
THE TEST OF EXPONENTIALITY BASED ON THE MEAN RESIDUAL LIFE FUNCTION REVISITED

A PREPRINT

Bruno Ebner
 Institute of Stochastics,
 Karlsruhe Institute of Technology (KIT),
 Englerstr. 2, 76133 Karlsruhe,
 Germany
 Bruno.Ebner@kit.edu

February 7, 2023

ABSTRACT

We revisit the family of goodness-of-fit tests for exponentiality based on the mean residual life time proposed by Baringhaus & Henze (2008). We motivate the test statistic by a characterisation of Shanbhag (1970) and provide an alternative representation, which leads to simple and short proofs for the known theory and an easy to access covariance structure of the limiting Gaussian process under the null hypothesis. Explicit formulas for the eigenvalues and eigenfunctions of the operator associated with the limit covariance are given using results on weighted Brownian bridges. In addition we derive further asymptotic theory under fixed alternatives as well as approximate Bahadur efficiencies, which provide an insight into the choice of the tuning parameter with regard to the power performance of the tests.

1 Introduction

We revisit the family of tests for exponentiality based on the residual life time as proposed by Baringhaus & Henze (2008), and we provide further theoretical insight into the asymptotic behaviour of the test under alternatives. The problem of interest is testing the assumption that data are distributed according to the exponential distribution with unknown scale parameter. To be precise, let X be a positive random variable and we write shorthand $X \sim \text{Exp}(\lambda)$, $\lambda > 0$, if X follows an exponential distribution with scale parameter λ , hence the density is given by

$$f(x, \lambda) = \lambda \exp(-\lambda x), \quad x > 0. \quad (1)$$

Note that $X \sim \text{Exp}(\lambda)$ if, and only if, $X/\lambda \sim \text{Exp}(1)$, which shows that the exponential distribution belongs to the scale family of distributions, for a detailed discussion see Johnson et al. (1994), Chapter 19. In the following we denote the family of exponential distributions by $\mathcal{E} := \{\text{Exp}(\lambda) : \lambda > 0\}$. Let X, X_1, X_2, \dots be positive independent and identically distributed (iid.) random variables with distribution \mathbb{P}^X defined on an underlying probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We test the composite hypothesis

$$H_0 : \mathbb{P}^X \in \mathcal{E} \quad (2)$$

against general alternatives, based on the sample X_1, \dots, X_n .

This testing problem has been studied extensively in the literature, and it is of ongoing interest, see Cuparić et al. (2021); Cuparić et al. (2020); Jiménez-Gamero et al. (2020); Milošević & Obradović (2016b); Xiong et al. (2022); Villaseñor & González-Estrada (2020) for some recent related publications, Ascher (1990); Henze & Meintanis (2005); Spurrier (1984) for surveys on the topic, and Allison et al. (2017); Ossai et al. (2022) for reviews of testing procedures as well as of extensive competitive Monte Carlo power studies. The focus of this article is on the procedure proposed

in Baringhaus & Henze (2008). That paper studies a family of tests of exponentiality based on the mean residual life function with test statistic

$$G_{n,a} = n \int_0^\infty \left[\frac{1}{n} \sum_{j=1}^n \min(Y_j, z) - \frac{1}{n} \sum_{j=1}^n \mathbf{1}\{Y_j \leq z\} \right]^2 \exp(-az) \, dz, \quad a > 0.$$

Here, $Y_j = X_j/\bar{X}_n$, $j = 1, \dots, n$, is the scaled data, where $\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j$. This family of tests is an extension of the test in Baringhaus & Henze (2000) insofar as the exponential weight function in the original test statistic is replaced with a more flexible weight function $w(z) = e^{-az}$, $z > 0$, which depends on some so-called tuning parameter $a > -1$. In Baringhaus & Henze (2008) the authors provide the limiting null distribution of the test statistics $G_{n,a}$ and a simulation study including simulated critical values, as well as a power study for different values of the tuning parameter a . Further insight on the behaviour of the tests under alternatives is hitherto missing. We fill this gap in the literature.

We start our investigation by providing an alternative representation of $G_{n,a}$ motivated by a characterisation of the exponential law due to Shanbhag (1970), which is obviously related to the mean residual life function. Shanbhag (1970) states that for a random variable X with $\mathbb{P}(X > y) > 0$ for $y > 0$, we have $X \sim \text{Exp}(\lambda)$ if and only if

$$\mathbb{E}[X|X > y] = y + \frac{1}{\lambda}, \quad \text{for all } y > 0, \lambda > 0.$$

Direct calculations show that under the same condition this characterisation can be restated in the following way: $X \sim \text{Exp}(\lambda)$ if and only if

$$\mathbb{E} \left[\left(X - y - \frac{1}{\lambda} \right) \mathbf{1}\{X > y\} \right] = 0, \quad \text{for all } y > 0, \lambda > 0. \quad (3)$$

In view of the scale invariance of the family \mathcal{E} , we consider the scaled data $Y_j = X_j/\bar{X}_n$, $j = 1, \dots, n$, and propose the weighted L^2 -type test statistic (fixing w.l.o.g. $\lambda = 1$ in the characterisation)

$$T_{n,a} = T_{n,a}(Y_1, \dots, Y_n) = n \int_0^\infty \left| \frac{1}{n} \sum_{j=1}^n (Y_j - y - 1) \mathbf{1}\{Y_j > y\} \right|^2 \exp(-ay) \, dy, \quad (4)$$

which depends on the tuning parameter $a \geq 0$. Note that $T_{n,a}$ is scale invariant, i.e. invariant w.r.t. transformations of the form $x \mapsto bx$, $b > 0$, since it only depends on the scaled data Y_1, \dots, Y_n . Scale invariance is indeed a desirable property, since the family \mathcal{E} is closed under such transformations. Putting $x \wedge y = \min(x, y)$ for real numbers x, y and using $n^{-1} \sum_{j=1}^n Y_j = 1$, some algebra shows

$$T_{n,a} = \frac{1}{n} \sum_{j,k=1}^n \frac{1}{a} (|Y_j - Y_k| - 1) e^{-aY_j \wedge Y_k} + \frac{1}{a^2} (Y_j + Y_k - 2(Y_j \wedge Y_k + 1)) e^{-aY_j \wedge Y_k} - \frac{2}{a^3} (e^{-aY_j \wedge Y_k} - 1), \quad (5)$$

for each $a > 0$, as well as

$$T_{n,0} = \frac{1}{n} \sum_{j,k=1}^n \frac{1}{3} (Y_j \wedge Y_k)^3 - \frac{Y_j + Y_k - 2}{2} (Y_j \wedge Y_k)^2 + (Y_j - 1)(Y_k - 1)(Y_j \wedge Y_k).$$

An efficient implementation of (5) in the statistical computing language R, see R Core Team (2021), can be found in Appendix B.

- Remark 1.1.**
1. Some algebra shows that $T_{n,a} \equiv G_{n,a}$ holds for all $a > -1$, hence both test statistics are identical.
 2. Representation (4) leads to a new compact and formula of the covariance kernel of the limiting Gaussian process, which admits direct conclusions for the limiting distribution of the test statistic.
 3. Obviously the restriction $a \geq 0$ can be generalised to negative values. In this article we mostly focus on the positive half axis, since the power results in the simulation study of Baringhaus & Henze (2008) suggest that negative tuning parameters lead to low power of the tests. The authors suggest to use the tuning parameter $1 \leq a \leq 2$ as a generally good choice.

4. Note that the test statistic $T_{n,0}$ is related to the test of exponentiality based on the integrated distribution function proposed in Klar (2001), Section 2. The main difference in the definitions of the tests lies in the fact, that $\mathbb{P}(X > y)$ in (3) is estimated by and not replaced with the theoretical known value under H_0 . This difference has an impact on the limiting distribution and the power of the tests as is shown in the sequel.

The rest of the paper is organised as follows. In Section 2 we provide a simple and direct proof of the asymptotic limiting distribution of $T_{n,a}$ under the null hypothesis and provide the first four cumulants of the limiting null distribution as well as explicit formulas of the eigenvalues and eigenfunctions associated with the integral operator induced by the covariance structure of the limiting Gaussian process. These findings corroborate the results of Baringhaus & Henze (2008) in a very short and direct way. In Section 3 we provide new limit results under fixed alternatives as well as a proof of the consistency of the testing procedures. Local approximate Bahadur efficiencies are deduced and interpreted in connection to empirical power results in Section 4. We conclude the article by pointing out open problems for further research in Section 5 and by an Appendix containing explicit formulas and some source code.

