



HAL
open science

A new weighted integral goodness of fit statistic for exponentiality

L. Baringhaus, N. Henze

► **To cite this version:**

L. Baringhaus, N. Henze. A new weighted integral goodness of fit statistic for exponentiality. *Statistics and Probability Letters*, 2010, 78 (8), pp.1006. 10.1016/j.spl.2007.09.060 . hal-00616538

HAL Id: hal-00616538

<https://hal.science/hal-00616538>

Submitted on 23 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

A new weighted integral goodness of fit statistic for exponentiality

L. Baringhaus, N. Henze

PII: S0167-7152(07)00360-4
DOI: 10.1016/j.spl.2007.09.060
Reference: STAPRO 4800

To appear in: *Statistics and Probability Letters*

Received date: 8 May 2007
Revised date: 27 July 2007
Accepted date: 25 September 2007

Please cite this article as: Baringhaus, L., Henze, N., A new weighted integral goodness of fit statistic for exponentiality. *Statistics and Probability Letters* (2007), doi:10.1016/j.spl.2007.09.060

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



A NEW WEIGHTED INTEGRAL GOODNESS OF FIT STATISTIC FOR EXPONENTIALITY

L. Baringhaus

N. Henze

Leibniz Universität Hannover
and
Universität Karlsruhe (TH)

Abstract

We propose a new weighted integral goodness of fit statistic for exponentiality. The statistic is motivated by a characterization of the exponential distribution via the mean residual life function. Its limit null distribution is seen to be the same as that of a certain weighted integral of the squared Brownian bridge. The Laplace transform and cumulants of the latter are expressible in terms of Bessel functions.

Key words and phrases: Exponential distribution; mean residual life function; Brownian bridge; Bessel functions

AMS 1991 subject classifications: Primary 62G10; Secondary 62E20

1 Introduction

Let X_1, \dots, X_n be independent and identically distributed non negative random variables with some unknown distribution function $F(x) = P(X_1 \leq x)$, $x \geq 0$. It is well-known that, under the assumption $0 < E(X_1) < \infty$, the distribution of X_1 is exponential, i.e.

$$F(x) = 1 - \exp(-\lambda x), \quad x \geq 0,$$

for some $\lambda > 0$, if and only if the mean residual life function is constant, i.e. we have

$$E(X_1 - z | X_1 > z) = E(X_1) \quad \text{for each } z > 0. \quad (1)$$

Baringhaus and Henze (2000) noted that (1) is equivalent to

$$E(\min(X_1, z)) = E(X_1)F(z) \quad \text{for each } z > 0. \quad (2)$$

Arguing that, under the assumptions $X_1 \geq 0$ and $0 < E(X_1) < \infty$, (2) is a characteristic property of the class $\{Exp(\lambda) : \lambda > 0\}$ of exponential distributions, they suggested a new approach to assess exponentiality. In particular, they proposed the Cramér-von Mises type statistic

$$G_n = n \int_0^\infty \left[\frac{1}{n} \sum_{k=1}^n \min(U_k, z) - \frac{1}{n} \sum_{k=1}^n I(U_k \leq z) \right]^2 e^{-z} dz, \quad (3)$$

where $\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k$ and $U_k = X_k/\bar{X}$, $k = 1, \dots, n$. Interestingly, the limit null distribution of G_n is the same as the limit null distribution of the classical Cramér-von Mises statistic when testing the simple hypothesis of uniformity on the interval $[0, 1]$. The goodness of fit test based on G_n rejecting the hypothesis of exponentiality if $G_n > g_n$ where g_n is the $(1 - \alpha)$ -quantile of the null distribution of G_n for some given $\alpha \in (0, 1)$, is consistent against any fixed alternative distribution. The test is discussed in the recent review paper on goodness of fit tests for exponentiality by Henze and Meintanis (2005). From the comparative simulation study given there one may conclude that it is a serious competitor, although there are other procedures showing a better power performance for special alternatives. However, the favorable behavior of these procedures is recognized by varying some weight parameter built in the corresponding test statistic. So, to obtain a possible gain in power performance we suggest to generalize (3) in a natural way by using the more general weight function e^{-az} , $z \geq 0$, where $a > -1$ is some real parameter. Then we have the test statistic

$$G_{n,a} = n \int_0^\infty \left[\frac{1}{n} \sum_{k=1}^n \min(U_k, z) - \frac{1}{n} \sum_{k=1}^n I(U_k \leq z) \right]^2 e^{-az} dz. \quad (4)$$

By some elementary calculations we see that in the case $a \neq 0$ the statistic can be written in the form

$$G_{n,a} = \frac{1}{n} \sum_{k,\ell=1}^n \left[\frac{2}{a^3} - \left(\frac{1}{a^2} + \frac{1}{a} \right) \min(U_k, U_\ell) (\exp(-aU_k) + \exp(-aU_\ell)) \right. \\ \left. - \left(\frac{2}{a^3} + \frac{1}{a^2} \right) \exp(-a \min(U_k, U_\ell)) + \left(\frac{1}{a^2} + \frac{1}{a} \right) \exp(-a \max(U_k, U_\ell)) \right].$$

