Volume 35 (2), pp. 145–160
http://orion.journals.ac.za

ORiON
ISSN 0259–191X (print)
ISSN 2224–0004 (online)

New goodness-of-fit tests for exponentiality based on a conditional moment characterisation

M Smuts* JS Allison† L Santana‡

Received: 23 August 2019; Accepted: 30 September 2019

Abstract

The exponential distribution plays a key role in the practical application of reliability theory, survival analysis, engineering and queuing theory. These applications often rely on the underlying assumption that the observed data originate from an exponential distribution. In this paper, two new tests for exponentiality are proposed, which are based on a conditional second moment characterisation. The proposed tests are compared to various established tests for exponentiality by means of a simulation study where it is found that the new tests perform favourably relative to the existing tests. The tests are also applied to real-world data sets with independent and identically distributed data as well as to simulated data from a Cox proportional hazards model, to determine whether the residuals obtained from the fitted model follow a standard exponential distribution.

Key words: Characterisation, Cox's proportional hazards model, exponential distribution, goodness-of-fit test.

1 Introduction

The exponential distribution is an important and commonly used statistical model for a multitude of real-life phenomena, such as lifetimes, time to default of loans, and many other *time-to-event* scenarios. As a result, this distribution plays a vital role in the practical application of reliability theory, survival analysis, engineering, and queuing theory

*Subject group Statistics, Department of Statistics, North-West University, Potchefstroom, South Africa email: smuts.marius@nwu.ac.za

†Subject group Statistics, Department of Statistics, North-West University, Potchefstroom, South Africa email: james.allison@nwu.ac.za

‡Corresponding author: Subject group Statistics, Department of Statistics, North-West University, Potchefstroom, South Africa email: leonard.santana@nwu.ac.za

http://dx.doi.org/10.5784/35-2-661

(to name only a few), as the underlying theory governing these applications often assumes an exponential distribution for the data. Therefore, to effectively implement these applications, it is necessary to perform goodness-of-fit tests to determine whether these fundamental distributional assumptions are satisfied or not. Examples where the assumption of exponentiality is necessary (and hence the need to test for this assumption) range from the analysis of queuing networks [26], to cancer clinical trials [15], and the time-to-failure of systems of machines and operators [12]; for further examples of data sets see the papers by Shanker, Fesshaye, and Selvaraj [23, 24].

Suppose that a random variable X follows an exponential distribution with scale parameter λ (written $X \sim Exp(\lambda)$). This random variable has a number of unique distributional properties which include the forms of its cumulative distribution function (CDF), survival function, probability density function, and characteristic function (CF), which are given by

$$F(x) = \mathrm{P}(X < x) = 1 - e^{-\lambda x},$$
$$S(x) = \mathrm{P}(X > x) = 1 - F(x) = e^{-\lambda x},$$
$$f(x) = \lambda e^{-\lambda x},$$

and

$$\phi(x) = \frac{\lambda}{\lambda - ix}, \quad i = \sqrt{-1},$$

respectively, with $x > 0$ and where $\lambda > 0$ is the scale parameter, with $\mathrm{E}(X) = 1/\lambda$. In addition, the exponential distribution also exhibits many other unique distributional properties, called *characterisations*. These characterisations help in the development of tests for exponentiality since, if one can verify that the data has these properties, then one can conclude that the data were obtained from an exponential distribution. One such property is the so-called 'memoryless' property which states that, if X follows an exponential distribution, then we can write

$$\mathrm{P}(X > s + t \,|\, X > s) = \mathrm{P}(X > t), \tag{1}$$

for $s, t > 0$. This property implies that, if X represents the lifetime of a certain component, then the remaining lifetime of that component is independent of its current age. For components that suffer from wear-and-tear (*i.e.*, where the lifetime is dependent on its current age), the exponential distribution would not be an appropriate model. A second property states that the exponential distribution uniquely has the feature that the *hazard rate* is a constant, that is,

$$h(x) = \frac{f(x)}{1 - F(x)} = \lambda.$$

This feature is directly tied to the memoryless property since the failure rate is constant throughout the lifetime of the component.

Suppose now that $X_1, X_2, \ldots, X_n$ are realisations from some random variable X with unknown distribution function F, then the process of testing whether or not this data are realisations from an exponential distribution with parameter λ involves the use of statistical inference via goodness-of-fit tests. The inferential question can be framed in the form of the following composite hypothesis statement:

$$H_0 : \text{The distribution of } X \text{ is } \exp(\lambda), \tag{2}$$

for some $\lambda > 0$, against the alternative hypothesis that the distribution of X is something other than exponential. Note that the tests that will be discussed in this paper all make use of a scaled version of the original data, defined as $Y_j = X_j\widehat{\lambda}$, $j = 1, 2, \ldots, n$, where $\widehat{\lambda}$ denotes the maximum likelihood estimator (MLE) of the parameter λ and is given by $\widehat{\lambda} = 1/\bar{X}_n$ with $\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$. The motivation for the use of this scaling factor primarily comes from the fact that the distribution of exponential random variables is invariant to simple scale transformations, that is, X is exponentially distributed if, and only if, cX is also exponentially distributed for every constant $c > 0$. Therefore, conclusions drawn regarding exponentiality based on the sample $Y_1, Y_2, \ldots, Y_n$ can reasonably be extended to the exponentiality of X (from which $X_1, X_2, \ldots, X_n$ was obtained). Furthermore, many statistics discussed will also employ the order statistics of X_j and Y_j, defined as $X_{(1)}<X_{(2)}<\ldots<X_{(n)}$ and $Y_{(1)}<Y_{(2)}<\ldots<Y_{(n)}$, respectively.