2 Limiting distribution under the null hypothesis

In this section we assume that X_1, X_2, \dots is an iid. sequence of random variables with $X_1 \sim \text{Exp}(1)$. Due to the L^2 -structure of the test statistics, a convenient setting –is in dependence of the constant $a \geq 0$ – the separable Hilbert space $\mathbb{H}_a = L^2([0, \infty), \mathcal{B}, \exp(-at)dt)$ of (equivalence classes of) measurable functions $f : [0, \infty) \rightarrow \mathbb{R}$ satisfying $\int_0^\infty |f(t)|^2 \exp(-at) dt < \infty$. Here, \mathcal{B} stands for the Borel sigma-field on $[0, \infty)$. The scalar product and the norm in \mathbb{H}_a will be denoted by

$$\langle f, g \rangle_{\mathbb{H}_a} = \int_0^\infty f(t)g(t) \exp(-at) dt, \quad \|f\|_{\mathbb{H}_a} = \langle f, f \rangle_{\mathbb{H}_a}^{1/2}, \quad f, g \in \mathbb{H}_a,$$

respectively.

A change of variable shows that

$$T_{n,a} = \bar{X}_n^{-3} \int_0^\infty Z_n^2(y) \exp(-ay/\bar{X}_n) dy,$$

where

$$Z_n(y) = \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j - y - \bar{X}_n) \mathbf{1}\{X_j > y\}, \quad y > 0.$$

Define the auxiliary process

$$Z_n^*(y) = \frac{1}{\sqrt{n}} \sum_{j=1}^n (X_j - y - 1) \mathbf{1}\{X_j > y\} - (X_j - 1) \mathbb{P}(X_1 > y), \quad y > 0.$$

Notice that Z_n^* is a sum of centred iid. random variables, and we have $\mathbb{E}[Z_1^*(s)Z_1^*(t)] = K_0(s, t)$, $s, t > 0$, where K_0 defined below. By the central limit theorem in Hilbert spaces there exists a centred Gaussian process Z in \mathbb{H}_a having covariance kernel

$$K_0(s, t) = \exp(-s \vee t) - \exp(-(s+t)), \quad s, t > 0, \quad (6)$$

such that $Z_n^* \xrightarrow{\mathcal{D}} Z$ in \mathbb{H}_a as $n \rightarrow \infty$. Here and in what follows, we write $s \vee t = \max(s, t)$, and $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution. Likewise $\xrightarrow{\mathbb{P}}$ will denote convergence in probability. Note that K_0 is an alternative representation of the limiting covariance kernel ρ in (7) of Baringhaus & Henze (2008).

Theorem 2.1. *Under the standing assumptions, there exists a centred Gaussian process Z in \mathbb{H}_a with covariance kernel K_0 defined in (6), such that*

$$T_{n,a} \xrightarrow{\mathcal{D}} \|Z\|_{\mathbb{H}_a}^2, \quad \text{as } n \rightarrow \infty.$$

Proof. We first consider the case $a = 0$. Note that

$$\|Z_n - Z_n^*\|_{\mathbb{H}_0}^2 = \left(\frac{1}{n} \sum_{j=1}^n (X_j - 1) \right)^2 \int_0^\infty \left| \frac{1}{\sqrt{n}} \sum_{k=1}^n (\mathbf{1}\{X_k > y\} - \mathbb{P}(X_1 > y)) \right|^2 dy.$$

By the strong law of large numbers $n^{-1} \sum_{j=1}^n (X_j - 1)$ converges to 0 a.s., and the central limit theorem in \mathbb{H}_0 implies that the second term is a tight sequence. Hence, $\|Z_n - Z_n^*\|_{\mathbb{H}_0}^2 = o_{\mathbb{P}}(1)$, and $T_{n,0} \xrightarrow{\mathcal{D}} \|Z\|_{\mathbb{H}_0}^2$, as $n \rightarrow \infty$ follows from $Z_n^* \xrightarrow{\mathcal{D}} Z$ in \mathbb{H}_0 , Slutski's lemma and the continuous mapping theorem.

For the case $a > 0$ define

$$\tilde{T}_{n,a} = \bar{X}_n^{-3} \int_0^\infty Z_n^2(y) \exp(-ay) \, dy.$$

A first-order Taylor expansion yields

$$\exp(-ay/\bar{X}_n) = \exp(-ay) + \frac{ay}{\Delta_n^2} \exp\left(-\frac{ay}{\Delta_n}\right) (\bar{X}_n - 1), \quad (7)$$

where $\Delta_n \in (\min(\bar{X}_n, 1), \max(\bar{X}_n, 1))$. From the Cauchy-Schwarz inequality and (7) we obtain

$$\begin{aligned} |T_{n,a} - \tilde{T}_{n,a}| &= \bar{X}_n^{-3} \int_0^\infty Z_n^2(y) |\exp(-ay/\bar{X}_n) - \exp(-ay)| \, dy \\ &\leq \bar{X}_n^{-3} \int_0^\infty Z_n^2(y) \, dy \int_0^\infty |\exp(-ay/\bar{X}_n) - \exp(-ay)| \, dy \\ &= \bar{X}_n^{-3} T_{n,0} \int_0^\infty \frac{ay}{\Delta_n^2} \exp\left(-\frac{ay}{\Delta_n}\right) |\bar{X}_n - 1| \, dy \\ &= \frac{1}{a} \bar{X}_n^{-3} T_{n,0} |\bar{X}_n - 1|. \end{aligned}$$

The strong law of large numbers implies $\bar{X}_n \rightarrow 1$ a.s., and since $T_{n,0}$ is a tight sequence (see case $a = 0$), $|T_{n,a} - \tilde{T}_{n,a}| = o_{\mathbb{P}}(1)$ follows. By the central limit theorem in \mathbb{H}_a there exists a centred Gaussian process \tilde{Z} with covariance kernel K_0 such that $Z_n^* \xrightarrow{\mathcal{D}} \tilde{Z}$ as $n \rightarrow \infty$ in \mathbb{H}_a and the same reasoning as for the case $a = 0$ yields $\|Z_n - Z_n^*\|_{\mathbb{H}_a}^2 = o_{\mathbb{P}}(1)$. Invoking Slutski's lemma and the continuous mapping theorem, we obtain $\tilde{T}_{n,a} \xrightarrow{\mathcal{D}} \|\tilde{Z}\|_{\mathbb{H}_a}^2$ from which the claim follows. \square

Remark 2.2. By direct evaluation of integrals the first two cumulants of the distribution of $\|Z\|_{\mathbb{H}_a}^2$ are

$$\kappa_{1,a} = \mathbb{E}\|Z\|_{\mathbb{H}_a}^2 = \int_0^\infty K_0(t, t) \exp(-at) \, dt = \frac{1}{(a+1)(a+2)}$$

and

$$\kappa_{2,a} = \mathbb{V}(\|Z\|_{\mathbb{H}_a}^2) = 2 \int_0^\infty \int_0^\infty K_0^2(s, t) \exp(-a(s+t)) \, ds \, dt = \frac{2}{(a^2 + 3a + 2)(2a + 3)(a + 2)},$$

where $\mathbb{V}(\cdot)$ denotes the variance. Following the methodology in Henze (1990); Stephens (1976), we calculate the third and fourth cumulants according to

$$\kappa_{j,a} = 2^{j-1} (j-1)! \int_0^\infty K_j(t, t) \exp(-at) \, dt,$$

where $K_j(s, t)$, the j^{th} iterate of $K_0(s, t)$, is given by

$$\begin{aligned} K_j(s, t) &= \int_0^\infty K_{j-1}(s, u) K_0(u, t) \exp(-au) \, du, \quad j \geq 2, \\ K_1(s, t) &= K_0(s, t). \end{aligned}$$

Direct calculation shows that

$$\kappa_{3,a} = \frac{16}{(1+a)(2a+3)(a+2)^3(3a+4)}$$

and

$$\kappa_{4,a} = \frac{48(11a+16)}{(3a+4)(2a+3)^2(a+2)^4(1+a)(4a+5)}.$$

These formulas are very useful in order to fit a Pearson system of distributions to approximate the critical values of the test statistic $T_{n,a}$. The approximated quantiles of the limiting distribution are found in Table 1 and source code is provided in Appendix B. A comparison with Table 1 in Baringhaus & Henze (2008) shows that the approximation of the critical values is a reasonable fit to the empirical critical values for $a \geq 0$.

q/a	-0.99	-0.9	-0.5	0	0.5	1	1.5	2	5	10
0.9	117.263	13.919	2.522	1.009	0.553	0.351	0.243	0.178	0.052	0.017
0.95	123.673	16.162	3.189	1.309	0.725	0.463	0.322	0.237	0.069	0.022
0.99	136.942	21.396	4.813	2.045	1.149	0.739	0.516	0.381	0.112	0.036

Table 1: Approximated q -quantiles of the limiting distribution in Theorem 2.1 by a Pearson system of distributions using the formulas of the cumulants in Remark 2.2.

