}x = 0.0138$$

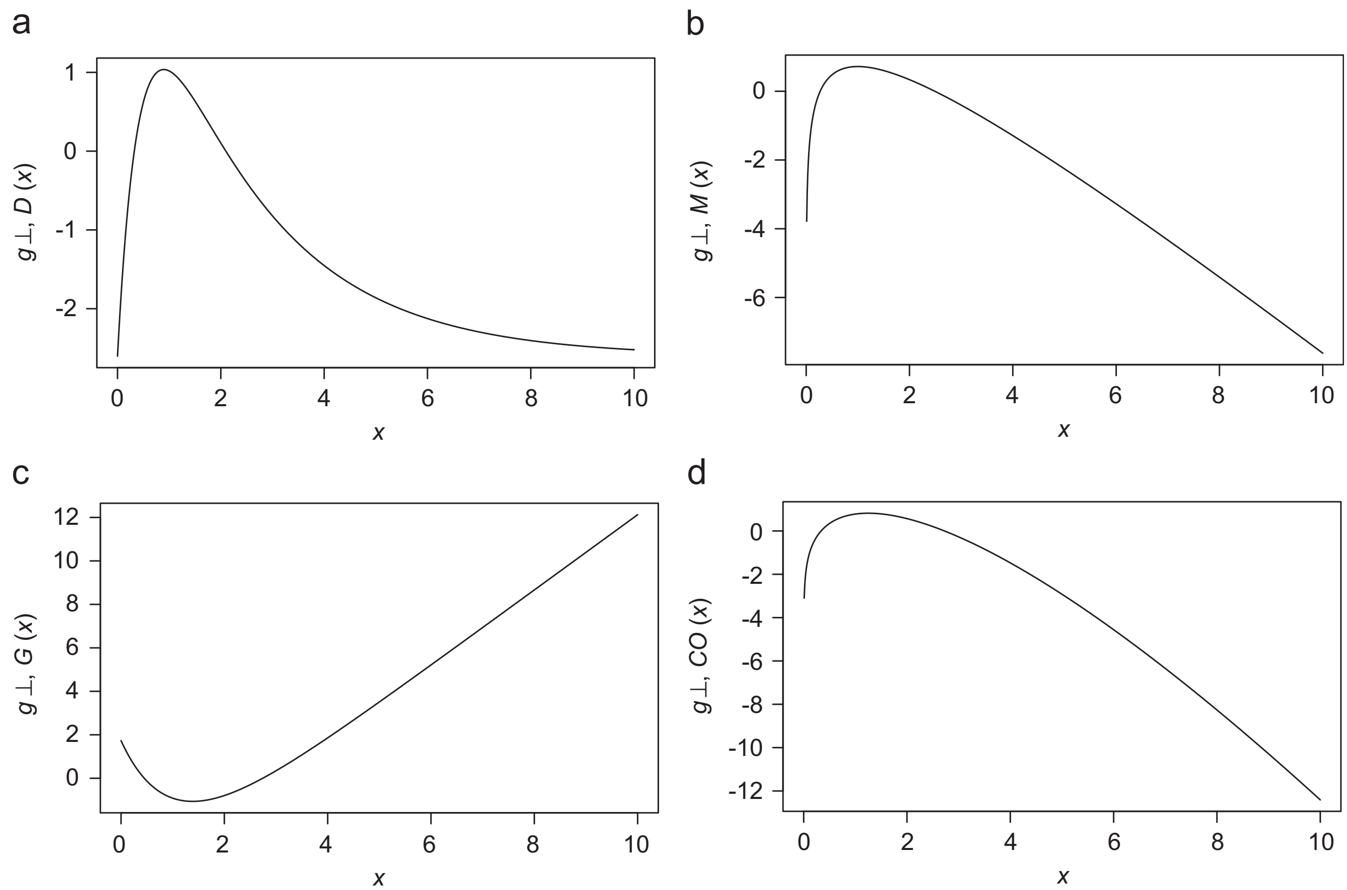

Fig. 2. Kernel functions of four asymptotically linear statistics: (a) Deshpande; (b) Moran; (c) Gini; (d) Cox–Oakes.

