

Testing adequacy of fit in the random right censoring framework

E Bothma



orcid.org/0000-0002-8604-0753

Thesis accepted in fulfilment of the requirements for the degree
Doctor of Philosophy in Science with Statistics at the North-
West University

Promoter:	Prof IJH Visagie
Co-promoter:	Prof JS Allison
Assistant Promoter:	Dr M Cockeran

Graduation May 2022
26071134

Acknowledgements

I wish to express appreciation and gratitude to the following individuals:

- Prof Jaco Visagie and Prof James Allison, my supervisors, for their continuous words of encouragement, unwavering support and guidance during this study, without you none of this would have been possible. I am eternally grateful that I get to share in your passion and knowledge for statistics.
- Dr Marike Cockeran and Dr Marius Smuts for your guidance and for insightful discussions.
- My mother, brother, and sister-in-law for their unconditional love, motivation, endless patience and support. Especially my mom, for the privilege of an excellent education, without her love and encouragement I would not have made it this far.
- My friends, for always being there when I needed you.

And most importantly to God for the privilege and opportunity to be able to do what I love through His strength, guidance, love and grace.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Abstract

Statistical inference of observed lifetimes in survival analysis, reliability theory, medical studies and numerous engineering fields are frequently of interest for researchers. In order to perform inference, it is important to test the hypothesis that observed times are realised from a specified class of distributions. Testing this hypothesis is complicated by the fact that random right censoring often occurs in the mentioned fields of study; i.e., not all of the lifetimes of interest are observed. Goodness-of-fit testing has been studied in depth in the statistical literature; however, tests in the presence of random right censoring are relatively scarce, especially for more flexible distributions such as the Weibull and gamma distributions. Therefore, the main goal of this thesis is to modify existing goodness-of-fit tests to accommodate random right censoring, specifically for three classes of distributions; the exponential, Weibull and gamma distributions. We compare the finite sample performance of these tests to existing tests, such as the classical Kolmogorov-Smirnov and Cramér-von Mises tests which have been modified to allow for random censoring. The majority of the newly modified tests are based on either the characteristic function or the Laplace transform. Some of the tests developed for the Weibull distribution for use with censored data are also new in the full sample case.

We compare the finite sample performance of the tests mentioned above against a broad range of alternatives, in both the complete sample case as well as in the presence random right censoring. These classes of tests are based on Stein's method for the approximation of integrals. Furthermore, we modify existing tests based on Stein's method and the Laplace transform for the gamma distribution to allow for the presence of random right censoring. We compare the performance of these tests against a wide range of alternatives as well as three censoring distributions.

In addition, we explore the computational assumptions relating to the Kaplan-Meier estimator which is commonly used in statistical literature. The effect of these assumptions are investigated in an extensive Monte Carlo simulation study along with a practical example. The practical example under consideration is used throughout the study; the data set contains 66 observations of which 14 are censored. A further contribution is found in new bootstrap algorithms for testing the adequacy of fit of survival models. The thesis concludes with a model based application of goodness-of-fit testing for the Cox proportional hazards model.

Key words: *Censoring, Exponential distribution, Gamma distribution, Goodness-of-fit, Kaplan-Meier estimator, Weibull distribution*

Preface

This thesis is written in article format and consists of one published article, one article accepted for publication, one draft article and one accepted proceedings. A general overview of the thesis is given in Chapter 1, along with the objectives of the study. The three articles as well as the proceedings can be found in Chapters 2, 3, 4 and 5. In Chapter 6 we report the main findings along with conclusions and avenues for future research.

The first article *Characteristic function and Laplace transform based tests for exponentiality in the presence of random right censoring*, has been published in *Stat.*

The second article *New classes of tests for the Weibull distribution using Stein's method in the presence of random right censoring*, has been accepted for publication in *Computational Statistics*.

The third article *On the effect of the Kaplan-Meier estimator's assumed tail behaviour on goodness-of-fit testing*, is to be submitted to *Computational Statistics and Data Analysis*.

The conference proceedings *On an omnibus test for the parametric proportional hazards model*, has been accepted for publication in the Proceedings of the 62nd Annual Conference of the South African Statistical Association.

The promoters agreed on co-authorship and gave consent for the use of these articles as part of the final thesis. The first author was solely responsible for the initial planning and proposal of the thesis, literature searches, writing all programs used in the Monte Carlo studies, interpretation of results, as well as planning and writing of the articles and the entire thesis.

STATEMENT BY CO-AUTHORS

Herewith is a statement of the co-authors giving permission that the articles may form part of this thesis.

I hereby approve the articles and give my consent that these articles may be published as part of the thesis for the degree Doctorate of Philosophy in Science with Statistics of Miss E Bothma.



Prof JS Allison

18/10/2021

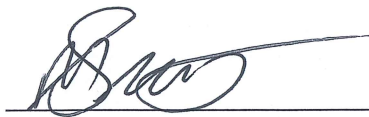
Date



Dr M Cockeran

18/10/2021

Date



Dr M Smuts

19/10/2021

Date



Prof IJH Visagie

19/10/2021

Date

Table of Contents

1	Introduction	6
1.1	Overview	6
1.2	Objectives	7
1.3	Thesis outline	8
2	Article 1: Characteristic function and Laplace transform based tests for exponential- ity in the presence of random right censoring	11
2.1	Addendum to Article 1	26
2.1.1	Epps and Pulley (1986)	26
2.1.2	Baringhaus and Henze (1991)	28
3	Article 2: New classes of tests for the Weibull distribution using Stein's method in the presence of random right censoring	30
3.1	Addendum to Article 2	51
4	Article 3: On the effect of the Kaplan-Meier estimator's assumed tail behaviour on goodness-of-fit testing	54
5	Article 4: On an omnibus test for the parametric proportional hazards model . . .	86
6	Conclusion	95
6.1	Overview of the goals of the papers presented	95
6.2	Overview of results	96
6.3	Concluding remarks and future research	98

CHAPTER 1

Introduction

1.1 Overview

In this section we provide a framework and general composition of the entire thesis. The research aims and objectives of the thesis are also stated.

It is often of interest for statistical practitioners as well as for other researchers to model survival, reliability or failure times. These lifetimes play a central role in various fields including survival analysis, reliability theory and numerous engineering fields. Before any inference is done, it is important to test the goodness-of-fit hypothesis that the observed lifetimes are realisations from a specific class of distributions. In this thesis, we focus on three classes of distributions, namely the exponential, Weibull and gamma distributions. In the full sample case, there is an abundance of tests available for all three of these classes of distributions. However, due to the nature of these studies, random right censoring often arises in these fields of study. For example, we may be interested in the time period that antibodies are detectable in a patient's blood after being given a particular type of Covid-19 vaccine. When the relevant data are collected, we will not necessarily be able to measure this time period for all of the patients in the study. Some may leave the study for different reasons, such as emigrating to another country while they still have detectable antibodies. In this case we do not observe the precise time of interest. As a result, testing the hypothesis of whether these observed lifetimes are realised from a specific distribution is complicated by the fact that not all of these times are observed.

Tests for lifetime distributions in the presence of random right censoring are relatively scarce. In this thesis we adapt standard tests used in the full sample case to accommodate for random right censoring. For the three classes of distributions under consideration we propose newly modified tests for censored samples. We introduce two new classes of tests for the Weibull distribution that are applicable in both the full sample case as well as in the case of random censoring. We investigate the finite sample performance using a comprehensive Monte Carlo study. Most of the tests that we modify or introduce are based on either the empirical versions of the characteristic function or the Laplace transform.

Throughout the study, we use a single censored data set to demonstrate the application of the newly modified tests. The practical example consists of 66 observed values, 14 of which are censored. This example is also used to demonstrate the required computational assumptions when estimating the lifetime and censoring distribution functions using the Kaplan-Meier estimator. The influence of these assumptions on the sizes as well as the empirical powers are studied extensively using Monte Carlo simulations.

Additionally, we consider goodness-of-fit testing for survival models by developing an omnibus test of fit for the parametric Cox proportional hazards model in the presence of random censoring. The test is able to detect deviations from the hypothesised model when the baseline distribution or the regression component of the model is misspecified.

In the last section of this thesis we include some concluding remarks along with some avenues for future research.

1.2 Objectives

The main goal of this thesis is to modify existing tests for the complete sample case to accommodate random right censoring for three classes of distributions; the exponential, Weibull and gamma classes. We further aim to propose new classes of tests for the Weibull distribution that can be applied using both complete samples as well as censored samples. We evaluate the finite sample performance of all these newly modified or proposed tests in extensive Monte Carlo simulation studies. Additionally, we are interested in the impact of some of the computational assumptions that are made when estimating the distribution function using the Kaplan-Meier estimator of the distribution function. The main objectives and aims of the thesis can be summarised as follows:

- To explore the existing literature to identify work that has already been done on this topic as well as any open questions relating to this.
- Modify existing goodness-of-fit tests for exponentiality to accommodate random right censoring and compare the finite sample performance of these tests to existing tests.
- Develop new tests for the Weibull distribution and compare the performance of these tests against a broad range of alternatives, in both the complete sample case as well as in the presence of random right censoring.
- Explore theoretical aspects regarding the method the newly proposed test is based on.

- Propose newly modified tests for the gamma distribution in the presence of random right censoring and compare the performance of these tests to those of existing tests for a broad range of alternatives.
- Explore and evaluate some of the computational assumptions, relating to the Kaplan-Meier estimator.
- Provide new bootstrap algorithms to calculate critical values when testing the adequacy of fit of survival models are proposed and implemented.
- Develop a new omnibus test for the parametric proportional hazards model

1.3 Thesis outline

In this section the general themes, titles, abstracts as well as the connections between the three articles and the conference proceedings are presented. The three articles and the proceedings included contributes to the literature on goodness-of-fit testing in the presence of random right censoring. In Chapter 6 we present some final findings and conclusions as well as some avenues for future research.

The titles and abstracts of the three articles as well as the proceedings, presented in Chapters 2, 3, 4 and 5 respectively, are given below:

Characteristic function and Laplace transform based tests for exponentiality in the presence of random right censoring

We test the composite hypothesis that lifetimes follow an exponential distribution based on observed randomly right censored data. Testing this hypothesis is complicated by the presence of this censoring, due to the fact that not all lifetimes are observed. To account for this complication, we propose modifications to tests based on the empirical characteristic function and Laplace transform. In the full sample case these empirical functions can be expressed as integrals with respect to the empirical distribution function of the lifetimes. We propose replacing this estimate of the distribution function by the Kaplan-Meier estimate. The resulting test statistics can be expressed in easily calculable forms in terms of summations of functionals of the observed data. Additionally, a general framework for goodness-of-fit testing, in the presence of random right censoring, is outlined. A Monte Carlo study is performed, the results of which indicate that the newly modified tests generally outperform the existing tests. A practical application, concerning initial remission times of leukemia patients, is discussed along with some concluding remarks and avenues for future research.

The scarcity of goodness-of-fit tests in the presence of random right censoring, particularly for more flexible distributions such as the Weibull distribution, leads us to develop new classes of goodness-of-fit tests for the Weibull distribution in the second article.

New classes of tests for the Weibull distribution using Stein's method in the presence of random right censoring

We develop two new classes of tests for the Weibull distribution based on Stein's method. The proposed tests are applied in the full sample case as well as in the presence of random right censoring. We investigate the finite sample performance of the new tests using a comprehensive Monte Carlo study. In both the absence and presence of censoring, it is found that the newly proposed classes of tests outperform competing tests against the majority of the distributions considered. In the cases where censoring is present we consider various censoring distributions. Some remarks on the asymptotic properties of the proposed tests are included. We present another result of independent interest; a test initially proposed for use with full samples is amended to allow for testing for the Weibull distribution in the presence of censoring. The techniques developed in the paper are illustrated using two practical examples.

In the first two articles we made some computational assumptions relating to the Kaplan-Meier estimator. We explore the effect of these assumptions in Article 3.

On the effect of the Kaplan-Meier estimator's assumed tail behaviour on goodness-of-fit testing

When analysing lifetime data in the presence of censoring one is often required to estimate the distribution function of the lifetimes nonparametrically. The most popular estimator used for this purpose is the Kaplan-Meier estimator. Interestingly, in its initial formulation this estimator is only defined up to the observed sample maximum. For values larger than the sample maximum two different assumptions are commonly used in the statistical literature. The first is to set the value of the estimate to one while the second is to use the value of the estimate at the sample maximum when estimating the tail of the distribution function. This paper illustrates the profound effect of these assumptions on the sizes and powers of goodness-of-fit tests for three classes of distributions often used in survival analysis. These differences are illustrated using observed remission time data. The considered classes of distributions are the exponential, Weibull and gamma. As a result of independent interest, we amend two classes of tests developed for the gamma distribution in the full sample case for use with censored data.

In the three main articles we only consider goodness-of-fit testing for censored samples in the independently and identically distributed (i.i.d.) case. In the conference proceedings we develop a goodness-of-fit test for a survival model, specifically for the parametric proportional hazards model.