From the theory of Gaussian processes it is well known (see Shorack & Wellner (1986), p. 206) that for each $a > 0$ an orthogonal decomposition of the process Z yields $\|Z\|_{\mathbb{H}_a}^2 = \sum_{j=1}^{\infty} \lambda_j(a) N_j^2$, where N_1, N_2, \dots are iid. standard normal, and $\lambda_1(a), \lambda_2(a), \dots$ is the decreasing sequence of positive eigenvalues of the integral operator $\mathcal{K}_a : \mathbb{H}_a \rightarrow \mathbb{H}_a$, $f \mapsto \mathcal{K}_a f(\cdot) = \int_0^{\infty} K_0(\cdot, t) f(t) \exp(-at) dt$. To calculate the eigenvalues $\lambda_j(a)$, $j = 1, 2, \dots$, of \mathcal{K}_a , one has to solve the homogeneous Fredholm integral equation of the second kind

$$\int_0^{\infty} K_0(x, t) f(t) \exp(-at) dt = \lambda f(x), \quad x > 0,$$

see, for example, Kac & Siegert (1947). Usually these problems are very hard to solve explicitly and numerical or Monte Carlo simulation techniques are used to obtain approximations of $\lambda_j(a)$, see Section 5 in Ebner & Henze (2021) for a stochastic approximation method or the method presented in Božin et al. (2020).

In the following we give explicit formulas for general tuning parameters a . Firstly, note that K_0 admits the representation

$$K_0(s, t) = \min(\exp(-s), \exp(-t)) - \exp(-s) \exp(-t), \quad s, t > 0. \quad (8)$$

Let $a = 1$. Using the alternative representation in (8) together with $x = \exp(-s)$ and a substitution ($y = \exp(-t)$), we have

$$\int_0^{\infty} K_0(x, t) f(t) \exp(-t) dt = \int_0^1 (\min(x, y) - xy) f(-\log(y)) dy.$$

Since $\mathbb{K}(x, y) = \min(x, y) - xy$, $x, y \in [0, 1]$, is the covariance kernel of the Brownian bridge $B(t)$, say, the eigenvalue problem is solved in this case, see e.g. Anderson & Darling (1952). We conclude that the eigenvalues and corresponding eigenfunctions are

$$\lambda_k(1) = \frac{1}{(k\pi)^2} \quad \text{and} \quad f_{k,1}(t) = \sin(k\pi \exp(-t)), \quad t > 0.$$

By analogy the general case $a > 0$ leads to

$$\int_0^{\infty} K_0(x, t) f(t) \exp(-at) dt = \int_0^1 (\min(x, y) - xy) f(-\log(y)) y^{a-1} dy,$$

which is connected to the eigenvalue problem of the weighted Brownian bridge $t^{a-1}B(t)$. This problem is solved in Theorem 1.4 in Deheuvels & Martynov (2003), which states the eigenvalues and eigenfunctions explicitly. Let $\nu = (a + 1)^{-1}$ (note that there is a typographical error in Deheuvels & Martynov (2003)) and denote by J_ν the Bessel functions of the first kind and by $0 < z_{\nu,1} < z_{\nu,2} < \dots$ the ascending sequence of zeros of J_ν , for details on the zeros of Bessel functions see Watson (1995), Chapter XV. We have for $k = 1, 2, \dots$

$$\lambda_k(a) = \left(\frac{2\nu}{z_{\nu,k}} \right)^2 \quad \text{and} \quad f_{k,a}(t) = \frac{J_\nu(z_{\nu,k} \exp(-\frac{t}{2\nu}))}{\sqrt{\nu} J_{\nu-1}(z_{\nu,k})} \exp\left(-\left(\frac{1}{\nu} - 1\right) \frac{t}{2}\right), \quad t > 0. \quad (9)$$

By using the identities in Remark 2.2 (compare to Corollary 1.3 in Deheuvels & Martynov (2003)) we see that $\sum_{k=1}^{\infty} \lambda_k^j(a) = \kappa_{j,a} / (2^{j-1} (j-1)!)$, $j = 1, 2, 3, 4$. The largest twenty eigenvalues are given in Table 2, and for the sake of completeness, we approximated the scaled cumulants by the sum of the respective powers of the first 100 eigenvalues. We see that for the mean, there is still some difference in the first few digits, which is explained by the low speed of convergence to 0 of the eigenvalues. Interestingly, the eigenvalues in Table 2 in each row are strictly decreasing, which is explained by the fact that for fixed k the function $z_{\nu,k}/\nu$ is increasing for $\nu \downarrow 0$, see Olver et al. (2010), p. 236.

Remark 2.3. The stated formulas of eigenvalues and eigenfunctions were also derived in Baringhaus & Henze (2000) for the case $a = 1$ relating the statistics to the classical Cramér-von Mises test and for the general case $a > -1$ by solving a related differential equation in Baringhaus & Henze (2008). From these results we see that (9) also holds in these cases.

k/a	0	1	2	3	4	5
1	0.2724430	0.10132118	0.05275301	0.03232757	0.02183334	0.015732912
2	0.0812703	0.02533030	0.01221201	0.00716691	0.00470790	0.003327524
3	0.0386475	0.01125791	0.00528483	0.00305755	0.00199068	0.001398474
4	0.0225326	0.00633257	0.00293284	0.00168481	0.00109214	0.000764966
5	0.0147448	0.00405285	0.00186176	0.00106499	0.00068857	0.000481445
6	0.0103955	0.00281448	0.00128585	0.00073348	0.00047341	0.000330626
7	0.0077217	0.00206778	0.00094101	0.00053570	0.00034534	0.000240981
8	0.0059612	0.00158314	0.00071835	0.00040833	0.00026299	0.000183403
9	0.0047409	0.00125088	0.00056629	0.00032151	0.00020693	0.000144239
10	0.0038603	0.00101321	0.00045785	0.00025971	0.00016705	0.000116401
11	0.0032042	0.00083737	0.00037782	0.00021415	0.00013768	0.000095907
12	0.0027021	0.00070362	0.00031708	0.00017960	0.00011543	0.000080385
13	0.0023095	0.00059953	0.00026989	0.00015279	0.00009817	0.000068347
14	0.0019966	0.00051694	0.00023250	0.00013156	0.00008450	0.000058824
15	0.0017433	0.00045032	0.00020237	0.00011447	0.00007351	0.000051161
16	0.0015352	0.00039579	0.00017774	0.00010050	0.00006452	0.000044903
17	0.0013624	0.00035059	0.00015735	0.00008895	0.00005709	0.000039727
18	0.0012171	0.00031272	0.00014028	0.00007927	0.00005088	0.000035397
19	0.0010939	0.00028067	0.00012584	0.00007110	0.00004562	0.000031738
20	0.0009885	0.00025330	0.00011352	0.00006412	0.00004114	0.000028618
$\sum_{j=1}^{100} \lambda_j(a)$	0.4959773	0.16565850	0.08288489	0.04974765	0.03317179	0.023697320
$\kappa_{1,a}$	0.5000000	0.16666667	0.08333333	0.05000000	0.03333333	0.023809524
$\sum_{j=1}^{100} \lambda_j^2(a)$	0.08333333	0.01111111	0.00297619	0.00111111	0.00050505	0.000261643
$\kappa_{2,a}/2$	0.08333333	0.01111111	0.00297619	0.00111111	0.00050505	0.000261643
$\sum_{j=1}^{100} \lambda_j^3(a)$	0.0208333	0.00105820	0.00014881	0.00003419	0.00001052	0.000003934
$\kappa_{3,a}/8$	0.0208333	0.00105820	0.00014881	0.00003419	0.00001052	0.000003934
$\sum_{j=1}^{100} \lambda_j^4(a)$	0.0055556	0.00010582	0.00000777	0.00000109	0.00000023	0.000000061
$\kappa_{4,a}/48$	0.0055556	0.00010582	0.00000777	0.00000109	0.00000023	0.000000061

Table 2: First twenty eigenvalues $\lambda_k(a)$ for different tuning parameters a and sums of powers over the first 100 eigenvalues and the corresponding cumulants using the formulas in Remark 2.2.

3 Limiting distribution under fixed alternatives and consistency

In this section we assume that X_1, X_2, \dots is a sequence of iid. random variables with cumulative distribution function F , $\mathbb{E}[X_1] = 1$ and $\mathbb{E}[|X_1|^3] < \infty$. The moment condition is motivated by the scale invariance of $T_{n,a}$ and hence there is no loss of generality compared to fixed alternatives with existing third moment.