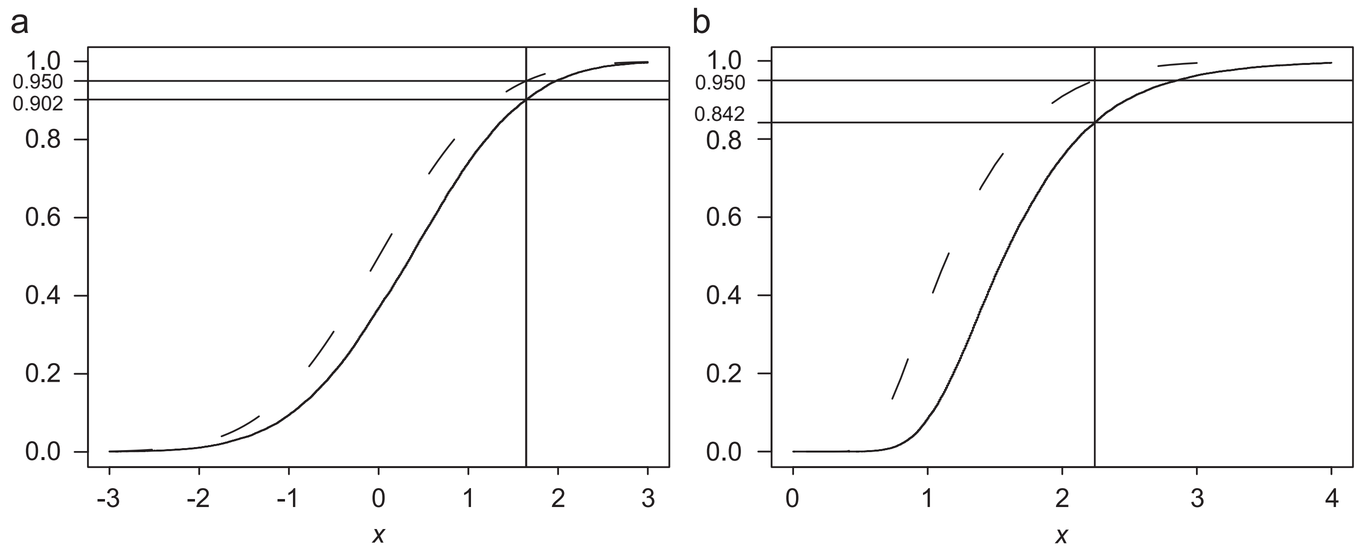

Fig. 3. The change in distribution of (a) the Cox–Oakes statistic and (b) the Kolmogorov–Smirnov statistic d_n from the transformed empirical process w_n, calculated from 20,000 samples of size 900. Distributions are shown under the hypothesis of exponentiality (dashed line) and under the local alternative mixture discussed in the text (solid line).

is quite small. Hence the asymptotic shift of the standardised Cox–Oakes statistic under alternatives a_n, being $0.0138c$, is also small for moderate c, which means that its power against a_n will be very low.

Fig. 3 compares the distributions of the Cox–Oakes statistic and the Kolmogorov–Smirnov statistic d_n calculated from the transformed empirical process w_n given in (3). The statistics are compared under the

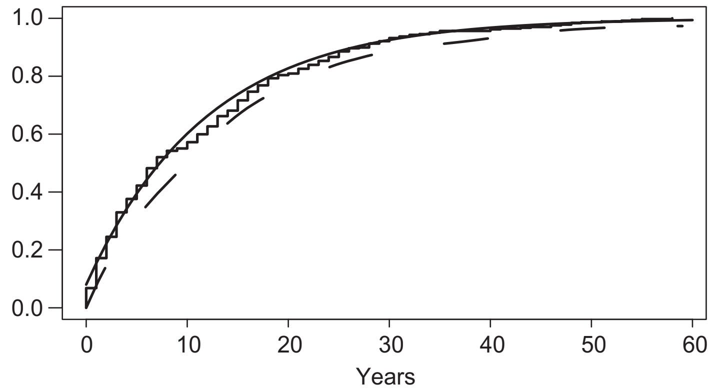

Fig. 4. Empirical distribution function of durations of rule of Chinese Emperors (steps) with an exponential approximation (solid) and the compensator $K(x, P_n)$ (dashed).

hypothesis and under the alternative a_n above, using 20,000 replications of size $n = 900$ with $c = 5$. That is, under the alternative, 150 of the 900 observations in each replicate sample come from the uniform distribution. It is clear that the change in distribution of the Kolmogorov–Smirnov statistic is essentially sharper, and the power of the Kolmogorov–Smirnov test based on the transformed version of the empirical process is larger.

Another example was found in our research on certain historical data (Khmaladze et al., 2007). Fig. 4 shows the empirical distribution function of the durations of rule for Chinese Emperors, as given in Encyclopaedia Britannica (2002), along with an exponential approximation. The approximation looks very good and the value of the standardised Cox–Oakes statistic is 0.08, which is obviously close to the median of the standard normal distribution. However, the value of the Kolmogorov–Smirnov statistic from the transformed empirical process is equal to $d_n = 2.46$, which corresponds to a p-value of 0.028. Hence the hypothesis of exponentiality should be rejected, which agrees with the more detailed analysis reported in Khmaladze et al. (2007).