On an omnibus test for the parametric proportional hazards model

We propose an omnibus test of fit for the parametric Cox proportional hazards model in the presence of random right censoring. The proposed test results from a modification of an existing test for the uniform distribution. This test is demonstrated to be able to detect deviations from the hypothesised model in two cases; first when the baseline distribution is misspecified and second when the regression component of the model is misspecified. Two modified classical tests are considered and a Monte Carlo study shows that the newly proposed test outperforms these tests for the majority of alternatives included. As a result of independent interest, we outline the procedure required to use the newly modified test in the framework of independent and identically distributed random variables.

CHAPTER 2

Article 1: Characteristic function and Laplace transform based tests for exponentiality in the presence of random right censoring

The first article, *Characteristic function and Laplace transform based tests for exponentiality in the presence of random right censoring*, has been published in *Stat*. A summary of the guidelines to authors from the journal is now presented.

Manuscript	Maximum length of 10 pages (excluding references, figures, tables and other forms of exhibit, but including text in appendices).
Title	A full title and a short title of up to 70 characters should be provided.
Abstract and keywords	A maximum of 200 words succinctly describing the article and five or six keywords or key phrases from the list of keywords on the ScholarOne site.
Tables	Tables must be on separate pages after the reference list, and not be incorporated into the main text. All tables should be referred to sequentially in the text.
References	References should be prepared according to the <i>Publication Manual of the American Psychological Association</i> (6th edition).
General formatting	A L ^A T _E X template is provided for submission.
Additional information	https://tinyurl.com/y4lu7f9f

ORIGINAL ARTICLE

Characteristic function and Laplace transform-based tests for exponentiality in the presence of random right censoring

Elzanie Bothma | James Samuel Allison | Marike Cockeran  |
Izak Jakobus Henning Visagie 

Subject Group Statistics, North-West
University, Potchefstroom, North West, 2520,
South Africa

Correspondence

Izak Jakobus Henning Visagie, Subject Group
Statistics, North-West University, 1 Hoffman
Street, Private Bag X6001, Potchefstroom,
North West, 2520, South Africa.
Email: jaco.visagie@nwu.ac.za

Funding information

National Research Foundation (NRF)

In this paper, the composite hypothesis that lifetimes follow an exponential distribution is tested based on observed randomly right censored data. Testing this hypothesis is complicated by the presence of this censoring, due to the fact that not all lifetimes are observed. To account for this complication, we propose modifications to tests based on the empirical characteristic function and Laplace transform. In the full sample case, these empirical functions can be expressed as integrals with respect to the empirical distribution function of the lifetimes. We propose replacing this estimate of the distribution function by the Kaplan–Meier estimate. The resulting test statistics can be expressed in easily calculable forms in terms of summations of functionals of the observed data. Additionally, a general framework for goodness-of-fit testing, in the presence of random right censoring, is outlined. A Monte Carlo study is performed, the results of which indicate that the newly modified tests generally outperform the existing tests. A practical application, concerning initial remission times of leukaemia patients, is discussed along with some concluding remarks and avenues for future research.

KEYWORDS

exponential distribution, goodness-of-fit testing, hypothesis testing, random right censoring, warp-speed bootstrap

1 | INTRODUCTION

The exponential distribution plays a central role in various fields such as survival analysis and reliability theory; see Klein and Moeschberger (2006). In the complete sample case, several tests for testing the hypothesis that the observed lifetimes are realizations from the exponential distribution have been developed. For an in-depth discussion of these tests, the reader is referred to Allison et al. (2017) as well as Henze and Meintanis (2005) and the references therein. In the mentioned fields, random right censoring often arises due to the nature of the study itself. Consider, for example, a medical study where the aim is to observe the lifetimes of patients with a specific disease. It often happens that a given patient is still alive at the end of the study or the patient leaves the study due to some other reason, such as dying in a car accident. Testing whether these observed lifetimes are realizations from a specified distribution is complicated by the fact that not all of these times are observed.

There is a relative scarcity of tests for exponentiality (or any other lifetime distribution) in the presence of random right censoring. One approach is to transform the censored sample to a complete sample and then use any of the existing tests for exponentiality developed for the full sample case. This approach is discussed in Balakrishnan et al. (2015). A more common approach is to use the standard tests applicable to complete samples and modify them to accommodate random right censoring. Koziol and Green (1976) derived a modification of the Cramér–von Mises test statistic in the case of a simple hypothesis, which is based on the Kaplan–Meier product limit estimate of the distribution function. Kim (2012) studied modified versions of the Kolmogorov–Smirnov and Cramér–von Mises test statistics for testing a composite hypothesis. However, this

study is limited by the fact that they are based on the random censoring model proposed by Koziol and Green (1976). This implies that, under the null hypothesis, the censoring distribution is known, which is an unrealistic assumption in practice. It is further well known that the null distribution of the test statistics is dependent on the unknown censoring distribution (see, e.g., D'Agostino & Stephens, 1986). Another test, specifically developed for the random censoring case, is the test proposed by Cox and Oakes (1984), which is based on a score function. This test has also been shown to be quite powerful in the full sample case; see Allison et al. (2017).

A number of powerful tests for exponentiality are based on either the empirical characteristic function or the empirical Laplace transform. These include tests proposed in Epps and Pulley (1986), Baringhaus and Henze (1991), Henze and Meintanis (2002a) and Henze and Meintanis (2002b). For an overview of testing procedures involving the characteristic function, see the discussion paper by Meintanis (2016). In this paper, we propose newly modified versions of these tests that can be used in the presence of random right censoring. We assume that the censoring distribution is unknown and this distribution is estimated non-parametrically. A bootstrap algorithm is employed to estimate the critical values associated with the tests. This algorithm assumes that the observed data are realized from an infinite population. The interested reader is referred to a list of references discussing finite population corrections which can be implemented if necessary.

Before proceeding, some notation is introduced. Let $X_1, \dots, X_n$ be independent and identically distributed (i.i.d.) lifetime variables with continuous distribution function F , and let $C_1, \dots, C_n$ be i.i.d. censoring variables with distribution function G , independent of $X_1, \dots, X_n$. Let

$$T_j = \min(X_j, C_j) \text{ and } \delta_j = \begin{cases} 1, & \text{if } X_j \leq C_j \\ 0, & \text{if } X_j > C_j. \end{cases}$$

Based on the observed pairs $(T_j, \delta_j), j = 1, \dots, n$, we wish to test the composite hypothesis

$$H_0 : F \text{ is the exponential distribution with expectation } 1/\lambda, \quad (1)$$

for some unknown $\lambda > 0$ against non-exponential alternatives. Denote the order statistics of $X_1, \dots, X_n$ and $T_1, \dots, T_n$ by $X_{(1)} < X_{(2)} < \dots < X_{(n)}$ and $T_{(1)} < T_{(2)} < \dots < T_{(n)}$, respectively. Note that $\delta_{(j)}$ represents the indicator variable corresponding to $T_{(j)}$. Using the notation introduced above, the Kaplan–Meier estimator, $\tilde{F}_n$, of the distribution function is given by

$$1 - \tilde{F}_n(t) = \begin{cases} 1, & t \leq T_{(1)} \\ \prod_{j=1}^{k-1} \left(\frac{n-j}{n-j+1} \right)^{\delta_{(j)}}, & T_{(k-1)} < t \leq T_{(k)}, k = 2, \dots, n. \\ \prod_{j=1}^n \left(\frac{n-j}{n-j+1} \right)^{\delta_{(j)}}, & t > T_{(n)}. \end{cases}$$

For more details about this estimator, see Kaplan and Meier (1958) and Efron (1967).

All of the test statistics under consideration make use of scaled lifetime values, denoted by $Y_j = T_j \hat{\lambda}$, where $\hat{\lambda} = \sum_{j=1}^n \delta_j / \sum_{j=1}^n T_j$ is the maximum likelihood estimate of λ . The invariance property of the exponential distribution justifies the use of these scaled values (see, e.g., Gupta & Richards, 1997).

The remainder of the paper is structured as follows. In Section 2, we indicate how four existing tests are modified to accommodate random right censoring. Since the null distribution of each of the test statistics depends on the unknown censoring distribution, we propose a parametric bootstrap procedure in Section 3 in order to compute critical values for the tests under consideration. Section 4 contains the results of a Monte Carlo study where the empirical powers of the newly modified tests are compared to those of existing tests. The paper concludes in Section 5 with an application to observed leukaemia remission times as well as some avenues for future research.

2 | PROPOSED TEST STATISTICS

The statistical literature contains several tests for exponentiality based on the empirical characteristic function and Laplace transform. The tests below are modifications of these tests obtained by estimating the mentioned functions using the Kaplan–Meier estimate of the distribution function instead of the empirical distribution function. Below, we consider a test based on the empirical characteristic function before turning our attention to tests based on the empirical Laplace transform.

Recall that the characteristic function of a random variable, X , with distribution, F , is defined to be

$$\phi(t) = E[e^{itY}] = \int e^{ity} dF(y),$$

where $i = \sqrt{-1}$. Denote the characteristic function of the standard exponential distribution by $\phi(t) = (1 + it)^{-1}$. Let F_n denote the empirical distribution function of the random variables $Y_1, \dots, Y_n$

$$F_n(y) = \sum_{j=1}^n I(Y_j \leq y),$$

where I denotes the indicator function. Using F_n to estimate F , we obtain the empirical characteristic function

$$\phi_n(t) = \int e^{ity} dF_n(y) = \frac{1}{n} \sum_{j=1}^n e^{itY_j}.$$

Upon replacing F_n by $\tilde{F}_n$, we obtain an estimate for the characteristic function based on a censored sample

$$\tilde{\phi}_n(t) = \int e^{ity} d\tilde{F}_n(y) = \sum_{j=1}^n \Delta_j e^{itY_j},$$

where Δ_j denotes the size of the jump in $\tilde{F}_n(T_{(j)}), j = 1, \dots, n$, given by

$$\Delta_1 = \frac{\delta_{(1)}}{n}, \Delta_n = \prod_{j=1}^{n-1} \left(\frac{n-j}{n-j+1} \right)^{\delta_{(j)}} \text{ and}$$

$$\Delta_j = \prod_{k=1}^{j-1} \left(\frac{n-k}{n-k+1} \right)^{\delta_{(k)}} - \prod_{k=1}^j \left(\frac{n-k}{n-k+1} \right)^{\delta_{(k)}} = \frac{\delta_{(j)}}{n-j+1} \prod_{k=1}^{j-1} \left(\frac{n-k}{n-k+1} \right)^{\delta_{(k)}}, j = 2, \dots, n-1.$$

Epps and Pulley (1986) introduced a test statistic based on the difference between the characteristic function and its empirical counterpart. The proposed test statistic is given by

$$EP_n = \frac{1}{2\pi} \int [\phi_n(t) - \phi(t)] \phi(-t) dt.$$

In the presence of random right censoring, we modify the test statistic by replacing ϕ_n with $\tilde{\phi}_n$. The resulting modified test statistic is given by

$$\tilde{EP}_n = \frac{1}{2\pi} \int [\tilde{\phi}_n(t) - \phi(t)] \phi(-t) dt.$$

After straightforward calculations, this test statistic can be expressed as

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

The null hypothesis in (1) is rejected for large values of $|\tilde{EP}_n|$.

We now turn our attention to tests based on the empirical Laplace transform. Using similar notation to that used for the various versions of the characteristic function, let ψ, ψ_n and $\tilde{\psi}_n$ respectively denote the Laplace transform, the empirical Laplace transform and the empirical version of this function obtained using the Kaplan–Meier estimate of the distribution. The resulting functions can be expressed as

$$\psi(t) = \int e^{-ty} dF(y), \psi_n(t) = \frac{1}{n} \sum_{j=1}^n e^{-tY_j} \text{ and } \tilde{\psi}_n(t) = \sum_{j=1}^n \Delta_j e^{-tY_j}.$$

Baringhaus and Henze (1991) proposed a test based on a partial differential equation involving the Laplace transform. In the presence of random right censoring, the modified version of this test becomes

$$\tilde{B}_{n,a} = n \int_0^{\infty} [(1+t)\tilde{\psi}'_n(t) + \tilde{\psi}_n(t)]^2 e^{-at} dt,$$

where $a > 0$ is a user-specified tuning parameter. After some algebra, this test statistic can be expressed as

$$\tilde{B}_{n,a} = n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{(1-Y_j)(1-Y_k)}{Y_j + Y_k + a} - \frac{Y_j + Y_k}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^3} \right].$$

The null hypothesis of exponentiality is rejected for large values of $\tilde{B}_{n,a}$.

Another test involving the Laplace transform was proposed in Henze and Meintanis (2002b). This test statistic is based on the squared difference between the Laplace transform and its empirical counterpart. In the presence of random censoring, this test can be modified to have test statistic

$$\tilde{L}_{n,a} = n \int_0^{\infty} [\tilde{\psi}_n(t) - \psi(t)]^2 (1+t)^2 e^{-at} dt,$$

with the following form that can easily be implemented:

$$\tilde{L}_{n,a} = n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{1 + (Y_j + Y_k + a + 1)^2}{(Y_j + Y_k + a)^3} \right] - 2n \sum_{j=1}^n \Delta_j \left[\frac{1 + Y_j + a}{(Y_j + a)^2} \right] + \frac{n}{a}.$$

The null hypothesis is rejected for large values of $\tilde{L}_{n,a}$.