In the special case $a = 0$ the statistic has the alternative representation

$$G_{n,0} = \frac{1}{n} \sum_{k,\ell=1}^n \left[\frac{1}{3} \min(U_k, U_\ell)^3 + (\min(U_k, U_\ell) - 1)(\max(U_k, U_\ell) - \frac{1}{2}U_k^2 - \frac{1}{2}U_\ell^2) \right].$$

Using the central limit theorem in Hilbert spaces as a method of proof different to that given in Baringhaus and Henze (2000) we show in the next section, that the limit null distribution of $G_{n,a}$ is the same as that of $\int_0^1 B(t)^2(1-t)^{a-1} dt$, where $(B(t), 0 \leq t \leq 1)$ is a Brownian bridge. Using this representation we shall be able to express the Laplace transform of the limit null distribution in terms of the Bessel function $I_{\frac{1}{a+1}}$, and, additionally, its cumulants via sums of the even powers of the reciprocal zeros of the Bessel function $J_{\frac{1}{a+1}}$. Critical values and empirical power values of the test obtained by simulation are shown in the last section.

2 The limit null distribution of $G_{n,a}$

The null distribution of $G_{n,a}$ does not depend on the parameter λ of the underlying exponential distribution. We therefore assume that $\lambda = 1$, i.e. that X_k has the distribution function $1 - \exp(-x)$, $x \geq 0$. Putting

$$R_n(z) = \sqrt{n} \left(\frac{1}{n} \sum_{k=1}^n [\min(X_k, z) - \bar{X}I(X_k \leq z)] \right), \quad z \geq 0,$$

it follows that

$$G_{n,a} = \frac{1}{\bar{X}^3} \int_0^\infty R_n^2(z) \left[\exp\left(-\left(\frac{1}{\bar{X}} - 1\right)az\right) - 1 \right] e^{-az} dz + \frac{1}{\bar{X}^3} \int_0^\infty R_n^2(z) e^{-az} dz. \quad (5)$$

We shall show below that for each $a > -1$ the statistic $\int_0^\infty R_n^2(z) \exp(-az) dz$ has a limit distribution. Using this in advance we first prove that

$$\int_0^\infty R_n^2(z) \left[\exp\left(-\left(\frac{1}{\bar{X}} - 1\right)az\right) - 1 \right] e^{-az} dz = o_P(1). \quad (6)$$

For, let us treat the cases $a \geq 0$ and $-1 < a < 0$ separately. If $a \geq 0$ we put $\eta = \frac{1}{2}$, $a^* = -\eta$ and $\epsilon_0 = 1$. For given $0 < \epsilon \leq \epsilon_0$ we consider the case where $|\frac{1}{\bar{X}} - 1| < \epsilon$. From

$$\left| \exp\left(-\left(\frac{1}{\bar{X}} - 1\right)az\right) - 1 \right| \leq za\epsilon \exp(\epsilon az), \quad z \geq 0,$$

we get using $z \leq e^{\eta z}/\eta$ for $z \geq 0$ that

$$\begin{aligned} \left| \int_0^\infty R_n^2(z) \left[\exp\left(-\left(\frac{1}{\bar{X}} - 1\right)az\right) - 1 \right] e^{-az} dz \right| &\leq a\epsilon \int_0^\infty R_n^2(z) z \exp(-a(1-\epsilon)z) dz \\ &\leq \frac{a\epsilon}{\eta} \int_0^\infty R_n^2(z) \exp(-[a(1-\epsilon) - \eta]z) dz. \\ &\leq \frac{a\epsilon}{\eta} \int_0^\infty R_n^2(z) \exp(-a^*z) dz. \end{aligned}$$

If $-1 < a < 0$ we put $\eta = \frac{a+1}{4}$, $a^* = \frac{a-3}{4}$ and $\epsilon_0 = \frac{1}{2}\left(-\frac{1}{a} - 1\right)$. For given $0 < \epsilon \leq \epsilon_0$ we consider the case where $|\frac{1}{\bar{X}} - 1| < \epsilon$. From

$$\left| \exp\left(-\left(\frac{1}{\bar{X}} - 1\right)az\right) - 1 \right| \leq z|a|\epsilon \exp(-\epsilon az), \quad z \geq 0,$$

we obtain

$$\begin{aligned} \left| \int_0^\infty R_n^2(z) \left[\exp\left(-\left(\frac{1}{\bar{X}} - 1\right)az\right) - 1 \right] e^{-az} dz \right| &\leq |a|\epsilon \int_0^\infty R_n^2(z) z \exp(-a(1+\epsilon)z) dz \\ &\leq \frac{|a|\epsilon}{\eta} \int_0^\infty R_n^2(z) \exp(-[a(1+\epsilon) - \eta]z) dz. \\ &\leq \frac{|a|\epsilon}{\eta} \int_0^\infty R_n^2(z) \exp(-a^*z) dz. \end{aligned}$$