To test the hypothesis in (2), formal test statistics are used, some of which are general statistics that can be applied to test for almost any distribution, whereas others exploit the various unique characteristics of exponentially distributed data, such as the memoryless or constant hazard properties. Examples of general tests employed to test exponentiality include the Kolmogorov-Smirnov test and the Cramér-von Mises test, both of which are based on the same basic principle of measuring the discrepancies between the theoretical CDF of the exponential distribution and its empirical equivalent (see, *e.g.*, Chapter 4 of D'Agostino and Stephens, 1986 [11]). Another class of general tests involves a similar approach, but replaces the CDF with the CF; examples of these include the test by Epps and Pulley [13], and the one by Meintanis, Swanepoel and Allison [21]. In addition, there are many goodness-of-fit tests based on the unique properties of the exponential distribution, which are occasionally desirable as they tend to focus on much more specific aspects of the distribution and are potentially more robust than the more general tests. For example, since the memoryless property uniquely characterises the exponential distribution, this implies that the exponentially distributed random variable X will have this property and conversely, if X exhibits this property it must be exponentially distributed. Therefore, a test based on this property will involve first determining sample estimates of the two probabilities appearing on either side of the expression of (1), and the test can then be designed to measure the equality of these two estimates. For examples of tests based on this characterisation, see [2], [5], and [6].

There are many more such unique characterisations of the exponential distribution and the literature on goodness-of-fit contains numerous test statistics based on these characterisations. For example, for tests based on the mean residual life, see [8], [25], [17], and [9]. For a test based on the Arnold-Villasenor characterisation, see [18], and for a test based on the Rossberg characterisation see [27]. For a comprehensive review of tests for exponentiality, the interested reader is referred to the review papers by Ascher [7], Henze and Meintanis [16], and Allison, Santana, Smit and Visagie [4].

The remainder of the paper is organised as follows: In Section 2 we propose new tests for exponentiality which are based on a conditional second moment characterisation of the exponential distribution and, in Section 3, the results of a brief Monte Carlo simulation are presented to compare the power performance of the newly proposed tests to some commonly used existing tests for exponentiality. The paper concludes in Section

4 where the tests are applied to some real-world data sets with independent and identically distributed random values, as well as to data simulated from a Cox proportional hazards model, to determine whether the residuals obtained from the fitted model follow a standard exponential distribution.

2 New tests for exponentiality based on a characterisation

Consider the following characterisation of the exponential distribution by Afify, Nofal and Ahmed [1]:

Characterisation. *Let X be a non-negative random variable with continuous distribution function F and density f. If $E(X^2) < \infty$, then X has an exponential distribution with parameter λ (that is, $F(x) = 1 - e^{-\lambda x}$) if, and only if, for all $t > 0$*

$$E\left[X^2 | X > t\right] = \frac{2}{\lambda^2} + h(t)\left(\frac{t^2}{\lambda} + \frac{2t}{\lambda^2}\right),$$

where $h(t) = \frac{f(t)}{S(t)}$ is the hazard rate.

From this characterisation we can deduce the following corollary.

Corollary 1 *Let X be a non-negative random variable with continuous distribution function F. If $E(X^2) < \infty$, then X has an exponential distribution with parameter λ if, and only if, for all $t > 0$*

$$E\left[X^2\,\mathrm{I}(X > t)\right] = S(t) r_\lambda(t),$$

where $r_\lambda(t) := \frac{2}{\lambda^2} + h(t)\left(\frac{t^2}{\lambda} + \frac{2t}{\lambda^2}\right)$ and $\mathrm{I}(\cdot)$ denotes the indicator function.

Proof: Straightforward calculations yield that, for all $t > 0$,

$$E\left[X^2 | X > t\right] = \frac{1}{P(X > t)} E\left[X^2\,\mathrm{I}(X > t)\right] = \frac{1}{S(t)} E\left[X^2\,\mathrm{I}(X > t)\right].$$

From the Characterisation, it follows that X has an exponential distribution with parameter λ if, and only if, for all $t > 0$,

$$\frac{1}{S(t)} E\left[X^2\,\mathrm{I}(X > t)\right] = r_\lambda(t),$$

or equivalently if, and only if, for all $t > 0$,

$$E\left[X^2\,\mathrm{I}(X > t)\right] = S(t) r_\lambda(t).$$

□

Now, note that if $X \sim Exp(\lambda)$ then

$$Y = \lambda X \sim \exp(1).$$

Based on Y, the characterisation in Corollary 1 can be restated as follows: Y is exponentially distributed if, and only if,

$$E\left[Y^2\,\mathrm{I}(Y>t)\right]=S(t)r_1(t),$$

where $S(t)=P(Y>t)$ and $r_1(t)=2+h(t)\left(t^2+2t\right)$, with $h(t)$ the hazard rate of Y.