Henze and Meintanis (2002a) proposed a goodness-of-fit test based on a characterization of the exponential distribution via the characteristic function. Modifying the proposed test, we obtain the test statistic

$$\tilde{H}_{n,a} = n \int_0^{\infty} [S_n(t) - tC_n(t)]^2 e^{-at} dt,$$

where $S_n(t) = \sum_{j=1}^n \Delta_j \sin(tY_j)$ and $C_n(t) = \sum_{j=1}^n \Delta_j \cos(tY_j)$. This statistic admits the following easily calculable form:

$$\begin{aligned} \tilde{H}_{n,a} = & \frac{an}{2} \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{1}{a^2 + (Y_j - Y_k)^2} - \frac{1}{a^2 + (Y_j + Y_k)^2} - \frac{4(Y_j + Y_k)}{(a^2 + (Y_j + Y_k)^2)^2} \right. \\ & \left. + \frac{2a^2 - 6(Y_j - Y_k)^2}{(a^2 + (Y_j - Y_k)^2)^3} + \frac{2a^2 - 6(Y_j + Y_k)^2}{(a^2 + (Y_j + Y_k)^2)^3} \right]. \end{aligned}$$

The null hypothesis is rejected for large values of $\tilde{H}_{n,a}$.

3 | BOOTSTRAP ALGORITHM

The null distribution of each of the test statistics considered depends on the unknown censoring distribution, even in the case of a simple hypothesis (see D'Agostino & Stephens, 1986). Since we will not assume any known form of the censoring distribution (e.g., the Koziol-Green model), we propose the following parametric bootstrap algorithm to estimate the critical values of the tests.

1. Based on the pairs $(T_j, \delta_j), j = 1, \dots, n$, estimate λ by $\hat{\lambda} = \sum \delta_j / \sum T_j$.
2. Obtain a parametric bootstrap sample $X_1^*, X_2^*, \dots, X_n^*$ by sampling from an exponential distribution with parameter $\hat{\lambda}$.
3. Obtain a non-parametric bootstrap sample $C_1^*, C_2^*, \dots, C_n^*$ by sampling from the Kaplan-Meier estimate of the distribution of the censoring times.
4. Set

$$T_j^* = \min(X_j^*, C_j^*) \text{ and } \delta_j^* = \begin{cases} 1, & \text{if } X_j^* \leq C_j^* \\ 0, & \text{if } X_j^* > C_j^* \end{cases}.$$

5. Calculate $\hat{\lambda}^* = \sum \delta_j^* / \sum T_j^*$, and obtain the scaled bootstrap values $Y_j^* = T_j^* \hat{\lambda}^*$.
6. Based on the data $(Y_j^*, \delta_j^*), j = 1, \dots, n$, calculate the value of the test statistic, say $S^* = S((Y_1^*, \delta_1^*), (Y_2^*, \delta_2^*), \dots, (Y_n^*, \delta_n^*))$.
7. Repeat steps 2–6 B times to obtain $S_1^*, S_2^*, \dots, S_B^*$. Obtain the order statistics, $S_{(1)}^* \leq S_{(2)}^* \leq \dots \leq S_{(B)}^*$. The estimated critical value is then $\hat{c}_n(\alpha) = S_{[B(1-\alpha)]}^*$ where $[A]$ denotes the floor of A .

The algorithm provided above is quite general and can easily be amended in order to test for any lifetime distribution in the presence of random censoring. The algorithm implicitly assumes an infinite underlying population. In the alternative case of a finite population, a finite population correction can be made; for more details, we refer to the recent paper Conti et al. (2017) and the references therein.

4 | MONTE CARLO STUDY

In this section, the power behaviour of the newly modified tests is investigated by means of a Monte Carlo study. The $\widetilde{EP}_n$, $\widetilde{L}_{n,a}$, $\widetilde{B}_{n,a}$ and $\widetilde{H}_{n,a}$ tests are compared to the modified Kolmogorov–Smirnov ($\widetilde{KS}_n$) and Cramér–von Mises ($\widetilde{CM}_n$) tests proposed in Koziol and Green (1976) as well as the Cox and Oakes (1984) test ($\widetilde{CO}_n$), which was originally proposed for use with censored data.

Let $\tilde{n}$ denote the number of observations which are uncensored, that is, $\tilde{n} = \sum \delta_j$. The calculable forms of the Kolmogorov–Smirnov, Cramér–von Mises and Cox and Oakes tests are, respectively, given by

$$\begin{aligned} \widetilde{KS}_n &= \sup_{x \geq 0} |\widetilde{F}_n(x) - (1 - e^{-x})| = \max \left[\max_{1 \leq j \leq n} \left\{ \widetilde{F}_n(Y_{(j)}) - (1 - e^{-Y_{(j)}}) \right\}, \max_{1 \leq j \leq n} \left\{ (1 - e^{-Y_{(j)}}) - \widetilde{F}_n(Y_{(j)}) \right\} \right], \\ \widetilde{CM}_n &= n \int_0^1 (t - \widetilde{F}_n(t))^2 dt = \frac{n}{3} + n \sum_{j=1}^{\tilde{n}+1} \widetilde{F}_n(\widetilde{Y}_{j-1}) (\widetilde{Y}_j - \widetilde{Y}_{j-1}) \left[\widetilde{F}_n(\widetilde{Y}_{j-1}) - (\widetilde{Y}_j + \widetilde{Y}_{j-1}) \right] \text{ and} \\ \widetilde{CO}_n &= \tilde{n} + \sum_{j=1}^n \log(Y_j) \delta_j - \frac{\tilde{n} \sum_{j=1}^n Y_j \log(Y_j)}{\sum_{j=1}^n Y_j}. \end{aligned}$$

The null hypothesis of exponentiality is rejected in the case of large values of $\widetilde{KS}_n$ and $\widetilde{CM}_n$, while the Cox and Oakes test rejects for small or large values of $\widetilde{CO}_n$.

4.1 | Simulation setting

The nominal significance level is set to 5% throughout. Empirical powers are presented for sample sizes of $n = 50$ and $n = 100$. The reported empirical powers are calculated in the case of 10%, 20% and 30% censoring. For each lifetime distribution considered, we include the power achieved using three different censoring distributions: the exponential, uniform and Lindley distributions. The alternative distributions used are listed in Table 1.

TABLE 1 Density functions of the alternative distributions

Alternative	Density	Notation
Exponential	$\theta \exp(-\theta x)$	$\text{Exp}(\theta)$
Gamma	$(\Gamma(\theta))^{-1} x^{\theta-1} \exp(-x)$	$\Gamma(\theta)$
Weibull	$\theta x^{\theta-1} \exp(-x^\theta)$	$W(\theta)$
Lognormal	$(\theta x \sqrt{2\pi})^{-1} \exp(-\log^2(x) / (2\theta^2))$	$LN(\theta)$
Chi square	$(2^{\theta/2} \Gamma(\theta/2))^{-1} x^{\theta/2-1} \exp(-x/2)$	$\chi^2(\theta)$
Beta	$x^{\alpha-1} (1-x)^{\theta-1} \Gamma(\alpha+\theta) (\Gamma(\alpha)\Gamma(\theta))^{-1}$	$\beta(\alpha, \theta)$

TABLE 2 Estimated powers for 10% censoring for a sample size of $n = 50$ with three different censoring distributions

F	$\widetilde{KS}$	$\widetilde{CM}$	$\widetilde{CO}$	$\widetilde{EP}$	$\widetilde{L}_{25}$	$\widetilde{L}_5$	$\widetilde{B}_{25}$	$\widetilde{B}_5$	$\widetilde{H}_5$	$\widetilde{H}_1$
$Exp(1)$	5	5	5	5	5	5	5	4	5	5
	5	5	5	5	5	5	4	4	5	5
	5	5	5	5	5	5	4	4	5	5
$\Gamma(0.6)$	56	63	81	63	82	79	80	75	55	56
	55	62	82	63	82	80	79	75	56	55
	55	62	82	63	82	79	79	75	56	56
$\Gamma(0.8)$	14	16	25	18	27	25	24	22	14	14
	14	16	25	18	27	24	24	22	13	14
	14	16	25	18	28	25	24	21	14	15
$\Gamma(1.2)$	10	12	14	11	11	14	11	10	9	11
	11	12	14	12	12	14	11	10	9	11
	10	11	14	12	11	13	10	8	9	11
$W(0.8)$	32	37	51	42	49	48	50	47	27	31
	31	37	50	42	49	49	48	47	27	31
	31	38	51	41	48	48	48	46	27	31
$W(1.2)$	21	25	29	26	23	28	23	23	16	22
	20	25	30	26	23	27	22	22	16	23
	21	25	29	26	22	27	21	20	15	22
$LN(1)$	21	26	14	13	27	20	24	18	29	16
	19	23	12	11	28	20	24	18	28	15
	19	23	13	11	28	20	23	16	29	15
$LN(1.5)$	82	87	83	90	70	80	85	87	59	78
	81	86	82	89	68	79	84	87	56	76
	79	84	81	88	66	78	83	85	54	75
$\chi^2(1)$	82	87	97	86	97	96	96	94	83	81
	81	86	97	86	97	95	96	94	82	81
	82	87	97	86	97	95	95	94	83	82
$\chi^2(3)$	33	40	51	40	46	50	44	42	30	37
	34	40	51	40	45	49	42	39	29	37
	34	41	50	39	46	50	39	34	29	36
$\beta(1, 1)$	89	97	85	95	54	75	82	88	78	97
	89	97	86	95	55	74	81	87	78	97
	89	97	85	95	54	74	81	87	78	97
$\beta(0.5, 1)$	35	47	49	8	69	54	60	43	55	49
	34	47	50	7	69	54	60	42	55	48
	35	46	50	8	69	54	60	43	55	48
$\beta(0.7, 1)$	38	49	13	32	9	9	14	15	28	49
	37	48	12	32	9	9	15	16	27	50
	37	49	12	31	9	9	14	15	28	49
$\beta(1, 1.5)$	60	76	63	77	32	48	52	61	34	70
	59	75	62	77	33	48	52	59	33	70
	59	75	63	77	32	48	52	60	33	70

TABLE 3 Estimated powers for 20% censoring for a sample size of $n = 50$ with three different censoring distributions

F	$\widetilde{KS}$	$\widetilde{CM}$	$\widetilde{CO}$	$\widetilde{EP}$	$\widetilde{L}_{25}$	$\widetilde{L}_5$	$\widetilde{B}_{25}$	$\widetilde{B}_5$	$\widetilde{H}_5$	$\widetilde{H}_1$
$Exp(1)$	4	4	5	5	5	5	4	3	4	4
	4	4	5	4	5	5	4	4	4	4
	4	4	5	5	5	5	4	3	4	4
$\Gamma(0.6)$	47	57	79	56	80	77	74	67	49	49
	48	56	80	55	80	77	73	63	48	48
	47	56	80	56	79	76	73	65	49	49
$\Gamma(0.8)$	11	14	23	15	26	23	21	17	12	13
	11	14	23	15	26	23	21	17	11	12
	11	14	24	15	26	23	20	16	12	13
$\Gamma(1.2)$	8	10	13	9	11	13	7	5	7	9
	8	10	13	10	11	13	4	2	7	9
	8	9	13	10	11	12	5	2	7	8
$W(0.8)$	25	31	46	35	45	44	42	37	21	26
	24	30	45	33	44	44	42	37	21	25
	24	31	46	34	45	44	41	35	21	26
$W(1.2)$	16	20	26	20	21	25	14	11	12	18
	16	19	26	20	20	24	9	5	12	18
	15	20	26	20	20	24	11	6	12	17
$LN(1)$	17	21	16	11	31	23	20	13	24	14
	14	18	16	8	34	25	16	9	23	13
	15	18	16	9	32	24	16	10	22	13
$LN(1.5)$	66	73	67	78	53	66	72	74	40	62
	60	69	60	72	45	59	69	69	33	57
	60	69	62	72	46	60	67	67	34	57
$\chi^2(1)$	75	82	96	79	96	94	94	90	76	75
	75	82	96	79	96	94	93	88	76	75
	75	82	96	79	96	94	93	88	76	75
$\chi^2(3)$	26	34	48	32	43	46	31	23	24	31
	26	32	47	32	42	46	21	10	22	30
	25	32	46	30	43	47	19	9	22	29
$\beta(1, 1)$	85	93	77	84	48	68	72	80	70	94
	85	93	77	80	48	66	68	76	69	94
	85	93	77	82	48	67	70	76	70	94
$\beta(0.5, 1)$	33	44	56	9	70	57	58	39	49	42
	31	43	58	10	71	58	55	33	49	41
	33	43	57	9	71	57	56	38	49	42
$\beta(0.7, 1)$	30	38	9	21	9	8	9	8	23	40
	29	35	9	18	10	8	8	6	21	39
	29	37	9	19	10	8	8	8	22	39
$\beta(1, 1.5)$	49	65	52	59	27	41	37	43	27	60
	48	62	51	57	28	40	32	33	26	58
	49	63	52	59	28	41	36	39	26	60