Theorem 3.1. *We have*

$$\frac{T_{n,a}}{n} \xrightarrow{\mathbb{P}} \int_0^\infty z^2(y) \exp(-ay) \, dy = \Delta_a, \quad a \geq 0, \quad \text{as } n \rightarrow \infty,$$

where $z(y) = \mathbb{E}[(X_1 - y - 1)\mathbf{1}\{X_1 > y\}]$, $y > 0$.

Proof. Note that by the same arguments as in the proof of Theorem 2.1 we have $n^{-1}\|Z_n - Z_n^*\|_{\mathbb{H}_a}^2 \rightarrow 0$ a.s. Since by the strong law of large numbers in Hilbert spaces $n^{-1/2}Z_n^* \rightarrow z$ a.s. in \mathbb{H}_a the claim follows by the continuous mapping theorem and Slutski's lemma. \square

Under the assumption $X_1 \sim \text{Exp}(1)$, the characterisation in (3) gives $z \equiv 0$ and hence $\Delta_a = 0$. Since Δ_a equals 0 if and only if X_1 follows an exponential distribution, we conclude that the tests $T_{n,a}$ are consistent against each alternative distribution with existing first moment.

Example 3.2. In this example we give explicit formulas for Δ_a for different $\Gamma(\beta, \beta)$ distributions. Note that for an suitable alternative we need $\mathbb{E}[X] = 1$ to be satisfied. Direct evaluations show that

1. if $X \sim \Gamma(2, 2)$, then $\Delta_a = 2/(a + 4)^3$,

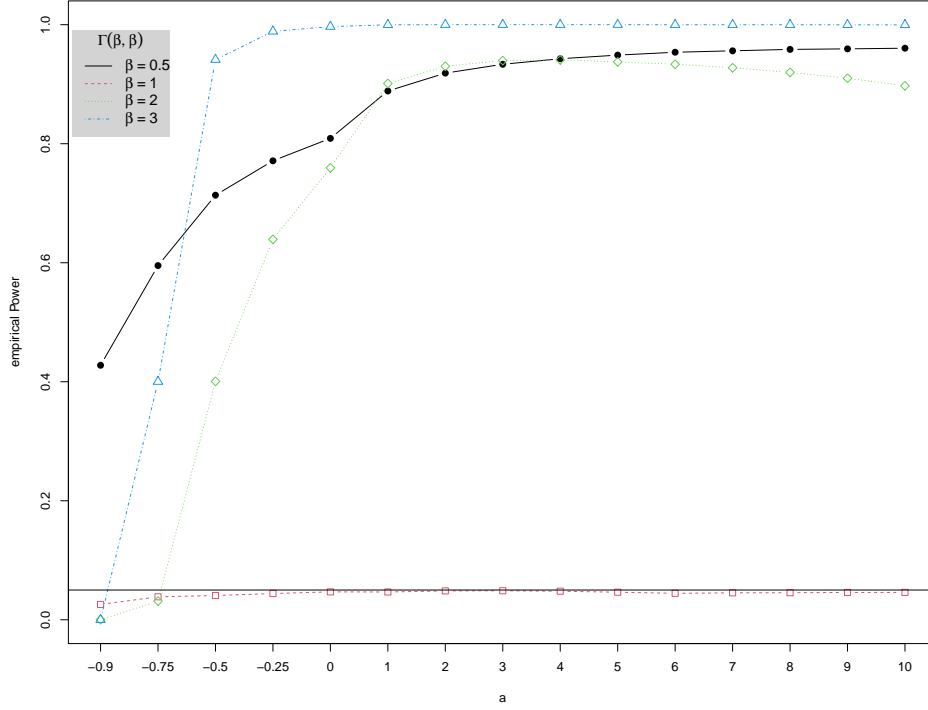


Figure 1: Empirical power of the test $T_{n,a}$ for different tuning parameters a . Each dot represents the empirical rejection rate of the test at a significance level of 0.05 under the stated $\Gamma(\beta, \beta)$ distribution (MC simulation with 10000 replications).

2. if $X \sim \Gamma(3, 3)$, then $\Delta_a = 2(a^2 + 30a + 252)/(a + 6)^5$, and

3. if $X \sim \Gamma(4, 4)$, then $\Delta_a = 2(a^4 + 56a^3 + 1344a^2 + 16640a + 94720)/(a + 8)^7$.

Remark 3.3. Note that higher values of Δ_a in dependence of the tuning parameter a do not imply greater power of the test against this alternative. To visualise this behaviour, we performed a Monte Carlo (MC) simulation study with different $\Gamma(\beta, \beta)$ distributions and plot the empirical rejection rate of the tests in Figure 1. Note that under $\Gamma(1, 1) = \text{Exp}(1)$, hence under the null hypothesis, the type I error is well calibrated below the significance level of 0.05 (here visualised by a solid line).

In the following we derive the limiting distribution of the test statistic under fixed alternatives under the stated assumptions in the beginning of the section. Define

$$W_n(y) = Z_n(y) - \sqrt{n}z(y), \quad W_n^*(y) = Z_n^*(y) - \sqrt{n}z(y), \quad y > 0.$$

It is easy to see that $\|W_n - W_n^*\|_{\mathbb{H}_a}^2 = o_{\mathbb{P}}(1)$.

Lemma 3.4. Under the standing assumptions there exists a centred Gaussian process W in \mathbb{H}_a with covariance kernel

$$\begin{aligned} K(s, t) &= \Psi_2(s \vee t) + st\mathbb{P}(X > s \vee t) - (s + t)\Psi_1(s \vee t) \\ &\quad - \mathbb{P}(X > s)(\Psi_2(t) - t\Psi_1(t)) - \mathbb{P}(X > t)(\Psi_2(s) - s\Psi_1(s)) \\ &\quad + \mathbb{V}(X)\mathbb{P}(X > s)\mathbb{P}(X > t) - z(s)z(t), \quad s, t > 0, \end{aligned}$$

where $s \vee t = \max(s, t)$, and $\Psi_\ell(s) = \mathbb{E}[(X - 1)^\ell \mathbf{1}\{X > s\}]$, $s > 0$, for $\ell = 1, 2$. In \mathbb{H}_a we have

$$W_n \xrightarrow{\mathcal{D}} W \quad \text{as } n \rightarrow \infty.$$

Proof. Since $\|W_n - W_n^*\|_{\mathbb{H}_a}^2 = o_{\mathbb{P}}(1)$ the limit process is determined by the limit behaviour of W_n^* . We have

$$W_n^*(y) = \frac{1}{\sqrt{n}} \sum_{j=1}^n [(X_j - y - 1)\mathbf{1}\{X_j > y\} - (X_j - 1)\mathbb{P}(X_j > y) - z(y)], \quad y \geq 0,$$

and $\mathbb{E}[W_n^*(y)] = 0$, $y > 0$, since $\mathbb{E}[X_1] = 1$. Hence W_n^* is a sum of iid. elements in \mathbb{H}_a and thus converges as $n \rightarrow \infty$ to a centred Gaussian process W with covariance kernel $K(s, t) = \mathbb{E}[W_1^*(s)W_1^*(t)]$, $s, t \geq 0$, in \mathbb{H}_a by the central limit theorem in Hilbert spaces. The formula for K is obtained by tedious but straightforward calculations. \square

By the results of Lemma 3.4 and Theorem 1 in Baringhaus et al. (2017), we thus have

$$\sqrt{n} \left(\frac{T_{n,a}}{n} - \Delta_a \right) \xrightarrow{\mathcal{D}} \mathbf{N}(0, \sigma_a^2), \quad (10)$$

where

$$\sigma_a^2 = 4 \int_0^\infty \int_0^\infty K(x, y) z(x) z(y) \exp(-a(x+y)) \, dx dy.$$

For some families of distributions, Δ_a and σ_a^2 can be calculated explicitly for fixed tuning parameters see Example 3.2.