Based on this, a random variable Y has a standard exponential distribution if, and only if,

$$E\left[Y^2\,\mathrm{I}(Y>t)\right]-S(t)r_1(t)=0,$$

or equivalently if, and only if,

$$\psi(t):=E\left[Y^2\,\mathrm{I}(Y>t)\right]-2S(t)-f(t)\left\{t^2+2t\right\}=0,$$

where $f(t)$ is the density function of Y.

Naturally, $\psi(t)$ is unknown and hence must be estimated from the data $Y_1,Y_2,\ldots,Y_n$, where $Y_j=X_j\widehat{\lambda}=X_j/\bar{X}_n$, $j=1,2,\ldots,n$. Define two possible estimators for $\psi(t)$ by

$$\widehat{\psi}_n^{(1)}(t)=\frac{1}{n}\sum_{i=1}^{n}Y_i^2\,\mathrm{I}\left(Y_i>t\right)-\frac{2}{n}\sum_{i=1}^{n}\mathrm{I}\left(Y_i>t\right)-e^{-t}\left(t^2+2t\right)$$

and

$$\widehat{\psi}_n^{(2)}(t)=\frac{1}{n}\sum_{i=1}^{n}Y_i^2\,\mathrm{I}\left(Y_i>t\right)-\frac{2}{n}\sum_{i=1}^{n}\mathrm{I}\left(Y_i>t\right)-\widehat{f}(t)\left(t^2+2t\right).$$

Here, $\widehat{f}(t)$ denotes the kernel density estimate of $f(t)$, which is defined as

$$\widehat{f}(t)=\frac{1}{nh}\sum_{i=1}^{n}\phi\left(\frac{t-Y_i}{h}\right),$$

where $\phi(\cdot)$ is the standard normal density function and h is a suitably chosen bandwidth (for an in-depth discussion on kernel density estimators, the interested reader is referred to the monograph by Wand and Jones [28]). The only difference between the estimators $\widehat{\psi}_n^{(1)}$ and $\widehat{\psi}_n^{(2)}$ is that in $\widehat{\psi}_n^{(1)}$ we choose $f(t)=e^{-t}$, the density function specified under the null hypothesis, whilst in $\widehat{\psi}_n^{(2)}$ we estimate f by $\widehat{f}$.

Now, if the observed data originated from an exponential distribution, then both $\widehat{\psi}_n^{(1)}$ and $\widehat{\psi}_n^{(2)}$ should be close to zero. This leads to the following two Cramér-von Mises type test statistics

$$S_n=n\int_0^\infty\left[\widehat{\psi}_n^{(1)}(t)\right]^2w(t)dF_n(t)$$

and

$$T_n=n\int_0^\infty\left[\widehat{\psi}_n^{(2)}(t)\right]^2w(t)dF_n(t),$$

where $F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{I}(Y_i \le t)$ is the empirical distribution function of $Y_1, Y_2, \ldots, Y_n$ and $w(t)$ is a suitable, positive weight function satisfying some standard integrability conditions. For implementation of the proposed test statistics, we will use $w(t) = e^{-at}$, where $a > 0$ is a user-defined tuning parameter.

With this choice of $w(t)$ the following easily calculable form of the proposed test statistics S_n and T_n is obtained.

$$S_{n,a} = \sum_{j=1}^{n} \left[\frac{1}{n} \sum_{i=1}^{n} Y_i^2 \,\mathrm{I}\left(Y_i > Y_j\right) - \frac{2}{n} \sum_{i=1}^{n} \mathrm{I}\left(Y_i > Y_j\right) - e^{-Y_j}\left(Y_j^2 + 2Y_j\right) \right]^2 e^{-aY_j}$$

and

$$T_{n,a} = \sum_{j=1}^{n} \left[\frac{1}{n} \sum_{i=1}^{n} Y_i^2 \,\mathrm{I}\left(Y_i > Y_j\right) - \frac{2}{n} \sum_{i=1}^{n} \mathrm{I}\left(Y_i > Y_j\right) - \widehat{f}(Y_j)\left(Y_j^2 + 2Y_j\right) \right]^2 e^{-aY_j}.$$

Both tests reject the null hypothesis in (2) for large values of the test statistics. The critical values for the test statistics can easily be calculated using the following Monte Carlo procedure.

1. Draw a random sample $X_1, X_2, \ldots, X_n$ from an exponential distribution with parameter 1.
2. Obtain the scaled observations $Y_i = X_i/\bar{X}_n, i = 1, 2, \ldots, n$.
3. Calculate the test statistic, say $S = S_n(Y_1, Y_2, \ldots, Y_n)$.
4. Repeat steps 1–3 a large number of times, say MC times, to obtain MC copies of S denoted $S_1, S_2, \ldots, S_{MC}$.
5. Obtain the order statistics $S_{(1)} \le S_{(2)} \le \cdots \le S_{(MC)}$.
6. The critical value at a α% significance level is then given by

$$\widehat{C}_n(\alpha) = S_{(\lfloor MC(1-\alpha)\rfloor)},$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x.