TABLE 4 Estimated powers for 30% censoring for a sample size of $n = 50$ with three different censoring distributions

F	$\widetilde{KS}$	$\widetilde{CM}$	$\widetilde{CO}$	$\widetilde{EP}$	$\widetilde{L}_{25}$	$\widetilde{L}_5$	$\widetilde{B}_{25}$	$\widetilde{B}_5$	$\widetilde{H}_5$	$\widetilde{H}_1$
$Exp(1)$	3	3	5	3	5	5	3	3	4	4
	3	3	5	2	5	5	4	4	4	4
	3	3	5	3	5	5	3	3	4	4
$\Gamma(0.6)$	33	46	76	42	77	73	62	49	38	40
	33	42	77	33	76	73	58	45	35	38
	33	45	77	41	76	72	60	46	38	40
$\Gamma(0.8)$	7	10	22	11	24	21	15	11	9	10
	7	9	22	6	25	21	16	13	8	9
	7	10	23	10	25	21	14	10	9	10
$\Gamma(1.2)$	5	7	12	6	10	11	2	1	5	7
	5	4	12	5	10	11	1	1	5	6
	5	6	12	5	10	11	1	1	5	6
$W(0.8)$	16	23	41	25	41	40	32	24	16	20
	14	20	40	15	40	38	30	25	15	18
	15	22	41	22	41	39	29	22	16	20
$W(1.2)$	9	13	23	11	19	21	5	1	8	12
	9	9	23	9	18	20	1	0	8	12
	9	11	23	10	18	21	2	1	8	12
$LN(1)$	10	14	21	8	34	26	10	6	16	10
	7	9	23	6	39	29	4	3	16	9
	8	11	21	7	36	27	6	4	15	9
$LN(1.5)$	41	53	47	55	36	48	52	51	22	42
	32	42	31	29	25	37	46	44	15	34
	35	46	40	41	29	42	46	44	18	37
$\chi^2(1)$	59	74	95	65	94	92	87	76	65	65
	61	71	94	55	94	92	83	70	63	64
	59	74	95	63	94	92	86	73	64	65
$\chi^2(3)$	16	23	44	18	40	42	12	3	16	21
	15	15	42	16	39	39	1	0	15	21
	13	17	43	15	39	41	3	1	13	19
$\beta(1, 1)$	76	85	67	54	42	59	55	59	58	87
	75	79	64	39	40	55	37	31	51	84
	77	84	66	49	41	57	51	53	56	86
$\beta(0.5, 1)$	25	37	63	9	72	59	47	25	40	34
	21	29	67	12	74	63	36	14	39	30
	24	35	64	10	73	59	46	23	40	33
$\beta(0.7, 1)$	21	24	7	9	10	7	4	2	16	27
	18	15	8	5	11	7	1	0	12	21
	20	22	8	8	10	7	4	1	15	26
$\beta(1, 1.5)$	35	46	42	33	23	33	18	14	19	45
	31	34	38	21	22	30	5	1	14	38
	34	45	42	30	23	33	15	10	17	43

TABLE 5 Estimated powers for 10% censoring for a sample size of $n = 100$ with three different censoring distributions

F	$\widetilde{KS}$	$\widetilde{CM}$	$\widetilde{CO}$	$\widetilde{EP}$	$\widetilde{L}_{25}$	$\widetilde{L}_5$	$\widetilde{B}_{25}$	$\widetilde{B}_5$	$\widetilde{H}_5$	$\widetilde{H}_1$
$Exp(1)$	5	5	5	5	5	5	5	4	5	5
	5	5	5	5	5	5	5	4	5	5
	5	5	5	5	5	5	4	4	5	5
$\Gamma(0.6)$	85	90	98	89	97	97	97	96	85	86
	85	90	98	89	98	97	97	96	85	86
	85	90	97	89	97	97	97	96	86	86
$\Gamma(0.8)$	24	28	42	31	44	41	41	37	24	25
	24	28	43	30	45	41	41	37	24	25
	24	29	43	30	44	42	41	38	24	26
$\Gamma(1.2)$	16	19	26	20	23	25	22	21	15	18
	16	19	25	20	22	24	21	19	15	18
	16	20	25	20	22	24	20	18	15	18
$W(0.8)$	56	65	78	69	75	75	76	75	49	57
	56	64	78	69	74	75	76	74	49	57
	56	64	77	69	74	75	75	74	50	57
$W(1.2)$	38	47	56	51	46	51	48	49	29	42
	39	46	56	51	46	52	48	47	29	41
	38	47	55	51	47	52	46	45	30	41
$LN(1)$	40	50	18	14	58	36	58	36	58	30
	38	48	16	11	60	37	57	37	58	31
	38	48	18	11	61	39	55	34	58	30
$LN(1.5)$	98	99	98	99	93	97	98	99	88	97
	98	99	98	99	92	97	98	99	87	97
	98	99	98	99	91	97	98	99	86	96
$\chi^2(1)$	98	99	100	99	100	100	100	100	98	98
	99	99	100	99	100	100	100	100	98	98
	99	99	100	99	100	100	100	100	99	98
$\chi^2(3)$	62	71	84	71	81	83	80	78	58	67
	62	71	83	71	81	83	79	75	58	66
	62	71	84	71	80	82	75	71	57	65
$\beta(1, 1)$	100	100	99	100	87	97	99	100	99	100
	100	100	99	100	86	97	99	100	99	100
	100	100	99	100	87	97	99	100	99	100
$\beta(0.5, 1)$	68	83	73	9	91	80	87	74	87	82
	67	83	73	9	91	79	87	73	86	82
	67	83	73	9	91	80	87	73	86	82
$\beta(0.7, 1)$	74	87	18	62	10	12	32	34	60	88
	74	87	18	61	11	12	31	34	59	88
	74	86	18	60	11	12	32	32	58	88
$\beta(1, 1.5)$	91	98	91	98	61	80	88	92	68	97
	92	98	91	98	61	81	88	92	68	97
	91	98	91	98	61	81	87	92	69	97

TABLE 6 Estimated powers for 20% censoring for a sample size of $n = 100$ with three different censoring distributions

F	$\widetilde{KS}$	$\widetilde{CM}$	$\widetilde{CO}$	$\widetilde{EP}$	$\widetilde{L}_{25}$	$\widetilde{L}_5$	$\widetilde{B}_{25}$	$\widetilde{B}_5$	$\widetilde{H}_5$	$\widetilde{H}_1$
$Exp(1)$	5	5	5	5	5	5	4	3	5	5
	4	5	5	5	5	5	4	4	4	4
	4	5	5	5	5	5	3	3	5	5
$\Gamma(0.6)$	80	86	97	84	97	96	95	92	81	80
	80	86	97	85	97	96	94	90	80	79
	79	86	97	84	96	96	95	92	81	80
$\Gamma(0.8)$	20	25	41	27	42	39	36	30	21	22
	20	25	41	26	42	38	34	28	20	22
	20	26	41	27	42	39	34	29	21	22
$\Gamma(1.2)$	14	17	24	18	22	23	15	11	13	15
	13	16	24	18	21	23	9	4	12	14
	13	17	23	17	21	22	12	6	12	14
$W(0.8)$	47	57	73	61	70	71	69	65	43	49
	46	55	73	59	69	70	67	61	42	48
	46	56	72	60	70	69	66	61	42	50
$W(1.2)$	31	40	51	42	42	47	35	29	24	34
	31	39	50	42	42	47	22	13	23	33
	31	39	50	43	43	47	30	20	24	34
$LN(1)$	35	45	25	12	65	44	53	30	53	27
	34	41	28	8	70	49	46	24	53	27
	33	43	29	10	69	48	46	23	52	26
$LN(1.5)$	92	95	92	97	81	91	95	95	72	91
	90	94	89	96	74	87	92	93	65	88
	90	94	88	95	74	87	92	92	67	88
$\chi^2(1)$	97	99	100	98	100	100	100	100	97	97
	97	99	100	98	100	100	100	99	97	97
	97	99	100	98	100	100	100	99	98	97
$\chi^2(3)$	54	65	81	63	78	80	68	58	51	58
	52	64	80	63	78	79	51	32	49	56
	50	63	81	64	77	79	51	32	47	54
$\beta(1, 1)$	100	100	97	99	82	94	98	99	98	100
	100	100	97	99	80	93	97	99	98	100
	100	100	97	99	81	94	98	99	98	100
$\beta(0.5, 1)$	66	80	80	11	92	83	86	71	83	76
	65	79	82	15	92	83	84	66	82	75
	66	80	81	13	92	83	86	69	82	76
$\beta(0.7, 1)$	69	78	11	45	11	9	20	21	50	80
	68	75	11	39	12	9	16	15	49	78
	69	76	11	42	11	9	20	18	51	79
$\beta(1, 1.5)$	86	95	84	94	54	73	77	83	60	94
	86	94	82	92	53	72	71	76	58	93
	86	95	83	93	54	73	74	81	59	93

TABLE 7 Estimated powers for 30% censoring for a sample size of $n = 100$ with three different censoring distributions

F	$\widetilde{KS}$	$\widetilde{CM}$	$\widetilde{CO}$	$\widetilde{EP}$	$\widetilde{L}_{25}$	$\widetilde{L}_5$	$\widetilde{B}_{25}$	$\widetilde{B}_5$	$\widetilde{H}_5$	$\widetilde{H}_1$
$Exp(1)$	4	4	5	4	5	5	2	2	4	4
	3	3	5	3	5	5	4	3	4	4
	3	4	5	4	5	5	3	2	4	4
$\Gamma(0.6)$	66	80	95	73	95	94	90	80	71	71
	66	78	96	65	95	94	82	67	68	69
	66	80	96	74	95	94	88	77	71	70
$\Gamma(0.8)$	14	20	39	22	40	37	26	17	17	18
	12	19	39	14	39	36	22	18	14	15
	14	20	39	21	39	36	24	15	16	18
$\Gamma(1.2)$	9	13	23	13	21	21	6	2	9	11
	7	8	22	13	19	20	1	1	8	10
	7	12	21	12	20	20	3	1	8	10
$W(0.8)$	34	46	67	48	65	65	55	44	33	39
	30	42	65	35	63	63	47	37	28	35
	31	45	66	46	65	64	50	38	32	38
$W(1.2)$	20	31	46	30	38	42	15	5	16	25
	17	21	44	28	37	39	1	0	14	23
	18	29	45	29	39	42	8	2	16	23
$LN(1)$	24	37	35	10	71	51	34	15	41	21
	18	28	44	14	77	60	11	5	38	19
	19	33	40	11	74	56	23	8	38	20
$LN(1.5)$	75	85	76	87	61	77	81	80	49	75
	65	77	59	62	46	64	70	66	35	66
	69	81	67	77	52	70	71	67	41	69
$\chi^2(1)$	93	97	100	93	100	100	99	98	94	93
	93	97	100	88	100	100	98	92	93	92
	92	97	100	93	100	100	99	96	94	93
$\chi^2(3)$	38	54	77	47	75	76	39	16	38	45
	31	41	75	46	73	73	4	0	34	41
	28	46	76	44	73	74	14	2	31	39
$\beta(1, 1)$	100	100	93	90	74	90	95	97	96	100
	100	99	92	78	72	86	88	92	94	100
	100	100	93	87	73	88	93	97	95	100
$\beta(0.5, 1)$	56	75	86	14	93	85	81	57	75	66
	48	66	90	28	94	87	59	22	71	58
	56	75	87	16	93	85	79	52	75	65
$\beta(0.7, 1)$	59	62	8	25	12	8	9	8	37	64
	52	46	8	12	13	8	2	0	29	52
	58	60	8	22	12	8	8	5	36	63
$\beta(1, 1.5)$	78	86	71	74	46	63	54	55	45	85
	73	76	67	54	42	58	12	3	36	77
	78	86	71	71	45	62	46	46	44	84

TABLE 8 Initial remission times of leukaemia patients, in days

4, 5, 8, 8, 9, 10, 10, 10, 10, 10, 11, 12, 12, 12*, 13, 14, 20, 20*, 23, 23, 25, 25, 25, 28, 28, 28,
 28, 29, 31, 31, 31, 32, 37, 40, 41, 41, 48, 48, 57, 62, 70, 74, 75, 89, 99, 100, 103, 124, 139, 143,
 159*, 161*, 162, 169, 190*, 195, 196*, 197*, 199*, 205*, 217*, 219*, 220, 245*, 258*, 269*

TABLE 9 p -values associated with the various tests used

Test	$\widetilde{KS}$	$\widetilde{CM}$	$\widetilde{CO}$	$\widetilde{EP}$	$\widetilde{L}_{.25}$	$\widetilde{L}_{.5}$	$\widetilde{B}_{.25}$	$\widetilde{B}_{.5}$	$\widetilde{H}_{.5}$	$\widetilde{H}_1$
p -value	<0.01	<0.01	0.03	0.11	0.13	0.03	<0.01	<0.01	0.06	<0.01

A warp-speed bootstrap methodology (Giacomini et al., 2013), which essentially entails using a single bootstrap replication for each Monte Carlo sample generated, is employed in order to calculate empirical powers. This methodology is used in order to reduce the computational cost of the calculation of these powers and has been employed by a number of authors in the literature to compare Monte Carlo performances (see, e.g., Hušková et al., 2020).