In general, however, we have to find a consistent estimator $\hat{\sigma}_{n,a}^2$ of σ_a^2 . In this spirit, we replace the probabilities by relative frequencies, the variance by the empirical variance $s_y^2 = \frac{1}{n} \sum_{j=1}^n (Y_j - 1)^2$, and the expectation by empirical counterparts

$$\Psi_{n,\ell}(s) = \frac{1}{n} \sum_{j=1}^n (Y_j - 1)^\ell \mathbf{1}\{Y_j > s\}, \quad \ell = 1, 2, \quad \text{and} \quad z_n(s) = \frac{1}{n} \sum_{j=1}^n (Y_j - s - 1) \mathbf{1}\{Y_j > s\}, \quad s > 0,$$

based on the scaled random variables Y_1, \dots, Y_n . Denoting by K_n the resulting estimator of K obtained from plugging-in the empirical counterparts into the formula of K , the estimator $\hat{\sigma}_{n,a}^2$ of σ_a^2 is

$$\hat{\sigma}_{n,a}^2 = 4 \int_0^\infty \int_0^\infty K_n(x, y) z_n(x) z_n(y) \exp(-a(x+y)) \, dx dy.$$

Writing $\hat{\tau}_n^2 = n^{-1} \sum_{j=1}^n (Y_j - 1)^2$, and

$$\begin{aligned} S_{1n} &= \int_0^\infty \int_0^\infty \Psi_{n,2}(s \vee t) z_n(s) z_n(t) \exp(-a(s+t)) \, dt ds, \\ S_{2n} &= \frac{1}{n} \sum_{j=1}^n \int_0^\infty \int_0^\infty st \mathbf{1}\{Y_j > s \vee t\} z_n(s) z_n(t) \exp(-a(s+t)) \, dt ds \\ S_{3n} &= \int_0^\infty \int_0^\infty (s+t) \Psi_{n,1}(s \vee t) z_n(s) z_n(t) \exp(-a(s+t)) \, dt ds, \\ S_{4n} &= \int_0^\infty (\Psi_{n,2}(t) - t \Psi_{n,1}(t)) z_n(t) \exp(-at) \, dt \\ S_{5n} &= \frac{1}{n} \sum_{j=1}^n \int_0^\infty \mathbf{1}\{Y_j > t\} z_n(t) \exp(-at) \, dt, \\ S_{6n} &= \int_0^\infty z_n^2(t) \exp(-at) \, dt, \end{aligned}$$

we have

$$\hat{\sigma}_{n,a}^2 = 4 [S_{1n} + S_{2n} - S_{3n} - 2S_{4n}S_{5n} + \hat{\tau}_n^2 S_{5n}^2 - S_{6n}^2].$$

Using the functions $v_{\ell,a}(x, y)$, $x, y > 0$, $\ell = 1, 2, 3$, from Appendix A, we have

$$\begin{aligned} S_{1n} &= \frac{1}{n^3} \sum_{j,k,l=1}^n (Y_j - 1)^2 v_{1,a}(Y_j, Y_k) v_{1,a}(Y_j, Y_l), \\ S_{2n} &= \frac{1}{n^3} \sum_{j,k,l=1}^n v_{2,a}(Y_j, Y_k) v_{2,a}(Y_j, Y_l), \\ S_{3n} &= \frac{2}{n^3} \sum_{j,k,l=1}^n (Y_j - 1) v_{2,a}(Y_j, Y_k) v_{1,a}(Y_j, Y_l), \\ S_{4n} &= \frac{1}{n^2} \sum_{j,k=1}^n [(Y_j - 1)^2 v_{1,a}(Y_j, Y_k) - (Y_j - 1) v_{2,a}(Y_j, Y_k)], \\ S_{5n} &= \frac{1}{n^2} \sum_{j,k=1}^n v_{1,a}(Y_j, Y_k), \quad \text{and} \quad S_{6n} = \frac{1}{n^2} \sum_{j,k=1}^n v_{3,a}(Y_j, Y_k). \end{aligned}$$

Since $\widehat{\sigma}_{n,a}^2$ is a consistent sequence of estimators of σ_a^2 for each fixed tuning parameter $a > 0$, we have (in the spirit of Corollary 1 in Baringhaus et al. (2017)) under the stated assumptions at the beginning of this section

$$\frac{\sqrt{n}}{\widehat{\sigma}_{n,a}} \left(\frac{T_{n,a}}{n} - \Delta_a \right) \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } n \rightarrow \infty. \quad (11)$$

This result has immediate consequences, see Section 3 of Baringhaus et al. (2017), which are detailed in the following subsections.

3.1 A confidence interval for Δ_a

For $\alpha \in (0, 1)$ let $u_\alpha = \Phi^{-1}(1 - \alpha/2)$ be the $(1 - \alpha/2)$ -quantile of the standard normal law. Putting

$$I_{n,a} = \left[\frac{T_{n,a}}{n} - \frac{u_\alpha \widehat{\sigma}_{n,a}}{\sqrt{n}}, \frac{T_{n,a}}{n} + \frac{u_\alpha \widehat{\sigma}_{n,a}}{\sqrt{n}} \right], \quad a \geq 0,$$

it follows from (11) that

$$\lim_{n \rightarrow \infty} \mathbb{P}_F(I_{n,a} \ni \Delta_a) = 1 - \alpha.$$

Hence $I_{n,a}$ is an asymptotic confidence interval at confidence level $1 - \alpha$ for Δ_a . In the following we revisit the gamma distributions of Example 3.2. Specific values of Δ_a , σ_a^2 and the corresponding estimators for $\Gamma(\beta, \beta)$ and $\beta \in \{2, 3, 4, 5, 10\}$ are found in Table 3. The discrepancy between estimators and true values are in line with the results of Table 5 in Baringhaus et al. (2017), which indicates a rather slow convergence to the limiting distribution under alternatives in (11). In Table 4 we present the empirical coverage probabilities of $I_{n,a}$ for Δ_a in the same setting as in Table 3. Critical values have been obtained by the Pearson-system approximation as presented in Table 1. Interestingly the confidence interval seems to be conservative, since the estimated probability of coverage of the true value Δ_a is disproportionately high in most cases. This is in contrast to the findings in Table 6 in Baringhaus et al. (2017), where lower coverage rates than indicated by the nominal level were reported.

3.2 Neighbourhood-of-model validation

A clear drawback in the field of goodness-of-fit testing is that if a level- α -test does not lead to the rejection of the hypothesis H_0 , the conclusion that H_0 is 'confirmed' is generally wrong. To overcome this problem, the results of (10) lead to a so-called 'neighbourhood-of-model validation', see Subsection 3.3 in Baringhaus et al. (2017). In this spirit one could see Δ_a as some sort of distance to the null hypothesis. If we argue to 'tolerate' a given value $\widetilde{\Delta}$, we can consider the testing problem

$$H_{\widetilde{\Delta}} : \Delta_a(F) \geq \widetilde{\Delta} \quad \text{versus} \quad K_{\widetilde{\Delta}} : \Delta_a(F) < \widetilde{\Delta}.$$

From (10) we obtain an asymptotic level- α -test by rejecting $H_{\widetilde{\Delta}}$ whenever

$$T_{n,a} \leq n\widetilde{\Delta} - \sqrt{n}\widehat{\sigma}_{n,a}\Phi^{-1}(1 - \alpha).$$

To prove this statement follow the reasoning in Baringhaus et al. (2017), subsection 3.3.

a/β	2	3	4	5	10
Δ_0	0.0312	0.0648	0.0903	0.1100	0.1661
$T_{n,0}/n$	0.0252	0.0672	0.0944	0.1179	0.1778
σ_0^2	0.0178	0.0300	0.0357	0.0382	0.0362
$\hat{\sigma}_{n,0}^2$	0.0134	0.0313	0.0361	0.0377	0.0331
Δ_1	0.0160	0.0337	0.0472	0.0575	0.0863
$T_{n,1}/n$	0.0135	0.0349	0.0493	0.0619	0.0923
σ_1^2	0.0039	0.0065	0.0076	0.0080	0.0070
$\hat{\sigma}_{n,1}^2$	0.0031	0.0069	0.0076	0.0080	0.0062
Δ_2	0.0092	0.0193	0.0268	0.0324	0.0475
$T_{n,2}/n$	0.0081	0.0199	0.0280	0.0348	0.0506
σ_2^2	0.0012	0.0018	0.0020	0.0020	0.0015
$\hat{\sigma}_{n,2}^2$	0.0010	0.0019	0.0019	0.0020	0.0013
Δ_3	0.0058	0.0119	0.0162	0.0194	0.0276
$T_{n,3}/n$	0.0053	0.0122	0.0170	0.0208	0.0292
σ_3^2	0.0004	0.0006	0.0006	0.0006	0.0003
$\hat{\sigma}_{n,3}^2$	0.0004	0.0006	0.0006	0.0006	0.0003
Δ_4	0.0040	0.0078	0.0104	0.0123	0.0168
$T_{n,4}/n$	0.0036	0.0079	0.0109	0.0130	0.0177
σ_4^2	0.0002	0.0002	0.0002	0.0002	0.0001
$\hat{\sigma}_{n,4}^2$	0.0002	0.0002	0.0002	0.0002	0.0001

Table 3: Values of Δ_a and σ_a^2 for different $\Gamma(\beta, \beta)$ distributions and simulated observations of $T_{n,a}/n$ and $\hat{\sigma}_{n,a}^2$ for different tuning parameters a ($n = 1000$).

4 Asymptotic Bahadur efficiencies

In this section we only consider the case that $a > 0$. We start the investigation by giving an alternative representation of Δ_a in Theorem 3.1.