3 Simulation study and results

In this section Monte Carlo simulations are used to compare the finite-sample performance of the newly proposed tests $T_{n,a}$ and $S_{n,a}$ to the following existing tests for exponentiality.

- The traditional tests of Kolmogorov-Smirnov (KS_n) and Cramér-von Mises (CM_n), where the test statistics for these tests are given by

$$KS_n = \max\left\{KS_n^+, KS_n^-\right\},$$

where

$$KS_n^+ = \max_{1\le j\le n}\left[\frac{j}{n} - \left(1 - e^{-Y_{(j)}}\right)\right],$$

$$KS_n^- = \max_{1\le j\le n}\left[\left(1-e^{-Y_{(j)}}\right)-\frac{j-1}{n}\right]$$

and

$$CM_n = \frac{1}{12n}+\sum_{j=1}^{n}\left[\left(1-e^{-Y_{(j)}}\right)-\frac{2j-1}{2n}\right]^2.$$

Both of these tests reject the null hypothesis for large values of the test statistics

- A Kolmogorov-Smirnov type test ($\overline{KS}_n$) and a Cramér-von Mises type test ($\overline{CM}_n$) based on the mean residual life, as developed by Baringhaus and Henze [8], with the following test statistics

$$\overline{KS}_n = \sqrt{n}\sup_{t\ge 0}\left|\frac{1}{n}\sum_{j=1}^{n}\min\{Y_j,t\}-\frac{1}{n}\sum_{j=1}^{n}I\left(Y_j\le t\right)\right| = \sqrt{n}\max\left\{\overline{KS}_n^+,\overline{KS}_n^-\right\},$$

where

$$\overline{KS}_n^+ = \max_{j\in\{0,1,\dots,n-1\}}\left[\frac{1}{n}\left(Y_{(1)}+\cdots+Y_{(j)}\right)+Y_{(j+1)}\left(1-\frac{j}{n}\right)-\frac{j}{n}\right],$$
$$\overline{KS}_n^- = \max_{j\in\{0,1,\dots,n-1\}}\left[\frac{j}{n}-\frac{1}{n}\left(Y_{(1)}+\cdots+Y_{(j)}\right)-Y_{(j)}\left(1-\frac{j}{n}\right)\right].$$

and

$$\overline{CM}_n = n\int_0^\infty\left[\frac{1}{n}\sum_{j=1}^{n}\min\left\{Y_j,t\right\}-\frac{1}{n}\sum_{j=1}^{n}I\left(Y_j\le t\right)\right]^2 e^{-t}dt$$
$$=\frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\left[2-3\exp\left(-\min\{Y_j,Y_k\}\right)-2\min\{Y_j,Y_k\}\left(e^{-Y_j}+e^{-Y_k}\right)+2\exp\left(-\max\{Y_j,Y_k\}\right)\right].$$

Both $\overline{KS}_n$ and $\overline{CM}_n$ reject the null hypothesis for large values.

- The Epps and Pulley test EP_n [13], which is based on the characteristic function, $\phi(x)$, and with the test statistic given by

$$EP_n = \sqrt{48n}\left[\frac{1}{n}\sum_{j=1}^{n}e^{-Y_j}-\frac{1}{2}\right].$$

The null hypothesis is rejected for large values of $|EP_n|$.

3.1 Simulation setting

A significance level of 5% was used throughout the study. Empirical critical values of all the tests were obtained from 10 000 independent Monte Carlo replications using the procedure given at the end of Section 2. Power estimates were calculated for sample sizes $n = 20$ and $n = 30$ using 10 000 independent Monte Carlo replications for the various alternative distributions given in Table 1.

The two new tests in which a tuning parameter appears were evaluated for $a = 0.25$ and $a = 1$. All calculations and simulations were performed in R [22].

3.2 Simulation results

Table 2 contains the percentage of the 10 000 Monte Carlo samples that resulted in the rejection of the null hypothesis in (2) rounded to the nearest integer. For each alternative the top row corresponds to the estimated powers obtained for $n = 20$ whereas the row below corresponds to the estimated powers for $n = 30$.

Alternative	$f(x)$	Notation
Gamma	$\frac{1}{\Gamma(\theta)}x^{\theta-1}e^{-x}$	$\Gamma(\theta)$
Weibull	$\theta x^{\theta-1}\exp(-x^{\theta})$	$W(\theta)$
Power	$\frac{1}{\theta}x^{(1-\theta)/\theta}, \quad 0 < x < 1$	$PW(\theta)$
Linear failure rate	$(1+\theta x)\exp(-x-\theta x^2/2)$	$LFR(\theta)$
Exponential logarithmic	$\frac{(\ln\theta)^{-1}(1-\theta)e^{-x}}{(1-\theta)e^{-x}-1}$	$EL(\theta)$
Exponential geometric	$\frac{(1-\theta)e^{-x}}{(1-\theta e^{-x})^2}$	$EG(\theta)$

Table 1: *Alternative distributions considered in the simulation study.*

The highest power for each alternative distribution is highlighted for ease of comparison. From Table 2 it is clear that there is no single test that dominates all of the other tests. However, $S_{n,a}$ outperforms all its competitors for the $EG(\theta)$, $EL(\theta)$ and $\Gamma(0.7)$ alternatives and for both sample sizes. No single test dominates for the majority of the other alternatives with the exception of $T_{n,a}$, which performs well for the alternatives $LF(4)$ and $PW(1)$. Overall, the two newly proposed tests produce estimated powers which are competitive relative to the other tests and, hence, this limited Monte Carlo study shows that they can be used in practice to test whether observed data are realised from an exponential distribution.