The estimated powers are reported in Tables 2–7. In the interest of continuity, these tables are deferred to the end of the paper. The reported empirical powers represent the percentages of 50,000 independent Monte Carlo samples that resulted in the rejection of the null hypothesis (rounded to the nearest integer). For each test considered, Tables 2–7 show three empirical powers against each of the lifetime distributions, corresponding to the three different censoring distributions used. In each case, the results for the exponential, uniform and Lindley distributions are shown in the first, second and third lines, respectively. For the reader's comfort, the highest power against each alternative distribution is displayed in bold in the simulation results.

In the discussion below, including the tables, the subscript n is suppressed. For $\widetilde{L}$ and $\widetilde{B}$, we include numerical powers in the cases where a is set to 0.25 and 0.5, while $a = 0.5$ and $a = 1$ are used in conjunction with $\widetilde{H}$. All calculations were performed in R (Core Team R, 2019). The LindleyR package was used to generate samples from censored distributions; see Mazucheli et al. (2016). The tables were produced using the Stargazer package (Hlavac, 2018).

4.2 | Simulation results

The size of the tests is maintained closely for all sample sizes and for censoring proportions of 10% and 20%, perhaps with the single exception of $\widetilde{B}$ in the case of 20% censoring. For 30% censoring, we find that $\widetilde{KS}$, $\widetilde{CM}$, $\widetilde{EP}$ and $\widetilde{B}$ are all slightly conservative, obtaining estimated sizes ranging between 2% and 4%. For small censoring proportions, the effect of the censoring distribution on the power of a given test is not pronounced. However, as the censoring proportion increases, these differences become more noticeable. As one would expect, the powers generally increase with the sample size and decrease slightly as the censoring proportion increases. We also estimated the powers for the complete sample case; the results that we obtained are similar to those of Allison et al. (2017) and are not included below. The choice of the tuning parameter, a , also plays a role in the performance of $\widetilde{L}$, $\widetilde{B}$ and $\widetilde{H}$. Overall, these tests have higher powers for smaller values of a .

In general, the $\widetilde{CO}$ and $\widetilde{L}$ tests outperform their competitors. As a result, we recommend using either the $\widetilde{CO}$ or $\widetilde{L}_{.25}$ test in practice. It is interesting to note that, when we disregard the $\widetilde{CO}$, the newly modified tests clearly outperform the modified tests based on the distribution function.

5 | PRACTICAL APPLICATION AND CONCLUSION

In this section, we apply each of the tests discussed in Section 4 to a real-world data set. The initial remission times of leukaemia patients, in days, were reported in Lee and Wang (2003); this data set can be found in Table 8. An asterisk is used in order to indicate that the observation is censored. The original data were segmented into three treatment groups. However, Lee and Wang (2003) showed that the data do not display significant differences among the various treatments. As a result, we treat the data as i.i.d. realizations from a single, censored, lifetime distribution. Again, we will assume an infinite population size. The estimated p -values for each of the tests, obtained using 1 million bootstrap replications, are displayed in Table 9.

The results from the practical example show that each of the tests, except $\widetilde{EP}$, $\widetilde{L}_{25}$ and $\widetilde{H}_5$, rejects the null hypothesis of exponentiality of the initial remission times at a 5% significance level. This provides some evidence that a more flexible distribution, perhaps with heavier tails, should be used as a model for this specific data set.

In this paper, we modify four tests for exponentiality, based on either the characteristic function or the Laplace transform, to accommodate random right censoring. These tests are compared to existing tests using a Monte Carlo study, and it is found that the newly modified tests are competitive in terms of their power performance. This is especially true for the test based on the squared difference between the Laplace transform and its empirical counterpart, $\widetilde{L}$, which, together with the $\widetilde{CO}$ test, is recommended for use in practice.

Although we only consider these tests in the i.i.d. set-up, these modified tests can also be used to test the adequacy of fit of various survival models when random right censoring is present. These models include the conventional Cox proportional hazards model (see, e.g., Cockeran et al., 2019) as well as the mixture cure model, which is currently receiving a lot of attention in the literature; see Amico and Van Keilegom (2018). Recently, Betsch and Ebner (2019) introduced a novel idea of basing goodness-of-fit tests on a fixed point property of a new transformation connected to the Stein characterization. These tests were found to be quite powerful in the full sample case, and it will be interesting to see whether they can be successfully modified to the censored case.

ACKNOWLEDGEMENTS

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

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

DATA AVAILABILITY STATEMENT

The initial remission times of leukemia patients, in days, were reported in Lee and Wang (2003), this data set can be found in Table 2.

ORCID

Marike Cockeran  <https://orcid.org/0000-0002-3990-8345>

Izak Jakobs Henning Visagie  <https://orcid.org/0000-0002-4661-6180>

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.
- Amico, M., & Van Keilegom, I. (2018). Cure models in survival analysis. *Annual Review of Statistics and Its Application*, 5, 311–342.
- Balakrishnan, N., Chimitova, E., & Vedernikova, M. (2015). An empirical analysis of some nonparametric goodness-of-fit tests for censored data. *Communications in Statistics-Simulation and Computation*, 44(4), 1101–1115.
- Baringhaus, L., & Henze, N. (1991). A class of consistent tests for exponentiality based on the empirical Laplace transform. *Annals of the Institute of Statistical Mathematics*, 43(3), 551–564.
- Betsch, S., & Ebner, B. (2019). A new characterization of the gamma distribution and associated goodness-of-fit tests. *Metrika*, 82(7), 779–806.
- Cockeran, M., Meintanis, S. G., & Allison, J. S. (2019). Goodness-of-fit tests in the Cox proportional hazards model. *Communications in Statistics-Simulation and Computation*, 1–12. <https://doi.org/10.1080/03610918.2019.1639738>
- Conti, P. L., Marella, D., Mecatti, F., & Andreis, F. (2017). A unified principled framework for resampling based on pseudo-populations: Asymptotic theory. arXiv preprint arXiv:1705.03827.
- Core Team R. (2019). *R: A language and environment for statistical computing*. Vienna, Austria, R Foundation for Statistical Computing. <https://www.R-project.org/>
- Cox, D. R., & Oakes, D. (1984). *Analysis of survival data* (Vol. 21). London: CRC Press.
- D'Agostino, R. B., & Stephens, M. A. (1986). *Goodness-of-fit techniques* (Vol. 68). London: CRC Press.
- Efron, B. (1967). The two sample problem with censored data. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, (Vol. 4, pp. 831–853). Berkeley, CA: University of California Press.
- Epps, T. W., & Pulley, L. B. (1986). A test of exponentiality vs. monotone-hazard alternatives derived from the empirical characteristic function. *Journal of the Royal Statistical Society: Series B (Methodological)*, 48(2), 206–213.
- Giacomini, R., Politis, D. N., & White, H. (2013). A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory*, 29(3), 567–589.
- Gupta, R. D., & Richards, D. S. P. (1997). Invariance properties of some classical tests for exponentiality. *Journal of statistical planning and inference*, 63(2), 203–213.
- Henze, N., & Meintanis, S. G. (2002a). Goodness-of-fit tests based on a new characterization of the exponential distribution. *Communications in Statistics-Theory and Methods*, 31(9), 1479–1497.
- Henze, N., & Meintanis, S. G. (2002b). Tests of fit for exponentiality based on the empirical Laplace transform. *Statistics*, 36(2), 147–161.
- Henze, N., & Meintanis, S. G. (2005). Recent and classical tests for exponentiality: A partial review with comparisons. *Metrika*, 61(1), 29–45.

- Hlavac, M. (2018). stargazer: Well-formatted regression and summary statistics tables. <https://CRAN.R-project.org/package=stargazer>
- Hušková, M., Meintanis, S. G., & Pretorius, C. (2020). Tests for validity of the semiparametric heteroskedastic transformation model. *Computational Statistics & Data Analysis*, 144. <https://doi.org/10.1016/j.csda.2019.106895>
- Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282), 457–481.
- Kim, N.-H. (2012). Testing exponentiality based on EDF statistics for randomly censored data when the scale parameter is unknown. *The Korean Journal of Applied Statistics*, 25(2), 311–319.
- Klein, J. P., & Moeschberger, M. L. (2006). *Survival analysis: Techniques for censored and truncated data*. New York: Springer Science & Business Media.
- Kozioł, J. A., & Green, S. B. (1976). A Cramér-von Mises statistic for randomly censored data. *Biometrika*, 63(3), 465–474.
- Lee, E. T., & Wang, J. (2003). *Statistical methods for survival data analysis* (Vol. 476). New Jersey: John Wiley & Sons.
- Mazucheli, J., Fernandes, L. B., & de Oliveira, R. P. (2016). LindleyR: The Lindley distribution and its modifications. <https://CRAN.R-project.org/package=LindleyR>. R package version 1.1.0.
- Meintanis, S. G. (2016). A review of testing procedures based on the empirical characteristic function. *South African Statistical Journal*, 50(1), 1–14.

How to cite this article: Bothma, E., Allison, J. S., Cockeran, M., & Visagie, I. J. H. (2021). Characteristic function and Laplace transform-based tests for exponentiality in the presence of random right censoring. *Stat*, 10(1), e394. <https://doi.org/10.1002/sta4.394>

2.1 Addendum to Article 1

Below, we include the derivations of the computational forms of several of the test statistics discussed in Section 2 of this paper.

2.1.1 Epps and Pulley (1986)

In the full sample case the Epps and Pulley (1986) test is given by

$$EP_n = \int_{-\infty}^{\infty} [\phi_n(t) - \phi(t)] \frac{1}{2\pi} \phi(-t) dt.$$

In the presence of random censoring this test is modified to be

$$\begin{aligned} \widetilde{EP}_n &= \int_{-\infty}^{\infty} [\widetilde{\phi}_n(t) - \phi(t)] \frac{1}{2\pi} \phi(-t) dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\sum_{j=1}^n \Delta_j e^{itY_j} - \frac{1}{1-it} \right) \frac{1}{1+it} dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\sum_{j=1}^n \Delta_j e^{itY_j} \frac{1}{1+it} - \frac{1}{1+t^2} \right) dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\frac{1-it}{1+t^2} \sum_{j=1}^n \Delta_j e^{itY_j} - \frac{1}{1+t^2} \right) dt \\ &= \frac{1}{2\pi} \left(\sum_{j=1}^n \Delta_j \int_{-\infty}^{\infty} \frac{1-it}{1+t^2} e^{itY_j} dt - \int_{-\infty}^{\infty} \frac{1}{1+t^2} dt \right) \\ &= \frac{1}{2\pi} \left(\sum_{j=1}^n \Delta_j I_1 - I_2 \right), \end{aligned}$$

where $I_1 = \int_{-\infty}^{\infty} (1-it)/(1+t^2) e^{itY_j} dt$ and $I_2 = \int_{-\infty}^{\infty} 1/(1+t^2) dt$.