Lemma 4.1. *Suppose that $\mu = \mathbb{E}[X]$ and $\mathbb{E}[|X|^3] < \infty$. If $a > 0$, we have*

$$\frac{T_{n,a}}{n} \xrightarrow{\mathbb{P}} \Delta_a = \frac{2}{a^3} - \frac{1}{a^3} \mathbb{E} \left[((1 - |Y_1 - Y_2|)(a^2 + a) + a + 2) \exp(-aY_1 \wedge Y_2) \right],$$

and if $a = 0$, we have

$$\frac{T_{n,0}}{n} \xrightarrow{\mathbb{P}} \Delta_0 = \mathbb{E} \left[(Y_1 \wedge Y_2)(Y_1 - 1)(Y_2 - 1) + (Y_1 \wedge Y_2)^2 \left(1 - \frac{1}{2}(Y_1 + Y_2) \right) + \frac{1}{3}(Y_1 \wedge Y_2)^3 \right],$$

where $Y_j = X_j/\mu$, $j = 1, 2$.

Proof. By Theorem 3.1, applied to the iid. random variables Y_1, \dots, Y_n with $\mathbb{E}[Y_1] = 1$, we have

$$\frac{T_n}{n} \xrightarrow{\mathbb{P}} \Delta_a = \int_0^\infty z^2(y) \exp(-ay) \, dy,$$

where $z(y) = \mathbb{E}[(Y_1 - y - 1)\mathbf{1}\{Y_1 > y\}]$, $y > 0$. Since Y_1 and Y_2 are independent, it follows that

$$z^2(y) = \mathbb{E}[(Y_1 - y - 1)(Y_2 - y - 1)\mathbf{1}\{Y_1 \wedge Y_2 > y\}].$$

a	n	$\beta = 2$			$\beta = 3$			$\beta = 4$			$\beta = 5$			$\beta = 10$		
		20	50	100	20	50	100	20	50	100	20	50	100	20	50	100
0		0.95	0.95	0.94	0.96	0.96	0.97	0.96	0.98	0.98	0.97	0.98	0.99	0.98	0.99	0.99
1		0.90	0.92	0.93	0.92	0.95	0.95	0.93	0.96	0.96	0.94	0.96	0.97	0.96	0.98	0.99
2		0.87	0.90	0.91	0.89	0.93	0.94	0.90	0.94	0.95	0.91	0.94	0.96	0.93	0.96	0.97
3		0.85	0.89	0.91	0.85	0.91	0.92	0.87	0.92	0.93	0.87	0.92	0.94	0.90	0.94	0.95
4		0.82	0.88	0.90	0.83	0.90	0.92	0.84	0.90	0.92	0.85	0.90	0.93	0.87	0.92	0.94

Table 4: Empirical relative frequencies of coverage of $I_{n,a}$ for Δ_a at a nominal level of 0.9 for different $\Gamma(\beta, \beta)$ distributions (10000 replications).

Fubini's Theorem yields

$$\begin{aligned}\Delta_a &= \int_0^\infty \mathbb{E}[(Y_1 - y - 1)(Y_2 - y - 1)\mathbf{1}\{Y_1 \wedge Y_2 > y\}] \exp(-ay) dy \\ &= \mathbb{E} \left[\int_0^{Y_1 \wedge Y_2} (Y_1 - y - 1)(Y_2 - y - 1) \exp(-ay) dy \right].\end{aligned}$$

Straightforward integration and some algebra concludes the proof in both cases. \square

Remark 4.2. An alternative way of proving Lemma 4.1 is to start by the representation (5) and to use symmetry arguments, the law of large numbers and Lebesgue's dominated convergence theorem, i.e. to adapt the lines of proof of Theorem 3.1 in Ebner & Henze (2021).

A useful tool for a theoretical comparison of the performance of two tests is the asymptotic relative Bahadur efficiency. This concept has been used throughout the literature on exponentiality tests, see Cuparić et al. (2019); Jovanović et al. (2015); Milošević (2016); Milošević & Obradović (2016a); Volkova & Nikitin (2015), and for more details on the theory we refer to Bahadur (1971); Nikitin (1995). For a brief introduction to the concept, see Section 5 of Jiménez-Gamero et al. (2020). In this spirit and using the same notations, we calculate the local approximate Bahadur slope of $T_{n,a}$. Assume that $0 \in \Theta \subset \mathbb{R}$, where Θ is an open parameter space, and $\mathcal{G} = \{G(x; \vartheta) : \vartheta \in \Theta\}$ is a family of distribution functions with density $g(x; \vartheta)$, such that $\vartheta = 0$ corresponds to the standard exponential density $g(x; 0) = \exp(-x)$, $x > 0$, and for each $\vartheta > 0$ the density $g(x; \vartheta)$ is not a density corresponding to an exponential distribution in \mathcal{E} . Moreover, we assume that the regularity assumptions WD in Nikitin & Peaucelle (2004) are satisfied. In the following assume that X_1, X_2, \dots are independent identical copies of X following the distribution with density $g(\cdot; \vartheta)$ with existing expectation $\mathbb{E}[X] = \mu(\vartheta) = \int_0^\infty xg(x; \vartheta)dx$. Then Lemma 4.1 yields

$$\frac{T_{n,a}}{n} \xrightarrow{\mathbb{P}_\vartheta} b_{T_a}(\vartheta),$$

where $\xrightarrow{\mathbb{P}_\vartheta}$ denotes convergence in probability under the true parameter ϑ , and

$$b_{T_a}(\vartheta) = \int_0^\infty \int_0^\infty h_a(x, y; \vartheta)g(x; \vartheta)g(y; \vartheta)dx dy$$

with

$$a^3 h_a(x, y; \vartheta) = 2 - \left(\left(1 - \frac{|x - y|}{\mu(\vartheta)} \right) (a^2 + a) + a + 2 \right) \exp\left(-\frac{a}{\mu(\vartheta)}x \wedge y\right), \quad x, y > 0, \quad a > 0,$$

and

$$h_0(x, y; \vartheta) = \frac{x \wedge y}{\mu(\vartheta)} \left(\frac{x}{\mu(\vartheta)} - 1 \right) \left(\frac{y}{\mu(\vartheta)} - 1 \right) + \left(\frac{x \wedge y}{\mu(\vartheta)} \right)^2 \left(1 - \frac{x + y}{2\mu(\vartheta)} \right) + \frac{1}{3} \left(\frac{x \wedge y}{\mu(\vartheta)} \right)^3, \quad x, y > 0.$$

Note that $\mu(0) = 1$, $b_{T_a}(0) = 0$, and after some algebra we have $b'_{T_a}(0) = 0$. Here and in the following, all derivatives are calculated w.r.t. ϑ . Writing $\mu_1 = \mu'(0)$, the same reasoning as in Appendix B of Cuparić et al. (2021) gives

$$\begin{aligned}b''_{T_a}(0) &= 2 \int_0^\infty \int_0^\infty h_a(x, y; 0)g'(x; 0)g'(y; 0)dx dy + 4\mu_1 \int_0^\infty \int_0^\infty h'_a(x, y; 0)g(x; 0)g'(y; 0)dx dy \\ &\quad + \mu_1^2 \int_0^\infty \int_0^\infty h''_a(x, y; 0)g(x; 0)g(y; 0)dx dy.\end{aligned}$$

Expanding $b_{T_a}(\vartheta)$ into a Taylor series around $\vartheta_0 = 0$, we obtain

$$b_{T_a}(\vartheta) = \frac{b''_{T_a}(0)}{2} \vartheta^2 + O(\vartheta^3), \quad \text{as } \vartheta \rightarrow 0.$$

From Section 2 we know that the limiting distribution of T_n is $\|Z\|_{\mathbb{H}_a}^2 = \sum_{j=1}^\infty \lambda_j(a)N_j^2$, where N_1, N_2, \dots are iid. standard normal, and $(\lambda_j(a))_{j \in \mathbb{N}}$ is the decreasing sequence of positive eigenvalues of the integral operator \mathcal{K}_a . Using the result in Zolotarev (1961), the logarithmic tail behaviour of the limiting distribution of $\tilde{T}_{n,a} = \sqrt{T_{n,a}}$ is

$$\log \left(1 - F_{\tilde{T}_a}(x) \right) = -\frac{x^2}{2\lambda_1(a)} + O(x^2), \quad x \rightarrow \infty.$$

Alt./ a	0	1	2	3	4	5
Weibull	0.722	0.834	0.865	0.868	0.859	0.843
Gamma	0.517	0.672	0.754	0.801	0.829	0.844
LFR	0.917	0.731	0.592	0.495	0.424	0.371
EMNW(3)	0.765	0.940	0.987	0.982	0.954	0.917
Makeham	0.918	0.987	0.948	0.884	0.818	0.757

 Table 5: Approximate Bahadur efficiencies $\text{eff}(g)$ of $T_{n,a}$ for different tuning parameters a .