4 Practical applications and conclusion

In this section all the tests considered in the simulation study will be applied to both real-world and simulated data sets. The two real-world data sets considered in this study respectively contain the failure times of air conditioning systems and the waiting times of bank customers. For these two data sets, the tests for exponentiality will be used to determine whether the observed values are realisations from an exponential distribution. On the other hand, the remaining two data sets that will be considered are simulated from a Cox proportional hazards (CPH) model and the tests for exponentiality will be used to determine the adequacy of a specific CPH model fitted to the data.

4.1 Practical application to real-world data sets

The first data set contains 30 failure times of the air conditioning system of an airplane as given by Linhart and Zucchini [20], whereas the second data set contains the waiting times

Distribution	$T_{n,0.25}$	$T_{n,1}$	$S_{n,0.25}$	$S_{n,1}$	KS_n	CM_n	$\overline{KS}_n$	$\overline{CM}_n$	EP_n
$EG(0.2)$	5	7	7	**8**	6	6	5	6	6
	6	7	7	**8**	6	6	5	6	6
$EG(0.5)$	10	15	17	**19**	11	12	9	14	15
	16	20	22	**23**	15	16	13	19	19
$EG(0.8)$	37	44	47	**51**	34	39	32	43	45
	53	61	61	**65**	49	55	48	60	60
$EL(0.2)$	14	20	22	**25**	16	18	12	19	20
	23	29	29	**33**	23	26	19	28	29
$EL(0.5)$	6	8	9	**10**	7	7	6	8	8
	8	10	11	**12**	8	8	7	9	9
$EL(0.8)$	5	**6**	**6**	**6**	5	5	5	5	6
	5	5	**6**	**6**	**6**	5	5	6	5
$\Gamma(0.7)$	13	17	19	**20**	15	18	11	18	19
	17	24	24	**27**	22	26	18	26	26
$\Gamma(1)$	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**
	5	**5**	**5**	**5**	**5**	**5**	**5**	**5**	**5**
$\Gamma(1.5)$	18	15	12	5	17	21	**22**	20	21
	24	19	16	12	25	**28**	27	27	**28**
$\Gamma(2)$	42	40	33	21	41	**48**	46	47	**48**
	59	56	48	42	60	**69**	64	**69**	**69**
$LFR(2)$	30	27	18	14	24	28	**32**	29	30
	44	39	35	24	35	43	46	**47**	46
$LFR(4)$	**45**	40	30	23	34	42	**45**	44	42
	63	57	53	41	51	60	61	62	**63**
$PW(1)$	**84**	73	65	49	54	68	74	71	68
	97	89	87	76	72	86	90	90	86
$W(1.2)$	13	11	8	6	12	14	**16**	13	14
	18	14	11	7	16	18	**20**	19	19
$W(1.5)$	45	40	35	29	40	48	**55**	49	51
	66	59	55	43	58	66	65	70	**72**

Table 2: *Estimated powers for $n = 20$ (top) and $n = 30$ (bottom).*

(in minutes) of 100 bank customers before service as obtained from Ghitany, Atieh and Nadarajah [14]; both data sets can be found in the Appendix in Tables 7 and 8. In Table 3 and 4 a summary of the results of all the different tests for exponentiality can be found. The summary contains the value of each test statistic and associated p-value used to test whether the data originated from an exponential distribution. For the failure time data, all of the tests, except the KS_n test, *do no reject* the null hypothesis of exponentiality using a 5% significance level. In contrast, for the waiting time data, all of the tests *reject* the null hypothesis of exponentiality at the same significance level. This illustrates that the newly proposed test at least agrees with the more traditional tests for exponentiality.

Test	**Test statistic value**	***p*-value**
$T_{n,0.25}$	13.592	0.153
$T_{n,1}$	8.446	0.089
$S_{n,0.25}$	8.336	0.117
$S_{n,1}$	6.032	0.079
KS_n	0.213	0.020
CM_n	0.214	0.053
$\overline{KS}_n$	1.325	0.053
$\overline{CM}_n$	0.397	0.072
EP_n	1.716	0.089

Table 3: *Summary of results for failure times of air conditioning system.*

Test	**Test statistic value**	***p*-value**
$T_{n,0.25}$	57.365	0.004
$T_{n,1}$	30.407	0.005
$S_{n,0.25}$	31.810	0.010
$S_{n,1}$	19.204	0.012
KS_n	0.173	<0.001
CM_n	0.715	<0.001
$\overline{KS}_n$	2.176	<0.001
$\overline{CM}_n$	1.480	<0.001
EP_n	-3.659	0.001

Table 4: *Summary of results for waiting times of bank customers.*

4.2 Practical application to simulated data sets

The following two data sets, given in the Appendix in Tables 9 and 10, contain simulated lifetimes $(t_i, i = 1, 2, \dots, 100)$ together with a single covariate $(x_i, i = 1, 2, \dots, 100)$ which can take on the values 0, 1, 2 or 3. The first data set was obtained by simulating data from a CPH model with a Weibull cumulative baseline hazard function, whereas the second data set was simulated from a CPH model with a log-normal cumulative baseline hazard function.