Next, we calculate these two integrals in turn. The second integral is

$$I_2 = \int_{-\infty}^{\infty} \frac{1}{1+t^2} dt = \tan^{-1}(t)|_{-\infty}^{\infty} = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi.$$

The first integral above can be divided into two integrals as follows;

$$I_1 = \int_{-\infty}^{\infty} \frac{1-it}{1+t^2} e^{itY_j} dt$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} \frac{1}{1+t^2} e^{itY_j} dt - i \int_{-\infty}^{\infty} \frac{t}{1+t^2} e^{itY_j} dt \\
&= I_{11} - iI_{12}.
\end{aligned}$$

When calculating I_{11} , we obtain

$$\begin{aligned}
I_{11} &= \int_{-\infty}^{\infty} \frac{1}{1+t^2} e^{itY_j} dt \\
&= \pi \int_{-\infty}^{\infty} \frac{1}{\pi} \frac{1}{1+t^2} e^{itY_j} dt \\
&= \pi E \left[e^{iY_j T} \right], \text{ with } T \sim \text{Cauchy}(1), \\
&= \pi e^{0-|Y_j|} \\
&= \pi e^{-Y_j}.
\end{aligned}$$

Turning our attention to I_{12} , we have

$$\begin{aligned}
I_{12} &= \int_{-\infty}^{\infty} \frac{t}{1+t^2} e^{itY_j} dt \\
&= \int_{-\infty}^{\infty} \frac{t}{1+t^2} [\cos(tY_j) + i \sin(tY_j)] dt \\
&= \int_{-\infty}^{\infty} \frac{t}{1+t^2} \cos(tY_j) dt + i \int_{-\infty}^{\infty} \frac{t}{1+t^2} \sin(tY_j) dt \\
&= 0 + i\pi e^{-Y_j} \\
&= i\pi e^{-Y_j}.
\end{aligned}$$

After putting all of these together we are left with the following easily calculable form for the modified Epps and Pulley (1986) test

$$\begin{aligned}
\frac{1}{4\sqrt{3n}} \widetilde{EP}_n &= \frac{1}{2\pi} \left(\sum_{j=1}^n \Delta_j I_1 - I_2 \right) \\
&= \frac{1}{2\pi} \left[\sum_{j=1}^n \Delta_j (I_{11} - iI_{12}) - \pi \right] \\
&= \frac{1}{2\pi} \left[\sum_{j=1}^n \Delta_j (\pi e^{-Y_j} - i i \pi e^{-Y_j}) - \pi \right] \\
&= \frac{1}{2\pi} \left[\sum_{j=1}^n \Delta_j 2\pi e^{-Y_j} - \pi \right]
\end{aligned}$$

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

2.1.2 Baringhaus and Henze (1991)

For the full sample case the test proposed by Baringhaus and Henze (1991) is

$$B_{n,a} = n \int_0^\infty \left[(1+t) \psi'_n(t) + \psi_n(t) \right]^2 e^{-at} dt,$$

where $\psi'_n(t)$ is the derivative of $\psi_n(t)$. In the presence of random right censoring this test becomes

$$\begin{aligned} \tilde{B}_{n,a} &= n \int_0^\infty \left[(1+t) \tilde{\psi}'_n(t) + \tilde{\psi}_n(t) \right]^2 e^{-at} dt \\ &= n \int_0^\infty \left[(1+t) \left(- \sum_{j=1}^n \Delta_j Y_j e^{-tY_j} \right) + \sum_{j=1}^n \Delta_j e^{-tY_j} \right]^2 e^{-at} dt \\ &= n \int_0^\infty \left[(1+t)^2 \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k Y_j Y_k e^{-t(Y_j+Y_k+a)} - (1+t) \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k Y_j e^{-t(Y_j+Y_k+a)} \right. \\ &\quad \left. - (1+t) \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k Y_k e^{-t(Y_j+Y_k+a)} + \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k e^{-t(Y_j+Y_k+a)} \right] dt \\ &= n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\int_0^\infty (1+t)^2 Y_j Y_k e^{-t(Y_j+Y_k+a)} dt - \int_0^\infty (1+t) Y_j e^{-t(Y_j+Y_k+a)} dt \right. \\ &\quad \left. - \int_0^\infty (1+t) Y_k e^{-t(Y_j+Y_k+a)} dt + \int_0^\infty e^{-t(Y_j+Y_k+a)} dt \right] \\ &= n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{Y_j Y_k}{Y_j + Y_k + a} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^3} - \frac{Y_j}{Y_j + Y_k + a} \right. \\ &\quad \left. - \frac{Y_j}{(Y_j + Y_k + a)^2} - \frac{Y_k}{Y_j + Y_k + a} - \frac{Y_k}{(Y_j + Y_k + a)^2} + \frac{1}{Y_j + Y_k + a} \right] \\ &= n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{1 - Y_j - Y_k + Y_j Y_k}{Y_j + Y_k + a} - \frac{Y_j + Y_k - 2Y_j Y_k}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^3} \right] \\ &= n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{(1 - Y_j)(1 - Y_k)}{Y_j + Y_k + a} - \frac{Y_j + Y_k}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^3} \right]. \end{aligned}$$

Henze and Meintanis (2002) proposes the following test in the complete sample case:

$$L_{n,a} = n \int_0^\infty [\psi_n(t) - \psi(t)]^2 (1+t)^2 e^{-at} dt.$$

In the presence of random right sampling, this test can be amended as follows;

$$\begin{aligned}
\tilde{L}_{n,a} &= n \int_0^\infty \left[\tilde{\psi}_n(t) - \psi(t) \right]^2 (1+t)^2 e^{-at} dt \\
&= n \int_0^\infty \left[\sum_{j=1}^n \Delta_j e^{-tY_j} - \frac{1}{1+t} \right]^2 (1+t)^2 e^{-at} dt \\
&= n \int_0^\infty \left[(1+t)^2 \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k e^{-t(Y_j+Y_k+a)} - 2(1+t) \sum_{j=1}^n \Delta_j e^{-t(Y_j+a)} + e^{-at} \right] dt \\
&= n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \int_0^\infty (1+t)^2 e^{-t(Y_j+Y_k+a)} dt - 2n \sum_{j=1}^n \Delta_j \int_0^\infty (1+t) e^{-t(Y_j+a)} dt + n \int_0^\infty e^{-at} dt \\
&= n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{1 + (Y_j + Y_k + a + 1)^2}{(Y_j + Y_k + a)^3} \right] - 2n \sum_{j=1}^n \Delta_j \left[\frac{1 + Y_j + a}{(Y_j + a)^2} \right] + \frac{n}{a}.
\end{aligned}$$

CHAPTER 3

Article 2: New classes of tests for the Weibull distribution using Stein's method in the presence of random right censoring

The second article, *New classes of tests for the Weibull distribution using Stein's method in the presence of random right censoring*, has been accepted for publication in *Computational Statistics*. A summary of the guidelines to authors from the journal is now presented.

Manuscript	No page limit.
Title	The title should be concise and informative.
Abstract and keywords	An abstract of 150 to 250 words and 4 to 6 keywords which can be used for indexing purposes.
Tables	All tables are to be numbered using Arabic numerals. Tables should always be cited in text in consecutive numerical order. For each table, please supply a table caption explaining the components of the table.
References	Cite references in the text by name and year in parentheses. The list of references should only include works that are cited in the text and that have been published or accepted for publication. Reference list entries should be alphabetized by the last names of the first author of each work.
General formatting	A L ^A T _E X template is provided for submission.
Additional information	https://tinyurl.com/486kbupt



New classes of tests for the Weibull distribution using Stein's method in the presence of random right censoring

E. Bothma¹ · J. S. Allison¹ · I. J. H. Visagie¹ 

Received: 19 May 2021 / Accepted: 2 November 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

We develop two new classes of tests for the Weibull distribution based on Stein's method. The proposed tests are applied in the full sample case as well as in the presence of random right censoring. We investigate the finite sample performance of the new tests using a comprehensive Monte Carlo study. In both the absence and presence of censoring, it is found that the newly proposed classes of tests outperform competing tests against the majority of the distributions considered. In the cases where censoring is present we consider various censoring distributions. Some remarks on the asymptotic properties of the proposed tests are included. We present another result of independent interest; a test initially proposed for use with full samples is amended to allow for testing for the Weibull distribution in the presence of censoring. The techniques developed in the paper are illustrated using two practical examples.

Keywords Goodness-of-fit testing · Hypothesis testing · Random right censoring · Warp-speed bootstrap · Weibull distribution

1 Introduction

The Weibull distribution is often used in survival analysis as well as reliability theory, see e.g., Kalbfleisch and Prentice (2011). This flexible distribution is a popular model which allows for constant, increasing and decreasing hazard rates. The Weibull distribution is also frequently applied in various engineering fields, including electrical and industrial engineering to represent, for example, manufacturing times, see Jiang and Murthy (2011). As a result of its wide range of practical uses, a number of goodness-of-fit tests have been developed for the Weibull distribution; see e.g., Mann et al. (1973), Tiku and Singh (1981), Liao and Shimokawa (1999), Cabaña and Quiroz (2005) as well as Krit (2014).

I. J. H. Visagie
jaco.visagie@nwu.ac.za

¹ Subject Group Statistics, North-West University, Potchefstroom, South Africa

The papers listed above deal with testing for the Weibull distribution in the full sample case; i.e., where all lifetimes are observed. However, random right censoring often occurs in the fields mentioned above. For example, we may study the duration that antibodies remain detectable in a patient's blood after receiving a specific type of Covid-19 vaccine, i.e. the duration of the protection that the vaccine affords the recipient. When gathering the relevant data, we will likely not be able to measure this duration in all of the patients. For example, some may leave the study by emigrating to a different country while still having detectable antibodies. In this case, the exact time of interest is not observed. This situation is referred to as random right censoring, see e.g., Cox and Oakes (1984).

In the presence of censoring, testing the hypothesis that the distribution of the lifetimes is Weibull is complicated by the fact that an incomplete sample is observed. Balakrishnan et al. (2015) suggests a way to perform the required goodness-of-fit tests by transforming the censored sample to a complete sample. Another approach is to modify the test statistics used in the full sample case to account for the presence of censoring. Although fewer in number, tests for the Weibull distribution in the presence of random censoring are available in the literature. For example, Koziol and Green (1976) and Kim (2017) propose modified versions of the Cramér-von Mises test and the test proposed in Liao and Shimokawa (1999), respectively, for use with censored data.

Throughout this paper we are primarily interested in the situation where censoring is present; the results relating to the full sample case are treated as special cases obtained when all lifetimes are observed. Before proceeding some notation is introduced. Let $X_1, \dots, X_n$ be independent and identically distributed (i.i.d.) lifetime variables with continuous distribution function F and let $C_1, \dots, C_n$ be i.i.d. censoring variables with distribution function H , independent of $X_1, \dots, X_n$. We assume non-informative censoring throughout. Let

$$T_j = \min(X_j, C_j) \quad \text{and} \quad \delta_j = \begin{cases} 1, & \text{if } X_j \leq C_j, \\ 0, & \text{if } X_j > C_j. \end{cases}$$

Note that in the full sample case $T_j = X_j$ and $\delta_j = 1$ for $j = 1, \dots, n$.

Based on the observed pairs (T_j, δ_j) , $j = 1, \dots, n$ we wish to test the composite hypothesis

$$H_0 : X \sim \text{Weibull}(\lambda, \theta), \quad (1)$$

for some unknown $\lambda > 0$ and $\theta > 0$. Here $X \sim \text{Weibull}(\lambda, \theta)$ refers to a Weibull distributed random variable with distribution function

$$F(x) = 1 - e^{-(x/\lambda)^\theta}, \quad x > 0.$$

This hypothesis is to be tested against general alternatives. We will make use of maximum likelihood estimation to estimate λ and θ . The log-likelihood of the Weibull distribution is

$$\mathcal{L}(\theta, \lambda | X_1, \dots, X_n) = d \log(\theta) - d\theta \log(\lambda) + (\theta - 1) \sum_{j=1}^n \delta_j \log(X_j) - \lambda^{-\theta} \sum_{j=1}^n X_j^\theta,$$

where $d = \sum_{j=1}^n \delta_j$. In the full sample case $d = n$. No closed form formulae for the maximum likelihood estimates $\hat{\lambda}$ and $\hat{\theta}$ exist, meaning that numerical optimisation techniques are required to arrive at parameter estimates.

Since the Weibull distribution has a shape parameter the first step for many goodness-of-fit tests for the Weibull distribution is to transform the data. If $X \sim \text{Weibull}(\lambda, \theta)$ then a frequently used transformation is $\log(X)$, which results in a random variable that is type I extreme value distributed with parameters $\log(\lambda)$ and $1/\theta$. The resulting transformed random variable is part of a location scale family, which is a desirable result when performing goodness-of-fit testing. We therefore have that if $X \sim \text{Weibull}(\lambda, \theta)$, then $X^{(t)} = \theta(\log(X) - \log(\lambda))$ follows a standard type I extreme value distribution with distribution function

$$G(x) = 1 - e^{-e^x}, \quad -\infty < x < \infty.$$

We denote a random variable with this distribution function by $EV(0, 1)$. As a result the hypothesis in (1) holds, if, and only if, $X^{(t)} \sim EV(0, 1)$. All of the test statistics considered make use of the transformed observed values

$$Y_j = \hat{\theta} \left[\log(T_j) - \log(\hat{\lambda}) \right], \quad (2)$$

with $\hat{\lambda}$ and $\hat{\theta}$ the maximum likelihood estimates of the Weibull distribution. Let

$$X_j^{(t)} = \hat{\theta} \left[\log(X_j) - \log(\hat{\lambda}) \right].$$

If $X_1, \dots, X_n$ are realised from a $\text{Weibull}(\lambda, \theta)$ distribution, then $X_1^{(t)}, \dots, X_n^{(t)}$ will approximately follow an $EV(0, 1)$ distribution see e.g. Kotz and Nadarajah (2000). The resulting random variables are no longer independent. However, several properties of the classical testing procedures remain unaffected when performing this type of transformation; the interested reader is referred to Baringhaus and Henze (1991) as well as Gupta and Richards (1997) and the references therein for more details. The tests employed below are based on discrepancy measures between the calculated values of $Y_1, \dots, Y_n$ and the standard type I extreme value distribution. The order statistics of $Y_1, \dots, Y_n$ are denoted by $Y_{(1)} < \dots < Y_{(n)}$, while $\delta_{(j)}$ represents the indicator variable corresponding to $Y_{(j)}$.