Thus in any case there is some $a^* > -1$ such for each $0 < \epsilon \leq \epsilon_0$

$$\left| \int_0^\infty R_n^2(z) \left[\exp\left(-\left(\frac{1}{\bar{X}} - 1\right)az\right) - 1 \right] \exp(-az) dz \right| \leq \frac{|a|\epsilon}{\eta} \int_0^\infty R_n^2(z) \exp(-a^*z) dz$$

if $|\frac{1}{\bar{X}} - 1| < \epsilon$. Using that $\int_0^\infty R_n^2(z) \exp(-a^*z) dz$ has a limit distribution, $\frac{1}{\bar{X}} = 1 + o_P(1)$ and $\frac{1}{\bar{X}^3} = 1 + o_P(1)$ we obtain (6) and, additionally,

$$G_{n,a} = \int_0^\infty R_n^2(z) \exp(-az) dz + o_P(1).$$

It remains to derive the limit distribution of $\int_0^\infty R_n^2(z) e^{-az} dz$. To this end note that

$$R_n(z) = H_n(z) + (\bar{X} - 1)L_n(z), \quad z \geq 0,$$

where

$$H_n(z) = \sqrt{n} \left(\frac{1}{n} \sum_{k=1}^n [\min(X_k, z) - I(X_k \leq z) - (X_k - 1)(1 - e^{-z})] \right), \quad z \geq 0,$$

and

$$L_n(z) = \sqrt{n} \left(\frac{1}{n} \sum_{k=1}^n [1 - e^{-z} - I(X_k \leq z)] \right), \quad z \geq 0.$$

H_n and L_n are random elements in the Hilbert space $L_2(\mathbb{R}_+, \mathfrak{B}_+, \mu_a)$, where $\mathbb{R}_+ = [0, \infty)$, \mathfrak{B}_+ is the Borel σ -field on \mathbb{R}_+ , and μ_a is the σ -finite measure having density $\exp(-ax)$, $x \geq 0$, with respect to Lebesgue measure on $(\mathbb{R}_+, \mathfrak{B}_+)$. Since

$$\int_0^\infty \text{Var} \left(\min(X_1, z) - I(X_1 \leq z) - (X_1 - 1)(1 - e^{-z}) \right) d\mu_a(z) = \int_0^\infty [1 - e^{-z}] e^{-(a+1)z} dz < \infty$$

and

$$\int_0^\infty \text{Var}\left(1 - e^{-z} - I(X_1 \leq z)\right) d\mu_a(z) = \int_0^\infty [1 - e^{-z}]e^{-(a+1)z} dz < \infty$$

the central limit theorem for Hilbert space-valued random variables applies (see Ledoux and Talagrand (1991), Corollary 10.9). Recognizing that

$$\rho(w, z) = \min(1 - e^{-w}, 1 - e^{-z}) - (1 - e^{-w})(1 - e^{-z}), \quad w, z \geq 0, \quad (7)$$

is the covariance function of the process $(H_n(z), z \geq 0)$ and also that of the process $(L_n(z), z \geq 0)$, there is a zero mean Gaussian process $(H(z), z \geq 0)$ with sample paths in $L_2(\mathbb{R}_+, \mathfrak{B}_+, \mu_a)$ and covariance kernel (7) such that

$$\int H_n^2(z) d\mu_a(z) \xrightarrow{\mathcal{D}} \int H^2(z) d\mu_a(z)$$

and

$$\int L_n^2(z) d\mu_a(z) \xrightarrow{\mathcal{D}} \int H^2(z) d\mu_a(z),$$

where “ $\xrightarrow{\mathcal{D}}$ ” means convergence in distribution. Due to $\bar{X} - 1 = o_P(1)$ this implies that

$$\int_0^\infty R_n^2(z) \exp(-az) dz \xrightarrow{\mathcal{D}} \int_0^\infty H^2(z) \exp(-az) dz.$$

The result just proved is summarized as follows.

Theorem 1. *The limit null distribution of the test statistic $G_{n,a}$ is that of $\int H^2(z) d\mu_a(z)$ where $(H(z), z \geq 0)$ is a zero mean Gaussian process with sample paths in $L_2(\mathbb{R}_+, \mathfrak{B}_+, \mu_a)$ and covariance function (7).*

The Gaussian process $(H(z), z \geq 0)$ has the same covariance function as the process $(B(1 - e^{-z}), z \geq 0)$ where $(B(t), 0 \leq t \leq 1)$ is the Brownian bridge. Consequently,

$$\int H^2(z) d\mu_a(z) \stackrel{\mathcal{D}}{=} \int_0^1 B^2(1-t)t^{a-1} dt \stackrel{\mathcal{D}}{=} \int_0^1 B^2(t)(1-t)^{a-1} dt$$

where “ $\stackrel{\mathcal{D}}{=}$ ” denotes equality in distribution. Since the limit distribution may be of independent interest, it is studied further. Let

$$G_a = \int_0^1 B(t)^2(1-t)^{a-1} dt.$$

Theorem 2. *a) The Laplace transform of G_a is*

$$\begin{aligned} \psi_a(z) &= \prod_{j=1}^{\infty} \left(1 + \left(\frac{2}{a+1}\right)^2 \frac{2z}{\gamma_j^2}\right)^{-1/2} \\ &= \left\{ \Gamma\left(1 + \frac{1}{a+1}\right) \frac{1}{\left(\frac{1}{a+1}\sqrt{2z}\right)^{\frac{1}{a+1}}} I_{\frac{1}{a+1}}\left(\frac{2}{a+1}\sqrt{2z}\right) \right\}^{-1/2}, \quad z \geq 0, \end{aligned}$$

where I_ν is the modified Bessel function of the first kind of order ν .