Recall that the cumulative hazard function of the j^{th} individual follows a CPH model with a single covariate if

$$\Lambda_j(t) = e^{\beta x_j} H(t),$$

where $H(\cdot)$ is some unspecified baseline cumulative hazard function, x_j is the value of the covariate of the j^{th} individual and β is an unknown regression parameter.

On the basis of the observed data $(t_j, x_j), j = 1, 2, \ldots, 100$ we wish to test the null hypothesis

$$\mathcal{H}_0 : H(t) = H_0(t; a, b), \tag{3}$$

where $H_0(t; a, b) = \left(\frac{t}{b}\right)^a$ is the Weibull cumulative baseline hazard function with unknown parameters a and b.

We can now estimate the parameters β, a and b by their maximum likelihood estimators $\widehat{\beta}$, $\widehat{a}$ and $\widehat{b}$. Based on these estimators we can obtain the (so-called) Cox-Snell residuals, defined as

$$\widehat{\varepsilon}_j = e^{\widehat{\beta} x_j} H_0(t_j; \widehat{a}, \widehat{b}).$$

If the null hypothesis is true (*i.e.*, if the cumulative baseline hazard was correctly specified as the Weibull cumulative baseline hazard) then the Cox-Snell residuals should (approximately) follow a standard exponential distribution (see, *e.g.*, Chapter 11 of Klein and Moeschberger [19]). Hence any exponential test on the basis of $\widehat{\varepsilon}_j, j = 1, 2, \ldots, 100$ constitutes in effect a goodness-of-fit test for the CPH model itself.

It is, therefore, expected that tests for exponentiality will not reject the null hypothesis for the first simulated data set (recall that this data was generated from a CPH model with a Weibull cumulative baseline hazard), whereas the tests should reject the null hypothesis of exponentiality for the second simulated data set (which was generated from a CPH model with a log-normal cumulative baseline hazard).

The results of all the different tests for exponentiality for the two simulated data sets are summarised in Table 5 and 6, which display both the test statistics and associated p-values used to test whether the residuals originate from a standard exponential distribution, *i.e.*, whether the cumulative baseline hazard is correctly specified as Weibull. Due to the fact that the null hypothesis in (3) involves unknown parameters — and hence must be estimated under $\mathcal{H}_0$ — the p-values had to be obtained using the bootstrap algorithm described in Cockeran, Allison and Meintanis [10]. For the first simulated data set, the MLEs are $\widehat{\beta} = 0.090$, $\widehat{a} = 0.880$ and $\widehat{b} = 0.763$. Table 5 shows that all tests correctly do not reject the null hypothesis, which is expected, as the data was known to be generated using a Weibull cumulative baseline hazard.

The second simulated data set produced the following MLEs: $\widehat{\beta} = 0.008$, $\widehat{a} = 0.854$ and $\widehat{b} = 3.302$, and the resulting p-values displayed in Table 6 indicate that the null hypothesis was rejected by all of the tests. This is not surprising, since a good test for exponentiality should have the ability to detect the mis-specification of the Weibull cumulative baseline hazard when the data originated from a CPH model with a log-normal cumulative baseline hazard.

To ultimately use the two new tests for exponentiality, one would need to make a choice regarding the value of the tuning parameter a, however, from extensive simulation studies

Test	**Test statistic value**	***p*-value**
$T_{n,0.25}$	2.824	0.549
$T_{n,1}$	0.713	0.692
$S_{n,0.25}$	0.328	0.745
$S_{n,1}$	0.093	0.872
KS_n	0.065	0.358
CM_n	0.040	0.679
$\overline{KS}_n$	0.692	0.447
$\overline{CM}_n$	0.053	0.674
EP_n	0.156	0.337

Table 5: *Summary of results for the first simulated data set (Weibull cumulative baseline hazard).*

Test	**Test statistic value**	***p*-value**
$T_{n,0.25}$	12.476	0.025
$T_{n,1}$	5.394	0.017
$S_{n,0.25}$	8.955	0.013
$S_{n,1}$	5.183	0.013
KS_n	0.119	0.004
CM_n	0.282	0.001
$\overline{KS}_n$	1.366	0.005
$\overline{CM}_n$	0.483	0.001
EP_n	1.289	0.001

Table 6: *Summary of results for the second simulated data set (log-normal cumulative baseline hazard).*

conducted (not displayed here), it was concluded that $a = 1$ produces satisfactory results. If, however, one would prefer to rather use a data dependent choice of this parameter, one can employ the method outlined in Allison and Santana [3].