The remainder of the paper is structured as follows. In Sect. 2, we propose two new classes of tests for the Weibull distribution for both the full sample and censored case. We also modify the test proposed in Krit (2014) to accommodate random right censoring. In the presence of censoring the null distribution of all the test statistics considered depend on the unknown censoring distribution, we therefore propose a parametric bootstrap procedure in Sect. 2 in order to compute critical values. Section 3 presents the results of a Monte Carlo study where the empirical powers of the newly

proposed classes of tests as well as the newly modified test are compared to those of existing tests. The paper concludes in Sect. 4 with two practical applications; one concerning the survival times of patients diagnosed with a certain type of leukemia (no censoring is present in these data) and the other relates to observed leukemia remission times (in the presence of censoring). Some avenues for future research are also discussed.

2 Proposed test statistics

Our newly proposed classes of tests are based on the following theorem, which characterises the standard type I extreme value distribution.

Theorem 1 *Let W be a random variable with absolutely continuous density and assume that $E[e^W] < \infty$. In this case $W \sim EV(0, 1)$, if, and only if,*

$$E \left[\left(it + 1 - e^W \right) e^{itW} \right] = 0, \quad \forall t \in R,$$

with $i = \sqrt{-1}$.

Proof The 'if' part of the theorem can easily be shown using direct calculation. The 'only if' part is shown below.

Assume $E \left[\left(it + 1 - e^W \right) e^{itW} \right] = 0$. From Fourier analysis we have that the Fourier transform of $f'(w)$ is $e^{itw} f'(w) = -it E[e^{itW}]$. This implies that

$$\begin{aligned} 0 &= E \left[\left(it + 1 - e^W \right) e^{itW} \right] \\ &= \int_{-\infty}^{\infty} \left(it + 1 - e^w \right) e^{itw} f(w) dw \\ &= \int_{-\infty}^{\infty} \left[-f'(w) + (1 - e^w) f(w) \right] e^{itw} dw, \end{aligned}$$

which is the Fourier transform of $h(w) := -f'(w) + (1 - e^w) f(w)$. The fact that the Fourier transform of h is 0 implies that $h(w) = 0$ for all $w \in R$. Thus, f satisfies the following differential equation $f'(w) - (1 - e^w) f(w) = 0$. Using separation of variables we have

$$\frac{f'(w)}{f(w)} = 1 - e^w \implies \log(f(w)) = w - e^w + c \implies f(w) = e^{w - e^w},$$

where the last step follows from the fact that f must be a density function and hence integrate to 1. This completes the proof. $\square$

Let $w(t)$ be a non-negative, symmetric weight function. From Theorem 1, we have that

$$\begin{aligned}\eta &= \int_{-\infty}^{\infty} E \left[(it + 1 - e^Y) e^{itY} \right] w(t) dt \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (it + 1 - e^y) e^{ity} dG(y) w(t) dt\end{aligned}\quad (3)$$

equals 0 if $Y \sim EV(0, 1)$. Note that the inclusion of the weight function, w , above is required to ensure that η is finite. Clearly η will be unknown because G is unknown. However, G can be estimated by the Kaplan-Meier estimator, G_n , of the distribution function given by

$$1 - G_n(t) = \begin{cases} 1, & t \leq Y_{(1)} \\ \prod_{j=1}^{k-1} \left(\frac{n-j}{n-j+1} \right)^{\delta_{(j)}}, & Y_{(k-1)} < t \leq Y_{(k)}, \quad k = 2, \dots, n. \\ \prod_{j=1}^n \left(\frac{n-j}{n-j+1} \right)^{\delta_{(j)}}, & t > Y_{(n)}. \end{cases}$$

More details about this estimator can be found in Kaplan and Meier (1958), Efron (1967) as well as Breslow and Crowley (1974). In the full sample case this estimator reduces to the standard empirical distribution function, $G_n(X_{(j)}) = j/n$.

Let Δ_j denote the size of the jump in $G_n(T_{(j)})$;

$$\Delta_j = G_n(T_{(j)}) - \lim_{t \uparrow T_{(j)}} G_n(t), \quad j = 1, \dots, n.$$

Simple calculable expressions for the Δ_j 's are

$$\begin{aligned}\Delta_1 &= \frac{\delta_{(1)}}{n}, \quad \Delta_n = \prod_{j=1}^{n-1} \left(\frac{n-j}{n-j+1} \right)^{\delta_{(j)}} \quad \text{and} \\ \Delta_j &= \prod_{k=1}^{j-1} \left(\frac{n-k}{n-k+1} \right)^{\delta_{(k)}} - \prod_{k=1}^j \left(\frac{n-k}{n-k+1} \right)^{\delta_{(k)}} \\ &= \frac{\delta_{(j)}}{n-j+1} \prod_{k=1}^{j-1} \left(\frac{n-k}{n-k+1} \right)^{\delta_{(k)}}, \quad j = 2, \dots, n-1.\end{aligned}$$

In the full sample case $\Delta_j = 1/n$, $j = 1, \dots, n$.

Estimating G by G_n in (3), we propose the test statistic

$$S_{n,a} = n \int_{-\infty}^{\infty} \left| \sum_{j=1}^n \Delta_j \left[ite^{itY_j} + (1 - e^{Y_j})e^{itY_j} \right] \right|^2 w_a(t) dt, \quad (4)$$

where $w_a(t)$ is a weight function containing a user-defined tuning parameter $a > 0$. The null hypothesis in (1) is rejected for large values of $S_{n,a}$.

Straightforward algebra shows that, if $w_a(t) = e^{-at^2}$, then the test statistic simplifies to

$$S_{n,a}^{(1)} = n \sqrt{\frac{\pi}{a}} \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k e^{-(Y_j - Y_k)^2 / 4a} \left\{ -\frac{1}{4a^2} \left((Y_j - Y_k)^2 - 2a \right) + 2 \left(1 - e^{Y_j} \right) \left(\frac{1}{2a} \right) (Y_j - Y_k) + \left(1 - e^{Y_j} \right) \left(1 - e^{Y_k} \right) \right\},$$

and if $w_a(t) = e^{-a|t|}$ the test statistic has the following easily calculable form

$$S_{n,a}^{(2)} = n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left\{ \frac{-4a \left(3(Y_j - Y_k)^2 - a^2 \right)}{\left((Y_j - Y_k)^2 + a^2 \right)^3} + \frac{8a(Y_j - Y_k) \left(1 - e^{Y_j} \right)}{\left((Y_j - Y_k)^2 + a^2 \right)^2} + \frac{2a \left(1 - e^{Y_j} \right) \left(1 - e^{Y_k} \right)}{(Y_j - Y_k)^2 + a^2} \right\}.$$

New goodness-of-fit tests containing a tuning parameter are often accompanied by a recommended value of this parameter; this choice is typically based on the finite sample power performance of the test. Another approach which may be used is to choose the value of the tuning parameter data-dependently; see e.g., Allison and Santana (2015). In this paper, we opt to use the values recommended in the literature for tests containing a tuning parameter.

The weight functions specified above correspond to scaled Gaussian and Laplace kernels. These weight functions are popular choices found in the goodness-of-fit literature; see e.g., Meintanis and Iliopoulos (2003), Allison et al. (2017), Betsch and Ebner (2018), Betsch and Ebner (2019) as well as Henze and Visagie (2020). The popularity of these weight functions is, at least in part, due to the fact that their inclusion typically results in simple calculable forms for L^2 -type statistics which do not require numerical integration. As an alternative to the weight functions used, one may employ a symmetric uniform kernel as a weight function; see e.g., Fernández et al. (2008). However, in the mentioned paper, the authors found that the computational time required for the test statistic obtained using the symmetric uniform kernel was substantial. This, coupled with the simple computational forms obtained using the weight functions defined above and the favourable power performance discussed in Sect. 3, motivated us to restrict our attention to the scaled Gaussian and Laplace kernels.

Although we do not derive the asymptotic results related to the proposed classes of test statistics we include some remarks in this regard. $S_{n,a}$ is a characteristic function based weighted L^2 -type statistic. The asymptotic properties of this class of statistics, in the complete sample case, are studied in detail in Feuerverger and Mureika (1977), while more recent references include Baringhaus and Henze (1988), Klar and Meintanis (2005) as well as Baringhaus et al. (2017). A convenient setting for the derivation of the asymptotic properties of these tests is the separable Hilbert space of square integrable functions. Typically, the asymptotic null distribution of $S_{n,a}$ corresponds to that of $\int_{-\infty}^{\infty} |Z(t)|^2 w_a(t) dt =: S_a$, where $Z(\cdot)$ is a zero-mean Gaussian process. The

distribution of S_a is the same as that of $\sum_{j=1}^{\infty} \lambda_j U_j$, where U_j are i.i.d. chi-squared random variables with parameter 1 and where λ_j are eigenvalues of an integral operator (see e.g., Allison et al. (2021)). These tests are consistent against a large class of fixed alternative distributions. Additionally, these tests are frequently consistent against contiguous alternatives converging to the null at a rate of $n^{-1/2}$.

In the case of random censoring, very little asymptotic results are available in the literature for test statistics of this type. Very recently, advances have been made in this regard; see e.g., Cuparić and Milošević (2021), in which a test for exponentiality is considered based on so-called inverse probability censoring weights, where the authors derive the asymptotic properties of the given test. In addition Fernández and Rivera (2020) studied Kaplan-Meier U - and V -statistics in order to derive some asymptotic results relating to the lifetime distribution. Some of these results may be helpful in deriving the asymptotic properties of the tests proposed in this paper in future research.

The null distribution of each of the test statistics considered depends on the unknown censoring distribution, even in the case of a simple hypothesis, see D'Agostino and Stephens (1986). Since we will not assume any known form of the censoring distribution, we propose the following parametric bootstrap algorithm to estimate the critical values of the tests.

1. Based on the pairs (T_j, δ_j) , $j = 1, \dots, n$ estimate θ and λ by $\hat{\theta}$ and $\hat{\lambda}$, respectively, using maximum likelihood estimation.
2. Transform T_j to Y_j using the transformation in (2) for $j = 1, \dots, n$.
3. Calculate the test statistic, say $W_n := W(Y_1, \dots, Y_n; \delta_1, \dots, \delta_n)$.
4. Obtain a parametric bootstrap sample $X_1^*, \dots, X_n^*$ by sampling from a Weibull distribution with parameters $\hat{\theta}$ and $\hat{\lambda}$.
5. Obtain a non-parametric bootstrap sample, $C_1^*, \dots, C_n^*$, by sampling from the Kaplan-Meier estimate of the distribution of $C_j^{(t)} = \hat{\theta} [\log(C_j) - \log(\hat{\lambda})]$.
6. Set

$$T_j^* = \min(X_j^*, C_j^*) \text{ and } \delta_j^* = \begin{cases} 1, & \text{if } X_j^* \leq C_j^* \\ 0, & \text{if } X_j^* > C_j^*. \end{cases}$$

7. Calculate $\hat{\theta}^*$ and $\hat{\lambda}^*$ based on (T_j^*, δ_j^*) , $j = 1, \dots, n$.
8. Obtain $Y_j^* = \hat{\theta}^* [\log(T_j^*) - \log(\hat{\lambda}^*)]$, $j = 1, \dots, n$.
9. Based on the pairs (Y_j^*, δ_j^*) , $j = 1, \dots, n$, calculate the value of the test statistic, say $W_n^* := W(Y_1^*, \dots, Y_n^*; \delta_1^*, \dots, \delta_n^*)$.
10. Repeat steps 4-9 B times to obtain $W_1^*, \dots, W_B^*$. Obtain the order statistics, $W_{(1)}^* \leq \dots \leq W_{(B)}^*$. The estimated critical value is then $\hat{c}_n(\alpha) = W_{[B(1-\alpha)]}^*$ where $[c]$ denotes the floor of c .

The algorithm provided above is quite general and can easily be amended in order to test for any lifetime distribution in the presence of random censoring. In the absence of censoring there is no need to implement this algorithm; in this case, the critical values can be obtained via Monte Carlo simulation by sampling from any Weibull distribution and effecting the transformation discussed above.