b) The first four cumulants of G_a are

$$\begin{aligned}\kappa_1 &= E(G_a) = \frac{1}{(a+1)(a+2)}, \\ \kappa_2 &= \text{Var}(G_a) = \frac{2}{(a+2)^2(a+1)(2a+3)}, \\ \kappa_3 &= \frac{16}{(a+2)^3(a+1)(2a+3)(3a+4)}, \\ \kappa_4 &= \frac{48(11a+16)}{(a+2)^4(2a+3)^2(a+1)(3a+4)(4a+5)}.\end{aligned}$$

Proof. Although part a) of the theorem can be obtained from the material presented in Deheuvels and Martynov (2003), we give a derivation for the special case considered here. Note that

$$G_a = \int B^2(t) d\nu_a(t),$$

where ν_a is the σ -finite measure with density $(1-t)^{a-1}$ with respect to Lebesgue measure on the Borel sets $\mathfrak{B}_{[0,1]}$ of the interval $[0,1]$. Let A be the integral operator on $L_2([0,1], \mathfrak{B}_{[0,1]}, \nu_a)$ associated with the covariance kernel $k(s,t) = \min(s,t) - st$, $0 \leq s, t \leq 1$, of the Brownian bridge. Thus

$$Af(t) = \int_0^1 k(t,s)f(s) d\nu_a(s), \quad 0 \leq t \leq 1,$$

for $f \in L_2([0,1], \mathfrak{B}_{[0,1]}, \nu_a)$. The integral operator A is positive definite. It is well known (see, e.g. Vakhania (1981), p. 58) that

$$G_a \stackrel{\mathcal{D}}{=} \sum_{k \geq 1} \lambda_k Z_k^2, \quad (8)$$

where the Z_k are independent unit normal random variables, and the λ_k are the eigenvalues of A . To obtain these eigenvalues, assume that an eigenfunction f of A with associated positive eigenvalue λ is smooth enough so that, starting with the equation

$$\int k(t,s)f(s) d\nu_a(s) = \lambda f(t), \quad 0 \leq t \leq 1, \quad (9)$$

we may differentiate twice on both sides of (9). This leads to the differential equation

$$\lambda f''(t) = -(1-t)^{a-1} f(t), \quad 0 < t < 1.$$

From (9) we also infer the boundary conditions $f(0) = 0$ and $f(1) = 0$. Putting $\varphi(t) = f(1-t)$, φ satisfies the differential equation

$$\lambda \varphi''(t) = -t^{a-1} \varphi(t), \quad 0 < t < 1, \quad (10)$$

subject to the boundary conditions

$$\varphi(0) = \varphi(1) = 0. \quad (11)$$

The general solution of (10) was found by Lommel (1868); for another reference see Nielsen (1904), p. 130. In view of (11), the general solution is

$$ct^{1/2} J_{\frac{1}{a+1}} \left(\frac{2}{a+1} \lambda^{-1/2} t^{(a+1)/2} \right), \quad 0 \leq t \leq 1,$$

where J_ν is the Bessel function of the first kind of order ν , and c is some constant. In what follows, $0 < \gamma_1 < \gamma_2 < \dots$ are the positive zeros of $J_{\frac{1}{a+1}}$. Putting

$$\lambda_j = \left(\frac{2}{a+1} \right)^2 \frac{1}{\gamma_j^2}, \quad j = 1, 2, \dots,$$

and

$$f_j(t) = \frac{(a+1)^{1/2}}{J_{\frac{a+2}{a+1}}(\gamma_j)} J_{\frac{1}{a+1}}(\gamma_j(1-t)^{(a+1)/2})(1-t)^{1/2},$$

($0 \leq t \leq 1$, $j = 1, 2, \dots$), it follows from

$$\begin{aligned} & \int f_j(t) f_k(t) d\nu_a(t) \\ &= \frac{a+1}{J_{\frac{a+2}{a+1}}(\gamma_j) J_{\frac{a+2}{a+1}}(\gamma_k)} \int_0^1 (1-t)^a J_{\frac{1}{a+1}}(\gamma_j(1-t)^{(a+1)/2}) J_{\frac{1}{a+1}}(\gamma_k(1-t)^{(a+1)/2}) dt \\ &= \begin{cases} 0, & j \neq k, \\ 1, & j = k, \end{cases} \end{aligned}$$

(see Erdélyi et al. (1953), p. 70) that $\{f_j, j \geq 1\}$ is an orthonormal set in $L_2([0, 1], \mathfrak{B}_{[0,1]}, \nu_a)$. For $f \in L_2([0, 1], \mathfrak{B}_{[0,1]}, \nu_a)$ the Fourier-Bessel series $\sum_{j \geq 1} \alpha_j f_j$ with

$$\alpha_j = \int f(t) f_j(t) d\nu_a(t), \quad j \geq 1,$$

converges in $L_2([0, 1], \mathfrak{B}_{[0,1]}, \nu_a)$, i.e.