Acknowledgements

The second author's work is based on research supported by the National Research Foundation (NRF). Any opinion, finding and conclusion or recommendation expressed in this material is that of the author and the NRF does not accept any liability in this regard.

References

[1] Afify AZ, Nofal ZM & Ahmed AH, 2013, *Two characterizations of gamma distribution in terms of sth conditional moments*, Journal of Advances in Mathematics, **5**, pp. 688–695.

[2] Ahmad I & Alwasel I, 1999, *A goodness-of-fit test for exponentiality based on the memoryless property*, Journal of the Royal Statistical Society Series B, **61(3)**, pp. 681–689.

[3] Allison JS & Santana L, 2014, *On a data-dependent choice of the tuning parameter appearing in certain goodness-of-fit tests*, Journal of Statistical Computation and Simulation, **85(16)**, pp. 3276–3288.

[4] Allison JS, Santana L, Smit N & Visagie IJH, 2017, *An 'apples to apples' comparison of various tests for exponentiality*, Computational Statistics, **32(4)**, pp. 1241–1283.

[5] Alwasel I, 2001, *On goodness of fit testing for exponentiality using the memoryless property*, Journal of Nonparametric Statistics, **13**, pp. 569–581.

[6] Angus JE, 1982, *Goodness-of-fit tests for exponentiality based on a loss-of-memory type functional equation*, Journal of Statistical Planning and Inference, **6**, pp. 241–251.

[7] Ascher S, 1990, *A survey of tests for exponentiality*, Communications in Statistics – Theory and Methods, **19(5)**, pp. 1881–1825.

[8] Baringhaus L & Henze N, 2000, *Tests of fit for exponentiality based on a characterization via the mean residual life function*, Statistical Papers, **41**, pp. 225–236.

[9] Baringhaus L & Henze N, 2008, *A new weighted integral goodness-of-fit statistic for exponentiality*, Statistics and Probability Letters, **78**, pp. 1006–1016.

[10] Cockeran M, Meintanis SG, & Allison JS, 2019, *Goodness-of-fit tests in the Cox proportional hazards model*, Communications in Statistics – Simulation and Computation, pp. 1–12.

[11] D'Agostino RB, & Stephens MA, 1986, *Goodness-of-fit Techniques.* Marcel Dekker, New York.

[12] Davis DJ, 1952, *An analysis of some failure data*, Journal of the American Statistical Association, **47(258)**, pp. 113–150.

[13] Epps TW & Pulley LB, 1986, *A test of exponentiality vs. monotone-hazard alternatives derived from the empirical characteristic function*, Journal of the Royal Statistical Society Series B, **48(2)**, pp. 206–213.

[14] Ghitany ME, Atieh B & Nadarajah S, 2008, *Lindley distribution and its application*, Mathematics and computers in simulation, **78(4)**, pp. 493–506.

[15] Han G, Schell MJ, Zhang H, Zelterman D, Pusztai L, Adelson K & Hatzis C, 2016, *Testing violations of the exponential assumption in cancer clinical trials with survival endpoints*, Biometrics, **73(2)**, pp. 687–695.

[16] Henze N, & Meintanis SG, 2005, *Recent and classical tests for exponentiality: a partial review with comparisons*, Metrika, **36**, pp. 29–45.

[17] Jammalamadaka SR, & Taufer E, 2006, *Use of mean residual life in testing departures from exponentiality*, Journal of Nonparametric Statistics, **18(3)**, pp. 277–292.

[18] Jovanović M, Milošević B, Nikitin YaYu, Obradović M & Volkova KYu, 2015, *Tests of exponentiality based on Arnold-Villasenor characterization and their efficiencies*, Computational Statistics and Data Analysis, **90**, pp. 100–113.

[19] Klein JP & Moeschberger ML, 2006, *Survival analysis: techniques for censored and truncated data*, Springer Science & Business Media.

[20] Linhart H & Zucchini W, 1986, *Model selection,* John Wiley & Sons.

[21] Meintanis SG, Swanepoel JWH & Allison JS, 2014, *The probability weighted characteristic function and goodness-of-fit testing*, Journal of Statistical Planning and Inference, **146**, pp. 122–132.

[22] R Core Team, 2018, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

[23] Shanker R, Fesshaye H & Selvaraj S, 2015, *On modeling of lifetimes data using exponential and Lindley distributions*, Biometrics & Biostatistics International Journal, **2(5)**, pp. 1/9 – 9/9.

[24] Shanker R, Fesshaye H, & Selvaraj S, 2016, *On modeling of lifetime data using one parameter Akash, Lindley and exponential distributions*, Biometrics & Biostatistics International Journal, **3(2)**.

[25] Taufer E, 2000, *A new test for exponentiality against omnibus alternatives*, Stochastic Modelling and Applications, **3**, pp. 23–36.

[26] Vinod B, & Altiok T, 1986, *Approximating unreliable queueing networks under the assumption of exponentiality*, Journal of the Operational Research Society, **37(3)**, pp. 309–316.