Since the limit in probability of $\tilde{T}_{n,a}/\sqrt{n}$ is $\sqrt{b_{T_a}(\vartheta)}$, the approximate local Bahadur slope is given by

$$c_{T_a}^*(\vartheta) = (\lambda_1(a))^{-1} b_{T_a}''(0) \vartheta^2 + o(\vartheta^2), \quad \text{as } \vartheta \rightarrow 0.$$

We compare the approximate Bahadur slope to the double Kullback-Leibler distance, also called Kullback-Leibler information numbers see Nikitin & Tchirina (1996),

$$KL(g) = \int_0^\infty \frac{(g'(x;0))^2}{g(x;0)} dx - \left(\int_0^\infty G'(x;0) dx \right)^2,$$

where $G(x; \vartheta) = \int_0^x g(t, \vartheta) dt$ is the cumulative distribution function of $g(\cdot; \vartheta)$. It is well known that the Kullback-Leibler information numbers are an upper bound for Bahadur efficiencies, see Bahadur (1971); Raghavachari (1970). Hence we compute the approximate Bahadur efficiencies, given by

$$\text{eff}(g) = \frac{b_{T_a}''(0)}{2\lambda_1(a)KL(g)}.$$

These are equivalent to the comparison of the local approximate Bahadur slopes of $T_{n,a}$ and the likelihood ratio test as in Cuparić et al. (2021). Note that the largest eigenvalues $\lambda_1(a)$ are given in (9) and for easy reference tabulated in the first row of Table 2. To simplify the comparison to many well known competing procedures treated in Cuparić et al. (2021), we consider the following examples of distributions, all being a member of the class \mathcal{G} and being standard references for the computation of Bahadur efficiencies of exponentiality tests, see Section 5 of Nikitin (1996):

1. the Weibull distribution with density

$$g(x; \vartheta) = (1 + \vartheta)x^\vartheta \exp(-x^{1+\vartheta}), \quad x \geq 0,$$

and $KL(g) = 1 - 2\gamma + \pi^2/6 + \gamma^2 - (1 - \gamma)^2$, where $\gamma = 0.5772156649\dots$ is the Euler-Mascheroni constant,

2. the gamma distribution with density

$$g(x; \vartheta) = x^\vartheta \exp(-x) / \Gamma(\vartheta + 1), \quad x \geq 0,$$

where $\Gamma(\cdot)$ denotes the gamma function, and $KL(g) = \pi^2/6 - 1$,

3. a linear failure rate (LFR) distribution with density

$$g(x; \vartheta) = (1 + \vartheta x) \exp(-x - \vartheta x^2/2), \quad x \geq 0,$$

and $KL(g) = 1$,

4. a mixture of exponential distributions with negative weights (EMNW(β)) with density

$$g(x; \vartheta) = (1 + \vartheta) \exp(-x) - \vartheta \beta \exp(-\beta x), \quad x \geq 0,$$

and $KL(g) = 16/45$ for $\beta = 3$,

5. and a Makeham distribution with density

$$g(x; \vartheta) = (1 + \vartheta(1 - \exp(-x))) \exp(-x - \vartheta(x - 1 + \exp(-x))), \quad x \geq 0,$$

and $KL(g) = 1/12$.

The results are reported in Table 5. Interestingly, there is a clear dependence of the efficiency of the tests on the tuning parameter a under all considered alternatives. Sometimes the highest efficiency is attained for the largest considered tuning parameter as in the gamma case, but for the LFR alternative for the lowest value of a . The Makeham alternative suggests to take 1 as best value for the tuning parameter. This behaviour is consistent with the empirical power study results in Tables 1 to 5 in Baringhaus & Henze (2008). A comparison to the Bahadur efficiencies stated in Tables 3 and 4 in Cuparić et al. (2021) for other tests of exponentiality shows that the considered procedures are competitive especially for the LFR and EMNW(3) alternatives. The results in Table 5 confirm the suggestion of Baringhaus & Henze (2008) that a tuning parameter a between 1 and 2 is a good choice, since it shows a robust approximate Bahadur efficiency over all considered alternatives.

5 Conclusions and Outlook

We revisited the family of tests of exponentiality of Baringhaus & Henze (2008) and provided new insight into the asymptotic behaviour of the tests under fixed alternatives as well as local Bahadur efficiencies. These results facilitate the comparison of the performances to other well known tests of exponentiality. As a result we visualised the dependence of the power of the tests on the choice of the tuning parameter a . This effect might be controlled by implementing a data dependent choice of the tuning parameter due to Tenreiro (2019). We leave this investigation open for further research.

We finish the article by pointing out other related open questions. As stated in Remark 1.1, the family of tests based on the integrated distribution function presented in Klar (2001) is very close to the test statistic $T_{n,a}$. For this family of tests corresponding theoretical results as in Section 3 and 4 are missing. There is little hope to solve the eigenvalue problem in this case, but results regarding Bahadur efficiency may be obtained by approximating the largest eigenvalue numerically. Another open question is due to the conservative behaviour of the confidence interval $I_{n,a}$ in Subsection 3.1, which suggests that an improvement in view of the length of the confidence interval might be found. We leave this investigation open for further research.

In Smuts et al. (2019) the authors propose a test of exponentiality based on a characterisation of the exponential law by a conditional second moment equation involving the hazard rate. Theoretical insight into this family of tests are hitherto missing, so it would be interesting to see corresponding results, since the tests based on the mean residual life function are tests based on a characterisation by a first conditional moment equation.

Acknowledgement

The author thanks Bernhard Klar and Bojana Milošević for fruitful discussions, and is grateful to Norbert Henze for numerous suggestions that all led to an improvement of the paper.

References

- Allison, J. S., Santana, L., Smit, N., & Visagie, I. J. H. (2017). An ‘apples to apples’ comparison of various tests for exponentiality. *Computational Statistics*, 32(4), 1241–1283.
- Anderson, T. W. & Darling, D. A. (1952). Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics*, 23(2), 193–212.
- Ascher, S. (1990). A survey of tests for exponentiality. *Communications in Statistics - Theory and Methods*, 19(5), 1811–1825.
- Bahadur, R. R. (1971). *Some limit theorems in statistics*, volume 4 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.
- Baringhaus, L., Ebner, B., & Henze, N. (2017). The limit distribution of weighted L^2 -goodness-of-fit statistics under fixed alternatives, with applications. *Annals of the Institute of Statistical Mathematics*, 69(5), 969–995.
- Baringhaus, L. & Henze, N. (2000). Tests of fit for exponentiality based on a characterization via the mean residual life function. *Statistical Papers*, 41(2), 225–236.
- Baringhaus, L. & Henze, N. (2008). A new weighted integral goodness-of-fit statistic for exponentiality. *Statistics & Probability Letters*, 78(8), 1006–1016.
- Becker, M. & Klöbner, S. (2017). *PearsonDS: Pearson Distribution System*. R package version 1.1.
- Božin, V., Milošević, B., Nikitin, Y. Y., & Obradović, M. (2020). New characterization-based symmetry tests. *Bulletin of the Malaysian Mathematical Sciences Society*, 43(1), 297–320.
- Cuparić, M., Milošević, B., Nikitin, Y. Y., & Obradović, M. (2020). Some consistent exponentiality tests based on Puri-Rubin and Desu characterizations. *Applications of Mathematics*, 65(3), 245–255.
- Cuparić, M., Milošević, B., & Obradović, M. (2019). New L^2 -type exponentiality tests. *SORT. Statistics and Operations Research Transactions*, 43(1), 25–50.
- Cuparić, M., Milošević, B., & Obradović, M. (2021). New consistent exponentiality tests based on V-empirical Laplace transforms with comparison of efficiencies. *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas*, 116(1), 42.
- Deheuvels, P. & Martynov, G. (2003). Karhunen-Loève expansions for weighted Wiener processes and Brownian bridges via Bessel functions. In J. Hoffmann-Jørgensen, J. A. Wellner, & M. B. Marcus (Eds.), *High Dimensional Probability III* (pp. 57–93). Basel: Birkhäuser Basel.