5. The process w_n and the problem of multiple testing

Goodness-of-fit test statistics are often used along with asymptotically linear, or directed statistics. It is quite common to apply a goodness-of-fit test to the sample first, to see if there is any disagreement with the hypothesis. If there is, one or two asymptotically linear statistics may then be used, to see if the disagreement can be attributed to deviations in particular directions. Alternatively, asymptotically linear statistics may be used to test whether there are deviations in one or more of a few particular directions initially, then a goodness-of-fit test is used to see if there are any other sorts of deviations from the hypothesis. In both these situations, but especially the latter, we should be interested in the conditional asymptotic distribution of the goodness-of-fit statistic, given the values of the asymptotically linear statistics already used, rather than the marginal asymptotic distribution.

While in principle the conditional asymptotic distribution of a statistic is the correct thing to study, in practice marginal asymptotic distributions are commonly used because the conditional distribution seems too difficult to use, or too computationally intensive. However, the approach based on the transformation (3), using (6), allows the resolution of this multiple testing problem in a simple and natural way. That is, it allows the derivation of a version of the transformed process w_n which, under the hypothesis is not only a standard Brownian motion, but is also asymptotically independent from any linear statistics already used. In this way there is no need for any conservative adjustments to test sizes, e.g., as used in Bonferroni corrections.

Indeed, suppose the maximum likelihood estimator $\hat{\theta}$ of the parameter θ has the asymptotic representation

$$\sqrt{n}(\hat{\theta} - \theta) = B^{-1}(\theta) \int \frac{\partial \log p(x, \theta)}{\partial \theta} \, \mathrm{d}v_n(x) + \mathrm{o_p}(1), \tag{19}$$

where $B(\theta)$ is the Fisher information, or Fisher information matrix if θ is a vector. Further, suppose that tests based on asymptotically linear statistics (12) were used. Then the corresponding kernels are added as coordinates to the vector function (4) and the compensator is still defined as (6). If θ is one-dimensional and if a single statistic of the form (12) was used, then we put

$$q(x,\theta)^{\mathrm{T}} = \left[1, \frac{\partial \log p(x,\theta)}{\partial\theta}, g(x,\theta)\right]$$

and construct $C(z,\theta)$ of (5), which now becomes a 3×3 matrix. Then following, e.g., Khmaladze (1993) or Khmaladze and Koul (2004, Sections 2 and 3), the obtained process w_n and the statistic $\int g\,\mathrm{d}v_n$ will be asymptotically independent, while w_n will still converge to Brownian motion.

Acknowledgements

We thank Ray Brownrigg for some assistance with statistical computing and Boris Andreyanov for useful discussions. The editor and two anonymous referees provided feedback on earlier versions which helped shape the current presentation. JH was supported by a Victoria University of Wellington Science Faculty grant during the completion of this work.