[27] Volkova KYu, 2010, *On asymptotic efficiency of exponentiality tests based on Rossberg's characterization*, Journal of Mathematical Sciences, **167(4)**, pp. 486–494.

[28] Wand MP & Jones MC, 1994, *Kernel smoothing.* Chapman and Hall/CRC.

Appendices

A. Real-world and simulated data sets

This appendix contains the data sets that were used for the practical applications in Section 4.

23	261	87	7	120	14	62	47	225	71
246	21	42	20	5	12	120	11	3	14
71	11	14	11	16	90	1	16	52	95

Table 7: *Failure times of air conditioning system of an airplane.*

0.8	0.8	1.3	1.5	1.8	1.9	1.9	2.1	2.6	2.7
2.9	3.1	3.2	3.3	3.5	3.6	4	4.1	4.2	4.2
4.3	4.3	4.4	4.4	4.6	4.7	4.7	4.8	4.9	4.9
5	5.3	5.5	5.7	5.7	6.1	6.2	6.2	6.2	6.3
6.7	6.9	7.1	7.1	7.1	7.1	7.4	7.6	7.7	8
8.2	8.6	8.6	8.6	8.8	8.8	8.9	8.9	9.5	9.6
9.7	9.8	10.7	10.9	11	11	11.1	11.2	11.2	11.5
11.9	12.4	12.5	12.9	13	13.1	13.3	13.6	13.7	13.9
14.1	15.4	15.4	17.3	17.3	18.1	18.2	18.4	18.9	19
19.9	20.6	21.3	21.4	21.9	23	27	31.6	33.1	38.5

Table 8: *Waiting times of bank customers (in minutes) before service.*

t	0.277	0.171	0.234	0.531	0.319	1.633	0.161	0.373	0.209	0.606
x	1	0	0	1	0	0	3	1	1	2
t	0.470	2.346	0.490	0.565	2.259	0.137	0.502	0.066	0.212	0.448
x	2	0	1	2	1	0	0	3	1	2
t	0.540	1.438	1.650	0.024	0.377	2.456	0.682	0.313	0.697	0.689
x	3	2	2	2	0	0	1	1	3	1
t	0.312	0.188	0.264	0.008	1.400	0.872	1.062	0.006	0.380	0.759
x	1	3	1	3	0	1	2	2	2	2
t	0.920	0.328	0.302	1.210	0.107	1.740	0.792	0.627	0.055	0.567
x	1	1	2	3	0	1	3	3	2	1
t	0.132	0.089	0.068	0.516	2.628	1.325	1.127	0.473	0.051	0.509
x	0	0	0	0	2	1	2	3	1	0
t	0.789	0.029	0.216	2.506	0.021	0.112	0.127	0.167	1.228	0.272
x	1	3	0	0	3	1	2	3	0	2
t	0.144	0.176	0.014	0.269	0.651	0.415	1.525	1.019	0.130	1.152
x	2	1	2	2	3	3	1	3	0	3
t	4.164	0.067	1.297	1.209	0.020	1.072	0.128	1.426	2.085	0.309
x	0	3	2	3	2	3	3	2	2	3
t	0.415	0.121	0.018	1.385	1.880	0.085	0.377	0.009	3.357	0.109
x	0	3	0	1	1	3	0	2	3	0

Table 9: *Simulated data set from a CPH model with a Weibull cumulative baseline hazard.*

t	0.454	0.925	0.215	2.831	20.276	0.418	2.076	0.546	1.473	6.037
x	1	3	3	1	0	0	2	2	3	3
t	0.795	0.232	1.423	3.064	5.358	0.630	2.762	2.225	0.339	2.964
x	3	3	1	3	2	0	3	0	1	2
t	9.573	0.289	12.025	1.040	8.959	2.353	8.887	2.544	1.580	0.634
x	3	2	3	3	0	1	0	2	0	3
t	0.434	0.426	0.766	12.508	1.220	1.250	0.284	0.656	1.778	0.736
x	1	1	2	2	2	1	3	1	1	2
t	1.588	13.494	3.873	3.931	0.843	2.385	2.243	1.087	1.583	3.441
x	0	0	3	0	3	0	0	3	1	2
t	1.677	2.294	1.056	1.072	5.101	1.631	1.449	9.263	3.322	0.820
x	0	0	1	2	1	3	1	1	2	0
t	2.267	7.378	10.503	1.043	0.862	0.670	2.078	5.113	2.014	5.540
x	0	1	3	0	0	2	0	1	1	0
t	0.453	2.290	1.065	1.101	0.294	12.021	0.569	0.211	0.277	0.479
x	2	2	3	3	0	1	2	1	3	0
t	2.094	11.347	3.797	27.351	11.561	3.542	0.753	0.479	0.156	5.678
x	1	3	2	2	1	2	2	0	1	0
t	0.375	19.239	1.701	1.210	7.755	4.850	1.830	1.022	11.083	0.563
x	0	2	3	1	1	3	3	2	2	2

Table 10: *Simulated data set using a CPH model with a log-normal cumulative baseline hazard.*