$$\lim_{n \rightarrow \infty} \int \left| \sum_{j=1}^n \alpha_j f_j(t) - f(t) \right|^2 d\nu_a(t) = 0.$$

This result can be proved in a rather elementary way by using the work of Hochstadt (1967), for example. Thus $\{f_j, j \geq 1\}$ is a complete orthonormal set in $L_2([0, 1], \mathfrak{B}_{[0,1]}, \nu_a)$. Using

$$-f_j(t)(1-t)^{a-1} = \lambda_j f_j''(t), \quad 0 < t < 1,$$

integration by parts gives

$$(1-t) \int_0^t s f_j(s) (1-s)^{a-1} ds + t \int_t^1 f(s) (1-s)^a ds = \lambda_j f_j(t), \quad 0 \leq t \leq 1,$$

or, equivalently,

$$\int k(t,s) f_j(s) d\nu_a(s) = \lambda_j f_j(t), \quad 0 \leq t \leq 1.$$

Thus the $f_j, j \geq 1$, form a complete orthonormal system of eigenfunctions of A with associated eigenvalues $\lambda_j, j \geq 1$. In view of (8) the Laplace transform of G_a is

$$\begin{aligned} \psi_a(z) &= \prod_{j=1}^{\infty} \left(1 + \left(\frac{2}{a+1} \right)^2 \frac{2z}{\gamma_j^2} \right)^{-1/2}, \quad z \geq 0, \\ &= \left\{ \Gamma \left(1 + \frac{1}{a+1} \right) \frac{1}{\left(\frac{1}{a+1} \sqrt{2z} \right)^{\frac{1}{a+1}}} I_{\frac{1}{a+1}} \left(\frac{2}{a+1} \sqrt{2z} \right) \right\}^{-1/2}, \quad z \geq 0, \end{aligned}$$

where I_ν is the modified Bessel function of the first kind of order ν . For the second equality see Nielsen (1904), p.358. This proves part a) of the theorem. To prove part b) which asserts an interesting connection between the cumulants κ_ν of G_a , the latter being an integral involving the Brownian bridge, and the sums of the even powers of the reciprocal zeros γ_j of the Bessel functions $J_{\frac{1}{a+1}}$ we note that

$$\kappa_\nu = 2^{\nu-1} (\nu-1)! \sum_{j=1}^{\infty} \lambda_j^\nu = 2^{\nu-1} (\nu-1)! \left(\frac{2}{a+1} \right)^{2\nu} \sum_{j=1}^{\infty} \left(\frac{1}{\gamma_j} \right)^{2\nu}, \quad \nu \geq 1.$$

We can derive the first four cumulants of G_a directly by using Fubini's theorem and calculating mixed moments of a zero mean multivariate normal vector with covariance structure specified by the Brownian bridge. Alternatively, we can use that

$$\begin{aligned} \sum_{j=1}^{\infty} \frac{1}{\gamma_j^2} &= \frac{2^{-2}}{1 + \frac{1}{a+1}} \\ \sum_{j=1}^{\infty} \frac{1}{\gamma_j^4} &= \frac{2^{-4}}{\left(1 + \frac{1}{a+1} \right)^2 \left(\frac{1}{a+1} + 2 \right)} \\ \sum_{j=1}^{\infty} \frac{1}{\gamma_j^6} &= \frac{2^{-6} \cdot 2}{\left(1 + \frac{1}{a+1} \right)^3 \left(\frac{1}{a+1} + 2 \right) \left(\frac{1}{a+1} + 3 \right)} \\ \sum_{j=1}^{\infty} \frac{1}{\gamma_j^8} &= \frac{2^{-8} \left(5 \frac{1}{a+1} + 11 \right)}{\left(1 + \frac{1}{a+1} \right)^4 \left(\frac{1}{a+1} + 2 \right)^2 \left(\frac{1}{a+1} + 3 \right) \left(\frac{1}{a+1} + 4 \right)} \end{aligned}$$

(see Nielsen (1904), p. 360) to obtain the first four cumulants stated in part b) of the theorem.

It is also possible to derive ‘limit test statistics’ in the cases $a \rightarrow -1$ and $a \rightarrow \infty$. For, it is immediately seen that

$$G_{n,-1} = \lim_{a \downarrow -1} G_{n,a} = -2n + \frac{1}{n} \sum_{k,\ell}^n \exp(\min(U_k, U_\ell)).$$

Of course, one can use $G_{n,-1}$ as test statistic for testing the hypothesis of exponentiality. However, presently we are unable to present any asymptotic theory for $G_{n,-1}$ as $n \rightarrow \infty$. In the second case, $a \rightarrow \infty$, putting $U_{1:n} = \min(U_1, \dots, U_n)$ we have

$$\lim_{a \rightarrow \infty} a^{-2} \exp(aU_{1:n}) \left(a^3 \frac{G_{n,a}}{n} - 2 \right) = \frac{1}{n} - 2U_{1:n}.$$

Thus one may suggest $U_{1:n}$ as test statistic rejecting the hypothesis of exponentiality for small values of $U_{1:n}$. We remark that a test based on this statistic is easily done because $U_{1:n}$ has the beta distribution $B(1, n-1)$ with the density $(n-1)(1-t)^{n-2}$, $0 \leq t \leq 1$. Finally, we remark that for each $-1 < a < \infty$ and given level α the test obtained by rejecting the hypothesis of exponentiality if $G_{n,a} > g_{n,a}(\alpha)$, where $g_{n,a}(\alpha)$ is the $(1-\alpha)$ -quantile of $G_{n,a}$ in the case where the hypothesis of exponentiality is true, is consistent against any fixed alternative distribution. A proof of this assertion is easily done by adapting the arguments used by Baringhaus and Henze (2000).