- Ebner, B. & Henze, N. (2021). Bahadur efficiencies of the Epps–Pulley test for normality. *Rossiiskaya Akademiya Nauk. Sankt-Peterburgskoe Otdelenie. Matematicheskii Institut im. V. A. Steklova. Zapiski Nauchnykh Seminarov (POMI)*, 30, 302–314.
- Henze, N. (1990). An approximation to the limit distribution of the Epps-Pulley test statistic for normality. *Metrika*, 37(1), 7–18.
- Henze, N. & Meintanis, S. G. (2005). Recent and classical tests for exponentiality: a partial review with comparisons. *Metrika*, 61(1), 29–45.
- Jiménez-Gamero, M. D., Milošević, B., & Obradović, M. (2020). Exponentiality tests based on Basu characterization. *Statistics*, 54(4), 714–736.
- Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions, volume 1*. John Wiley & Sons.
- Jovanović, M., Milošević, B., Nikitin, Y. Y., Obradović, M., & Volkova, K. Y. (2015). Tests of exponentiality based on Arnold-Villasenor characterization and their efficiencies. *Computational Statistics & Data Analysis*, 90, 100–113.
- Kac, M. & Siegert, A. J. F. (1947). An explicit representation of a stationary Gaussian process. *Annals of Mathematical Statistics*, 18(3), 438–442.
- Klar, B. (2001). Goodness-of-fit tests for the exponential and the normal distribution based on the integrated distribution function. *Annals of the Institute of Statistical Mathematics*, 53(2), 338–353.
- Milošević, B. (2016). Asymptotic efficiency of new exponentiality tests based on a characterization. *Metrika*, 79(2), 221–236.
- Milošević, B. & Obradović, M. (2016a). New class of exponentiality tests based on U-empirical Laplace transform. *Statistical Papers*, 57(4), 977–990.
- Milošević, B. & Obradović, M. (2016b). Some characterization based exponentiality tests and their Bahadur efficiencies. *Publications de l'Institut Mathématique. Nouvelle Série*, 100, 107–117.
- Nikitin, Y. (1995). *Asymptotic efficiency of nonparametric tests*. Cambridge: Cambridge University Press.
- Nikitin, Y. Y. (1996). Bahadur efficiency of a test of exponentiality based on a loss-of-memory type functional equation. *Journal of Nonparametric Statistics*, 6(1), 13–26.
- Nikitin, Y. Y. & Peaucelle, I. (2004). Efficiency and local optimality of nonparametric tests based on U - and V -statistics. *METRON*, 62(2), 185–200.
- Nikitin, Y. Y. & 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.
- Olver, F. W. J., Lozier, D. W., Bosvert, R. F., & Clark, C. W., Eds. (2010). *NIST handbook of mathematical functions*. Cambridge: Cambridge Univ. Press, 1st edition.
- Ossai, E. O., Madukaife, M. S., & Oladugba, A. V. (2022). A review of tests for exponentiality with Monte Carlo comparisons. *Journal of Applied Statistics*, 49(5), 1277–1304.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raghavachari, M. (1970). On a Theorem of Bahadur on the Rate of Convergence of Test Statistics. *The Annals of Mathematical Statistics*, 41(5), 1695–1699.
- Shanbhag, D. N. (1970). The characterizations for exponential and geometric distributions. *Journal of the American Statistical Association*, 65(331), 1256–1259.
- Shorack, G. R. & Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley series in probability and mathematical statistics. New York: Wiley.
- Smuts, M., Allison, J. S., & Santana, L. (2019). New goodness-of-fit tests for exponentiality based on a conditional moment characterisation. *Orion*, 35(2), 145–160.
- Spurrer, J. D. (1984). An overview of tests for exponentiality. *Communications in Statistics - Theory and Methods*, 13(13), 1635–1654.
- Stephens, M. A. (1976). Asymptotic results for goodness-of-fit statistics with unknown parameters. *The Annals of Statistics*, 4(2), 357–369.
- Tenreiro, C. (2019). On the automatic selection of the tuning parameter appearing in certain families of goodness-of-fit tests. *Journal of Statistical Computation and Simulation*, 89(10), 1780–1797.

- Villaseñor, J. A. & González-Estrada, E. (2020). On testing exponentiality based on a new estimator for the scale parameter. *Brazilian Journal of Probability and Statistics*, 34(4), 809–820.
- Volkova, K. Y. & Nikitin, Y. Y. (2015). Exponentiality tests based on Ahsanullah’s characterization and their efficiency. *Journal of Mathematical Sciences*, 204(1), 42–54.
- Watson, G. N. (1995). *A treatise on the theory of Bessel functions*. Cambridge Mathematical Library. Cambridge: Cambridge Univ. Press, 2nd edition.
- Xiong, P., Zhuang, W., & Qiu, G. (2022). Testing exponentiality based on the entropy of record values. *Journal of Applied Statistics*, 49(4), 782–802.
- Zolotarev, V. M. (1961). Concerning a certain probability problem. *Theory of Probability & Its Applications*, 6(2), 201–204.

A formulas for the functions in the estimator of the limiting variance

In this section we provide explicit formulas needed in Section 3. We write for $x, y > 0$ and $a > -1$

$$\begin{aligned} v_{1,a}(x, y) &= \int_0^\infty (y - t - 1) \mathbf{1}\{x \wedge y > t\} \exp(-at) dt, \\ v_{2,a}(x, y) &= \int_0^\infty t(y - t - 1) \mathbf{1}\{x \wedge y > t\} \exp(-at) dt, \quad \text{and} \\ v_{3,a}(x, y) &= \int_0^\infty (y - t - 1)(x - t - 1) \mathbf{1}\{x \wedge y > t\} \exp(-at) dt. \end{aligned}$$

Straightforward integration and some calculations show for $a \neq 0$

$$v_{1,a}(x, y) = a^{-2} [(a(x \wedge y) + 1 + (1 - y)a) \exp(-a(x \wedge y)) - 1 + (y - 1)a]$$

and $v_{1,0}(x, y) = -(x \wedge y)(-2y + (x \wedge y) + 2)/2$, $x, y > 0$, as well as

$$v_{2,a}(x, y) = a^{-3} \left[((x \wedge y)^2 a^2 + ((1 - y)a^2 + 2a)(x \wedge y) + 2 + (1 - y)a) \exp(-a(x \wedge y)) - 2 + (y - 1)a \right]$$

and $v_{2,0}(x, y) = -(x \wedge y)^2(2(x \wedge y) - 3y + 3)/6$, $x, y > 0$. The third function integrates to

$$\begin{aligned} v_{3,a}(x, y) &= a^{-3} \left[(-a^2(x \wedge y)^2 + a(-2 + (x + y - 2)a)(x \wedge y) - 2 - (y - 1)(x - 1)a^2 \right. \\ &\quad \left. + (x + y - 2)a) \exp(-a(x \wedge y)) + 2 + (y - 1)(x - 1)a^2 + (-x - y + 2)a \right], \end{aligned}$$

where $v_{3,0}(x, y) = (x \wedge y) \left((x \wedge y)^2/3 + (-x/2 - y/2 + 1)(x \wedge y) + (y - 1)(x - 1) \right)$.

B R source code

In the following we provide the source code written for the statistical computing language R, see R Core Team (2021). An efficient implementation of the test statistic in (5) is given by the following code.

```
T.n.a<-function(data,a)
{
  n=length(data)
  data=data/mean(data)
  datam=matrix(data,n,n)
  pjl=datam+t(datam)
  mjl=(data-1)%*(t(data)-1)
  minjl=pmin(datam,t(datam))
  if (a==0) {SUM=minjl^3/3-(pjl-2)*minjl^2/2+mjl*minjl} else {
  SUM=(-(datam-minjl-1)*(t(datam)-minjl-1)*exp(-a*minjl)+mjl)/a+((pjl
  -2*minjl-2)*exp(-a*minjl)-(pjl-2))/a^2+2*(1-exp(-a*minjl))/a^3}
  return(sum(SUM)/n)
}
```

The following code can be used for approximation of the critical values of the test statistic by a Pearson system of distributions using the R package PearsonDS, see Becker & Klößner (2017).

```
#Cumulants from Remark 2.2
kappa_1<-function(a) {return(1/((a+1)*(a+2)))}
kappa_2<-function(a) {return(2/((a^2+3*a+2)*(2*a+3)*(a+2)))}
kappa_3<-function(a) {return(16/((a+1)*(a+2)^3*(2*a+3)*(3*a+4)))}
kappa_4<-function(a) {return(48*(11*a+16)/((a+1)*(a+2)^4*(2*a+3)^2*(3*a+4)*(4*a+5)))}

#The function provides the approximation of the 1-alpha quantile of the limiting
#distribution for a tuning parameter a
cv.T<-function(alpha,a)
{
  require(PearsonDS)
  kum=c(kappa_1(a),kappa_2(a),kappa_3(a),kappa_4(a))
  mom.a=c(kum[1:2],kum[3]*kum[2]^(-3/2),3+kum[4]*kum[2]^(-2))
  return(qpearson(1-alpha,moments=mom.a))
}
```