References

Angus, J.E., 1982. Goodness-of-fit tests for exponentiality based on a loss-of-memory type functional equation. Journal of Statistical Planning and Inference 6 (3), 241–251.
Ascher, S., 1990. A survey of tests for exponentiality. Communications in Statistics—Theory and Methods 19 (5), 1811–1825.
Bandyopadhyay, D., Basu, A.P., 1989. A note on tests for exponentiality by Deshpande. Biometrika 76, 403–405.
Baringhaus, L., Henze, N., 1991. A class of consistent tests for exponentiality based on the empirical Laplace transform. Annals of the Institute of Statistical Mathematics 43, 179–192.
Baringhaus, L., Henze, N., 2000. Tests of fit for exponentiality based on a characterization via the mean residual life function. Statistical Papers 41, 225–236.
Borodin, A., Salminen, P., 2002. Handbook of Brownian Motion: Facts and Formulae (Probability and its Applications), second ed. Birkhauser Verlag AG, Basel.
Cox, D.R., Oakes, D., 1984. Analysis of Survival Data. Chapman & Hall, London.
Deheuvels, P., Martynov, G., 2003. Karhunen–Loève expansions for weighted Wiener processes and Brownian bridges via Bessel functions. Progress in Probability 55, 57–93.
Deshpande, J.V., 1983. A class of tests for exponentiality against increasing failure rate average alternatives. Biometrika 70, 514–518.
Dragulescu, A., Yakovenko, V.M., 2001a. Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States. Physica A 299, 213–221.
Dragulescu, A., Yakovenko, V.M., 2001b. Evidence for the exponential distribution of income in the USA. The European Physical Journal B 20, 585–589.
Durbin, J., 1973. Weak convergence of the sample distribution function when parameters are estimated. The Annals of Statistics 1, 279–290.
Einmahl, J.H.J., McKeague, I.W., 2003. Empirical likelihood based hypothesis testing. Bernoulli 9 (2), 267–290.
Encyclopaedia Britannica, 2002. Encyclopaedia Britannica, Chicago.
Epps, T.W., Pulley, L.B., 1986. A test for exponentiality vs. monotone hazard alternatives derived from the empirical characteristic function. Journal of the Royal Statistical Society B 48, 206–213.
Feller, W., 1971. An Introduction to Probability Theory and its Applications, vol. II, second ed. Wiley, New York.
Gail, M.H., Gastwirth, J.L., 1978. A scale-free goodness-of-fit test for exponentiality based on the Gini statistic. Journal of the Royal Statistical Society B 40, 350–357.
Granger, C.W.J., Ding, Z., 1995. Some properties of absolute return. An alternative measure of risk. Les Annales d'Economie et de Statistique 40, 67–91.
Henze, N., 1993. A new flexible class of omnibus tests for exponentiality. Communications in Statistics—Theory and Methods 22, 115–133.
Henze, N., Meintanis, S.G., 2005. Recent and classical tests for exponentiality: a partial review with comparisons. Metrika 61 (1), 29–45.
Janssen, A., 2000. Global power functions of goodness of fit tests. The Annals of Statistics 28 (1), 239–253.
Khmaladze, E.V., 1979. The use of Omega-square tests for testing parametric hypotheses. Theory of Probability and its Applications 24 (2), 283–302.
Khmaladze, E.V., 1981. Martingale approach in the theory of goodness of fit tests. Theory of Probability and its Applications 26 (1), 240–257.
Khmaladze, E.V., 1993. Goodness of fit problem and scanning innovation martingales. The Annals of Statistics 21 (2), 798–829.
Khmaladze, E.V., Koul, H.L., 2004. Martingale transforms goodness-of-fit tests in regression models. The Annals of Statistics 32 (3), 995–1034.

Khmaladze, E.V., Shinjikashvili, E., 2001. Calculation of noncrossing probabilities for Poisson processes and its corollaries. Advances in Applied Probability 33, 702–716.

Khmaladze, E.V., Brownrigg, R.D., Haywood, J., 2007. Brittle power: on Roman Emperors and exponential lengths of rule. Statistics and Probability Letters 77 (12), 1248–1257.

Koul, H.L., 1977. A test for new better than used. Communications in Statistics: Theory and Methods A6, 563–574.

Koul, H.L., 1978. A class of tests for testing new is better than used. The Canadian Journal of Statistics 6 (2), 249–271.

Martynov, G., 1977. Omega Square (Cramér–von Mises) Test (in Russian). Nauka, Moscow (in Russian).

Moran, P.A.P., 1951. The random division of an interval—part II. Journal of the Royal Statistical Society B 13, 147–150.

Nikitin, Ya.Yu., Tchirina, A.V., 1996. Bahadur efficiency and local optimality of a test for the exponential distribution based on the Gini statistic. Journal of the Italian Statistical Society 5 (1), 163–175.

Orlov, A.I., 1972. On testing the symmetry of distributions. Theory of Probability and its Applications 17, 357–361.

Pham, H. (Ed.), 2003. Handbook of Reliability Engineering. Springer, London.

Shiryaev, A.N., 1999. Essentials of Stochastic Finance. Facts, Models, Theory. World Scientific, Singapore.

Sornette, D., 2003. Why Stock Markets Crash: Critical Events in Complex Financial Systems. Princeton University Press, Princeton, NJ.

Stanley, M.H.R., Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M.A., Stanley, H.E., 1996. Scaling behaviour in the growth of companies. Nature 379, 804–806.

Stephens, M.A., 1986. Tests for the exponential distribution. In: D'Agostino, R.B., Stephens, M.A. (Eds.), Goodness-of-Fit Techniques. Marcel Dekker, New York, pp. 421–459.

Tchirina, A.V., 2003. Large deviations for a class of scale-free statistics under the gamma distribution. Journal of Mathematical Sciences 128 (1), 2640–2655, (2005) (in English); Translation from Zapiski Nauchnykh Seminarov POMI 298, 252–279 (in Russian).

Van den Berg, G.J., Ridder, G., 1998. An empirical equilibrium search model of the labor market. Econometrica 66, 1183–1221.