3 Empirical Results

For sample sizes $n = 10, 20, 30, 40, 50, 100, 200$, parameter values $a = -0.99, -0.9, -0.5, 0, 0.5, 1, 1.5, 2, 5, 10$ and levels $\alpha = 0.05, 0.1$ we got approximations of the critical values $g_{n,a}(\alpha)$ by simulation with 100000 replications. The results are shown in TABLE 1 and TABLE 2 of the appendix. An empirical power study based on simulations with 10000 replications was done for the sample sizes $n = 20, 50, 100$ and level $\alpha = 0.05$. The alternative distributions were chosen from distribution families considered also by Baringhaus and Henze (1991), (2000). The distributions included are the Gamma distributions (\mathcal{G}), Weibull distributions (\mathcal{W}), Lognormal distributions (\mathcal{LN}) with scale parameter 1 and shape parameter θ , the uniform distribution $\mathcal{U}[0, 1]$, the Half-Normal distribution (\mathcal{HN}), the Half-Cauchy distribution (\mathcal{HC}), the χ_1^2 -distribution, the Power distributions (\mathcal{PW}) with density $\theta^{-1}x^{1/\theta-1}$, $0 < x < 1$, the linear increasing failure rate distributions (LIFR) with density $(1 + \theta x) \exp(-(x + \theta x^2/2))$, $x > 0$, and the JSHAPE distributions (JS) with density $(1 + \theta x)^{-1/\theta-1}$, $x > 0$. The empirical power values (rounded to the nearest integer) are shown in TABLE 3, TABLE 4 and TABLE 5 of the appendix. For estimated power values of various other competitive procedures we refer to Baringhaus and Henze (1991) and also to Henze and Meintanis (2005), although there is merely a partial overlap with the alternative distributions considered in the latter paper. We conclude, that taking a weight parameter $1 \leq a \leq 2$ is a rather good choice. For some alternative distributions, the Gamma distributions or lognormal distributions with shape parameter $\theta < 1$ for example, a choice of $a > 2$ can be recommended. In these cases the power performance is also better than that of the top ranked test in the comparative study of Henze and Meintanis (2005).

Acknowledgement: The authors are indebted to Nora Gürtler for assistance in computation.

REFERENCES

- Baringhaus, L. and Henze, N. (1991). A class of consistent tests for exponentiality based on the empirical Laplace transform. *Ann. Inst. Statist. Math.* **43** 551-564.
- Baringhaus, L. and Henze, N. (2000). Tests of fit for exponentiality based on a characterization via the mean residual life function. *Statist. Papers.* **41** 225–236.
- Deheuvels, P. and Martynov, G. (2003). Karhunen-Loève expansions for weighted Wiener processes and Brownian bridges via Bessel functions. In: J. Hoffmann-Jørgensen, M. Marcus and J. Wellner, eds. *High Dimensional Probability III* (Birkhäuser, Basel) pp 57–93.
- Erdélyi, A., Magnus, W., Oberhettinger, F., Tricomi, F. (1953). *Higher transcendental functions, Vol. II*. McGraw-Hill, New York.
- Henze, N. and Meintanis, S. (2005). Recent and classical tests for exponentiality: a partial review with comparisons. *Metrika* **61** 29–45.
- Hochstadt, H. (1967). The mean convergence of Fourier-Bessel series. *SIAM Rev.* **9** 211-218.
- Kamke, E. (1977). *Differentialgleichungen. Lösungen und Lösungsverfahren, Band I*. 9. Auflage, Teubner, Stuttgart.
- Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*. Springer, New York.
- Nielsen, N. (1904). *Handbuch der Theorie der Zylinderfunktionen*. Teubner, Leipzig.
- Vakhania, N.N. (1981). *Probability Distributions on Linear Spaces*. North Holland, New York.

A Appendix: Tables of critical values and power values

TABLE 1
Critical values of $G_{n,a}$ for $\alpha = 0.05$

Sample size n	Parameter a									
	-0.99	-0.9	-0.5	0	0.5	1	1.5	2	5	10
10	5.056	3.944	1.846	1.017	0.634	0.427	0.305	0.226	0.065	0.018
20	8.069	5.983	2.336	1.147	0.678	0.449	0.313	0.230	0.067	0.021
30	9.886	7.282	2.597	1.193	0.691	0.450	0.316	0.231	0.068	0.021
40	11.299	8.107	2.723	1.224	0.702	0.446	0.316	0.234	0.069	0.022
50	12.433	8.793	2.846	1.249	0.699	0.453	0.320	0.232	0.069	0.022
100	15.848	10.654	3.072	1.279	0.713	0.457	0.316	0.233	0.069	0.022
200	18.338	11.900	3.233	1.296	0.711	0.464	0.320	0.235	0.069	0.022

TABLE 2
Critical values of $G_{n,a}$ for $\alpha = 0.1$

Sample size n	Parameter a									
	-0.99	-0.9	-0.5	0	0.5	1	1.5	2	5	10
10	3.063	2.580	1.406	0.793	0.490	0.332	0.237	0.174	0.049	0.014
20	4.395	3.526	1.705	0.875	0.515	0.340	0.238	0.176	0.050	0.016
30	5.341	4.180	1.847	0.908	0.528	0.341	0.240	0.176	0.050	0.016
40	6.033	4.581	1.945	0.927	0.531	0.343	0.240	0.176	0.051	0.016
50	6.569	4.946	2.042	0.948	0.536	0.346	0.238	0.175	0.051	0.016
100	8.200	5.929	2.226	0.977	0.539	0.346	0.240	0.177	0.051	0.016
200	9.917	6.961	2.359	0.991	0.544	0.345	0.239	0.176	0.051	0.017

TABLE 3
 Percentage of 10000 Monte Carlo samples declared significant;
 level $\alpha = 0.05$; sample size $n = 20$

Distribution	Parameters a									
	-0.99	-0.9	-0.5	0	0.5	1	1.5	2	5	10
$\mathcal{W}(0.6)$	59	61	65	69	69	70	72	73	78	80
$\mathcal{W}(1.2)$	1	1	4	9	12	13	14	14	12	8
$\mathcal{W}(1.4)$	0	0	11	26	33	35	37	37	32	20
$\mathcal{W}(1.6)$	0	0	26	50	60	63	64	65	57	39
χ_1^2	38	39	44	47	49	52	55	57	66	72
$\mathcal{PW}(0.8)$	1	13	87	94	94	93	92	89	69	40
$\mathcal{PW}(1.2)$	0	1	33	48	48	44	40	35	17	7
$\mathcal{PW}(1.4)$	0	0	16	28	27	24	21	18	9	7
$\mathcal{PW}(2.0)$	1	1	4	8	10	11	12	14	25	37
$\mathcal{PW}(3.0)$	13	17	28	38	46	53	59	64	79	87
LIFR(1)	0	0	6	14	18	19	19	19	13	7
LIFR(2)	0	0	10	24	29	30	30	30	20	11
LIFR(4)	0	0	17	36	42	43	43	42	30	16
LIFR(6)	0	0	23	43	50	51	52	50	37	20
LIFR(10)	0	0	30	53	60	61	61	60	47	26
$\mathcal{HN}(0,1)$	0	0	7	17	21	21	22	21	14	8
$\mathcal{LN}(0.7)$	2	3	16	34	46	53	59	64	73	70
$\mathcal{LN}(0.8)$	6	6	10	19	23	27	31	34	43	39
$\mathcal{LN}(1.0)$	21	21	20	18	16	15	14	13	11	8
$\mathcal{LN}(1.5)$	66	67	68	67	67	66	65	64	59	53
\mathcal{HC}	72	72	73	71	69	67	66	65	59	52
JS(0.5)	50	50	51	49	47	45	44	43	39	35
JS(1.0)	81	82	84	84	83	83	83	82	80	77
$\mathcal{U}[0,1]$	0	3	58	74	74	71	67	62	38	18
$\mathcal{G}(0.4)$	56	58	65	68	72	75	77	80	86	90
$\mathcal{G}(0.6)$	26	28	30	30	32	33	35	37	44	50
$\mathcal{G}(0.8)$	11	11	12	11	10	10	11	11	13	16
$\mathcal{G}(1.4)$	1	1	4	9	13	14	15	16	14	9
$\mathcal{G}(1.6)$	0	0	6	16	21	24	26	27	25	17
$\mathcal{G}(1.8)$	0	0	10	24	31	34	38	39	38	26
$\mathcal{G}(2.0)$	0	0	15	33	42	46	50	52	49	36
$\mathcal{G}(2.4)$	0	0	28	53	64	69	72	74	72	57
$\mathcal{G}(3.0)$	0	1	49	77	85	89	91	92	91	80

TABLE 4
 Percentage of 10000 Monte Carlo samples declared significant;
 level $\alpha = 0.05$; sample size $n = 50$

Distribution	Parameters a									
	-0.99	-0.9	-0.5	0	0.5	1	1.5	2	5	10
$\mathcal{W}(0.6)$	86	88	94	96	97	98	98	98	99	99
$\mathcal{W}(1.2)$	0	0	8	21	28	30	31	32	29	22
$\mathcal{W}(1.4)$	0	1	41	66	75	77	78	79	73	61
$\mathcal{W}(1.6)$	0	5	80	95	97	98	98	98	96	90
χ_1^2	58	62	75	82	87	89	90	92	95	97
$\mathcal{PW}(0.8)$	50	94	100	100	100	100	100	100	99	89
$\mathcal{PW}(1.2)$	1	20	92	96	94	90	84	78	40	16
$\mathcal{PW}(1.4)$	0	5	72	80	75	65	54	45	15	8
$\mathcal{PW}(2.0)$	0	0	19	30	30	28	29	32	50	66
$\mathcal{PW}(3.0)$	12	19	59	78	87	91	94	95	99	100
LIFR(1)	0	0	20	39	45	45	43	42	29	17
LIFR(2)	0	1	41	64	69	69	67	65	49	31
LIFR(4)	0	2	64	84	87	87	86	85	71	49
LIFR(6)	0	4	75	90	93	93	92	91	81	60
LIFR(10)	0	8	84	95	97	97	96	96	89	72
\mathcal{HN}	0	0	28	49	54	53	51	50	34	21
$\mathcal{LN}(0.7)$	4	5	43	75	88	94	96	98	100	100
$\mathcal{LN}(0.8)$	10	10	20	37	51	60	67	73	89	95
$\mathcal{LN}(1.0)$	36	37	36	34	29	25	22	21	22	29
$\mathcal{LN}(1.5)$	91	92	94	95	95	95	94	94	91	84
\mathcal{HC}	95	95	96	96	95	94	94	93	88	80
JS(0.5)	78	79	82	82	81	80	79	78	70	61
JS(1.0)	98	98	99	99	99	99	99	99	99	98
$\mathcal{U}[0, 1]$	9	56	99	100	100	99	98	97	79	49
$\mathcal{G}(0.4)$	80	83	93	97	98	99	99	99	100	100
$\mathcal{G}(0.6)$	38	40	51	58	63	67	69	72	78	81
$\mathcal{G}(0.8)$	13	13	16	17	17	18	18	19	23	27
$\mathcal{G}(1.4)$	1	1	7	21	28	32	33	35	35	30
$\mathcal{G}(1.6)$	0	0	18	40	51	56	59	62	62	56
$\mathcal{G}(1.8)$	0	0	33	61	72	77	79	81	82	77
$\mathcal{G}(2.0)$	0	1	49	78	86	90	91	93	93	90
$\mathcal{G}(2.4)$	0	5	79	95	98	99	99	99	99	99
$\mathcal{G}(3.0)$	1	20	97	100	100	100	100	100	100	100

TABLE 5
 Percentage of 10000 Monte Carlo samples declared significant;
 level $\alpha = 0.05$; sample size $n = 100$

Distribution	Parameters a									
	-0.99	-0.9	-0.5	0.1	0.5	1	1.5	2	5	10
$\mathcal{W}(0.6)$	98	99	100	100	100	100	100	100	100	100
$\mathcal{W}(1.2)$	0	0	23	44	51	54	56	56	52	44
$\mathcal{W}(1.4)$	0	8	84	96	98	98	98	98	97	93
$\mathcal{W}(1.6)$	8	53	100	100	100	100	100	100	100	100
χ_1^2	78	83	95	98	99	99	100	100	100	100
$\mathcal{PW}(0.8)$	100	100	100	100	100	100	100	100	100	100
$\mathcal{PW}(1.2)$	51	94	100	100	100	100	100	98	72	30
$\mathcal{PW}(1.4)$	14	66	100	100	99	96	90	82	28	10
$\mathcal{PW}(2.0)$	0	6	76	81	75	69	66	65	77	89
$\mathcal{PW}(3.0)$	23	47	95	99	100	100	100	100	100	100
LIFR(1)	0	2	57	76	78	77	76	73	56	37
LIFR(2)	0	10	85	95	95	95	95	93	82	61
LIFR(4)	2	31	98	100	100	100	100	99	96	84
LIFR(6)	6	47	99	100	100	100	100	100	99	91
LIFR(10)	14	66	100	100	100	100	100	100	100	97
$\mathcal{HN}(0, 1)$	0	4	70	85	87	86	84	82	64	43
$\mathcal{LN}(0.7)$	6	12	83	99	100	100	100	100	100	100
$\mathcal{LN}(0.8)$	15	16	39	71	85	92	96	97	100	100
$\mathcal{LN}(1.0)$	53	54	56	53	47	44	42	41	45	61
$\mathcal{LN}(1.5)$	99	99	100	100	100	100	100	100	99	98
\mathcal{HC}	100	100	100	100	100	100	100	100	99	96
JS(0.5)	94	95	97	97	97	97	96	96	92	85
JS(1.0)	100	100	100	100	100	100	100	100	100	100
$\mathcal{U}[0, 1]$	94	100	100	100	100	100	100	100	98	84
$\mathcal{G}(0.4)$	96	98	100	100	100	100	100	100	100	100
$\mathcal{G}(0.6)$	52	58	77	86	90	92	94	95	97	97
$\mathcal{G}(0.8)$	16	18	24	27	29	31	34	35	40	43
$\mathcal{G}(1.4)$	1	1	22	43	54	59	62	64	66	61
$\mathcal{G}(1.6)$	0	1	49	76	84	88	90	91	93	91
$\mathcal{G}(1.8)$	0	5	75	93	97	98	98	99	99	99
$\mathcal{G}(2.0)$	1	13	91	99	99	100	100	100	100	100
$\mathcal{G}(2.4)$	6	45	99	100	100	100	100	100	100	100
$\mathcal{G}(3.0)$	39	90	100	100	100	100	100	100	100	100