3 Numerical results

In this section, we compare the power performances of the newly proposed tests to those of existing tests via a Monte Carlo simulation study. The existing tests used include the classical Kolmogorov-Smirnov (KS_n) and Cramér-von Mises (CM_n) tests. These tests have been modified for use with censored data, see Koziol and Green (1976). The test introduced in Liao and Shimokawa (1999) is considered in the case of full samples. A modification making this test suitable for use with censored data is proposed in Kim (2017); we denote the test statistic by LS_n in both the full sample and censored cases. The calculable forms of the test statistics mentioned above are

$$\begin{aligned}
 KS_n &= \max \left[\max_{1 \leq j \leq n} \left\{ G_n(Y_{(j)}) - \left(1 - e^{-e^{Y_j}} \right) \right\}, \right. \\
 &\quad \left. \max_{1 \leq j \leq n} \left\{ \left(1 - e^{-e^{Y_j}} \right) - G_n^-(Y_{(j)}) \right\} \right], \\
 CM_n &= \frac{n}{3} + n \sum_{j=1}^{d+1} \left\{ G_n \left(X_{j-1}^{(t)} \right) \left(X_j^{(t)} - X_{j-1}^{(t)} \right) \right. \\
 &\quad \left. \times \left[G_n \left(X_{j-1}^{(t)} \right) - \left(X_j^{(t)} + X_{j-1}^{(t)} \right) \right] \right\}, \\
 LS_n &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \left(\frac{\max \left[j/n - G_n(Y_j), G_n(Y_j) - (j-1)/n \right]}{\sqrt{G_n(Y_j) \left[1 - G_n(Y_j) \right]}} \right).
 \end{aligned}$$

Krit (2014) proposes a test for the Weibull distribution in the full sample case. This test compares the empirical Laplace transform of the random variables resulting from the transformation in (2) to the Laplace transform of an $EV(0, 1)$ random variable; $\psi(t) = \Gamma(1-t)$ for $t < 1$. Let ψ_n be the empirical Laplace transform of the transformed observations, obtained using the Kaplan-Meier estimate of the distribution function;

$$\psi_n(t) = \int_{-\infty}^{\infty} e^{-tx} dG_n(x) = \sum_{j=1}^n \Delta_j e^{-tY_j}. \quad (5)$$

The resulting test statistic is

$$KR_n^* = n \int_I [\psi_n(t) - \Gamma(1-t)]^2 w_a(t) dt, \quad (6)$$

where $w_a(t) = e^{at-e^{at}}$ is a weight function, a a user-specified tuning parameter and I some interval. Based on numerical considerations, Krit (2014) suggests that $I = (-1, 0]$ should be used. The quantity in (6) can be approximated by a Riemann sum;

$$K R_n = n \sum_{k=-m}^{-1} \left[\sum_{j=1}^n \Delta_j e^{-Y_j k/m} - \Gamma(1 - k/m) \right]^2 e^{ak/m - e^{ak/m}},$$

where m is the number of points at which the integrand is evaluated. In the numerical results shown below, we use $a = -5$ and $m = 100$ as recommended in Krit (2014). In the full sample case, G_n , in (5), is taken to be the empirical distribution function. Upon setting $\Delta_j = 1/n$ we obtain the test statistic in Krit (2014).

For each of the tests considered above, the null hypothesis in (1) is rejected for large values of the test statistics.

3.1 Simulation setting

In the numerical results presented below, we use a nominal significance level of 10% throughout. Empirical powers are presented for sample sizes $n = 50$ and $n = 100$. The empirical powers for complete and censored samples are reported; censoring proportions of 10% and 20% are included. For each lifetime distribution considered, we report the powers obtained using three different censoring distributions. The first censoring distribution used is the exponential distribution, the parameter of which is chosen so as to obtain the specified level of censoring. The second censoring distribution used is the uniform distribution with support $(0, m)$; again, m is chosen such that the required censoring level is achieved. The final censoring distribution used is the Koziol-Green model proposed in Koziol and Green (1976). Denote the survival function of a given lifetime distribution by S . In this case, the Koziol-Green censoring distribution, indexed by β , has survival function $S^\beta(t)$. It can be shown that the censoring proportion is $\beta/(1 + \beta)$. We chose the value of β so as to ensure that the required level of censoring is achieved. The alternative lifetime distributions considered are listed in Table 1. Note that each of the alternatives considered have the same support as that of the Weibull distribution with the exception of the beta distribution which is restricted to the unit interval $(0, 1)$ (Tables 10, 11, 12 and 13).

The obtained empirical powers are presented in Tables 2, 3, 4, 5, 6, 7, 8 and 9. These tables report the percentages of 50 000 independent Monte Carlo samples that

Table 1 Density functions of the alternative distributions

Alternative	Density	Notation	Support
Weibull	$\theta x^{\theta-1} \exp(-x^\theta)$	$W(\theta)$	$(0, \infty)$
Gamma	$(\Gamma(\theta))^{-1} x^{\theta-1} \exp(-x)$	$\Gamma(\theta)$	$(0, \infty)$
Lognormal	$(\theta x \sqrt{2\pi})^{-1} \exp\left(-\log^2(x) (2\theta^2)^{-1}\right)$	$LN(\theta)$	$(0, \infty)$
Chi square	$(2^{\theta/2} \Gamma(\theta/2))^{-1} x^{\theta/2-1} \exp(-x/2)$	$\chi^2(\theta)$	$(0, \infty)$
Beta	$x^{\alpha-1} (1-x)^{\theta-1} \Gamma(\alpha+\theta) (\Gamma(\alpha)\Gamma(\theta))^{-1}$	$\beta(\alpha, \theta)$	$(0, 1)$
Lindley	$\frac{\theta^2}{\theta+1} (1+x) \exp(-\theta x)$	$Lind(\theta)$	$(0, \infty)$

Table 2 Estimated powers for the full sample case where $n = 50$

F	KS_n	CM_n	LS_n	KR_n	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,5}^{(1)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$	$S_{n,5}^{(2)}$
$W(0.5)$	10	10	10	10	10	10	10	10	10	10
$W(1)$	10	10	10	10	10	10	10	10	10	10
$W(1.5)$	10	11	10	11	11	10	10	11	11	10
$W(2)$	10	10	9	10	10	10	10	10	10	10
$\Gamma(2)$	13	15	17	14	17	17	16	14	16	17
$\Gamma(3)$	17	20	25	19	24	25	24	18	22	25
$LN(0.5)$	50	62	70	63	72	76	76	51	68	76
$LN(1)$	50	61	70	63	73	77	77	51	69	77
$\chi^2(8)$	21	24	31	23	30	32	31	20	28	31
$\chi^2(10)$	23	28	35	27	34	36	35	23	31	36
$\beta(1, 1)$	69	81	92	90	85	87	88	80	86	89
$\beta(0.5, 1)$	70	81	93	91	86	88	88	80	86	89
$Lind(0.5)$	11	11	12	13	11	13	14	10	11	13
$Lind(2)$	10	10	11	11	11	11	11	10	10	11

Table 3 Heatmap of the estimated powers for the full sample case where $n = 50$

F	KS_n	CM_n	LS_n	KR_n	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,5}^{(1)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$	$S_{n,5}^{(2)}$
$W(0.5)$										
$W(1)$										
$W(1.5)$										
$W(2)$										
$\Gamma(2)$										
$\Gamma(3)$										
$LN(0.5)$										
$LN(1)$										
$\chi^2(8)$										
$\chi^2(10)$										
$\beta(1, 1)$										
$\beta(0.5, 1)$										
$Lind(0.5)$										
$Lind(2)$										

lead to the rejection of the null hypothesis, rounded to the nearest integer. For ease of comparison, the highest power in each line is printed in bold. Tables 2 and 4 contain the results relating to full samples. In order to ease visual comparison of the results obtained, we include so called "heatmaps", see Döring and Cramer (2019), of these results in Tables 3 and 5. For each test considered, Tables 6, 7, 8 and 9 show three empirical powers against each lifetime distribution, corresponding to the three different censoring distributions used. In each case, the results for the exponential, uniform and Koziol-Green models are shown in the first, second and third lines, respectively.

In order to reduce the computational cost associated with the numerical powers a warp-speed bootstrap procedure, see Giacomini et al. (2013), is employed. This methodology has been employed by a number of authors in the literature to compare

Table 4 Estimated powers for the full sample case where $n = 100$

F	KS_n	CM_n	LS_n	KR_n	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,5}^{(1)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$	$S_{n,5}^{(2)}$
$W(0.5)$	10	10	10	10	10	10	10	9	10	10
$W(1)$	10	10	10	10	10	10	10	10	10	10
$W(1.5)$	10	10	10	10	10	10	10	10	10	10
$W(2)$	10	9	10	10	10	10	10	9	10	10
$\Gamma(2)$	17	19	26	19	24	26	27	17	22	26
$\Gamma(3)$	26	31	41	31	39	43	45	24	35	44
$LN(0.5)$	77	88	94	90	95	97	98	78	93	97
$LN(1)$	78	88	93	90	95	97	98	78	93	97
$\chi^2(8)$	32	39	51	41	49	54	56	31	45	55
$\chi^2(10)$	36	44	58	46	56	62	64	34	51	63
$\beta(1, 1)$	94	98	100	100	99	99	99	99	99	100
$\beta(0.5, 1)$	94	98	100	100	99	99	99	99	99	99
$Lind(0.5)$	12	12	13	13	12	14	16	11	12	15
$Lind(2)$	11	11	11	11	11	11	12	10	11	12

Table 5 Heatmap of the estimated powers for the full sample case where $n = 100$

F	KS_n	CM_n	LS_n	KR_n	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,5}^{(1)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$	$S_{n,5}^{(2)}$
$W(0.5)$										
$W(1)$										
$W(1.5)$										
$W(2)$										
$\Gamma(2)$										
$\Gamma(3)$										
$LN(0.5)$										
$LN(1)$										
$\chi^2(8)$										
$\chi^2(10)$										
$\beta(1, 1)$										
$\beta(0.5, 1)$										
$Lind(0.5)$										
$Lind(2)$										

Monte Carlo performances; see e.g., Meintanis et al. (2018), Allison et al. (2019) as well as Mijburgh and Visagie (2020). The bootstrap algorithm in Sect. 2 is implemented to calculate the critical values used to obtain the results in Tables 6, 7, 8 and 9.

For $S_{n,a}^{(1)}$ and $S_{n,a}^{(2)}$, we include numerical powers in the cases where a is set to 1, 2 and 5 in the full sample case and set to 0.75, 1 and 2 in the presence of censoring. The difference between these two sets of choices is due to the fact that the newly proposed tests exhibits higher powers in the presence of censoring if slightly smaller values of a are used. All calculations are performed in R, see [44]. The *LindleyR* package is used to generate samples from censored distributions, see Mazucheli et al. (2016). Parameter estimation is performed using the *parmsurvfit* package, see Jacobson et al. (2018), while the tables are produced using the *Stargazer* package, see Hlavac (2018).

Table 6 Estimated powers for 10% censoring for a sample size of $n = 50$ with three different censoring distributions

F	KS_n	CM_n	LS_n	KR_n	$S_{n,0.75}^{(1)}$	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,0.75}^{(2)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$
$W(0.5)$	9	9	9	7	8	8	8	8	8	8
	9	9	9	4	9	9	9	9	8	9
	10	9	10	9	8	8	9	9	9	8
$W(1)$	9	9	9	9	8	8	10	9	9	8
	9	9	9	8	8	8	10	9	9	8
	9	9	9	9	8	8	9	9	9	8
$W(1.5)$	10	10	10	9	9	9	9	9	9	9
	10	10	10	8	9	9	9	10	10	9
	10	10	10	9	8	8	10	9	9	8
$W(2)$	9	10	10	8	9	9	9	9	9	9
	10	10	10	9	9	9	10	10	9	9
	9	10	10	9	8	9	9	9	9	8
$\Gamma(2)$	13	14	13	21	15	15	15	12	13	14
	13	14	13	21	15	15	15	12	13	14
	12	14	12	21	14	14	15	12	12	14
$\Gamma(3)$	17	19	18	29	21	22	23	15	17	21
	16	19	17	28	20	21	21	14	16	20
	16	19	17	29	20	20	20	14	16	19
$LN(0.5)$	46	56	57	64	64	67	69	36	46	63
	45	56	57	63	64	66	68	36	46	62
	43	54	56	64	61	63	63	35	44	59
$LN(1)$	43	54	53	56	60	62	58	34	43	58
	42	53	49	48	58	59	53	31	41	56
	44	54	56	64	61	63	63	35	44	59
$\chi^2(8)$	19	23	22	34	26	27	27	16	20	25

3.2 Simulation results

First, we consider the results associated with the full sample case, given in Tables 2 and 4, together with the heatmaps shown in Tables 3 and 5. All of the tests considered attain the nominal size for both sample sizes used. The tests associated with the highest powers are $S_{n,2}^{(1)}$ and $S_{n,5}^{(1)}$, although LS_n and $S_{n,5}^{(2)}$ also performs well. When analysing complete samples, we recommend using $S_{n,2}^{(1)}$ or $S_{n,5}^{(1)}$.

We now turn our attention to the powers achieved in the presence of censoring. The size of the tests are maintained closely for all sample sizes for censoring proportions of 10% and 20%, with the single exception of $KR_{m,a}$ in the case of small sample sizes. As expected, the powers generally increase with sample size and decrease marginally as the censoring proportion increases.

Table 6 continued

F	KS_n	CM_n	LS_n	KR_n	$S_{n,0.75}^{(1)}$	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,0.75}^{(2)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$
$\chi^2(10)$	19	22	21	33	26	26	27	16	19	25
	19	22	21	34	24	25	24	16	18	23
	21	25	25	38	29	31	32	17	21	28
	21	25	25	38	29	31	32	18	21	28
	21	25	24	37	27	28	27	18	20	26
$\beta(1, 1)$	64	75	83	1	78	80	82	68	73	80
	63	75	83	1	78	79	81	68	72	80
	61	70	79	0	55	50	31	62	63	47
$\beta(0.5, 1)$	61	72	79	0	74	75	76	66	70	76
	60	72	79	0	73	74	74	65	69	74
	61	69	79	0	54	49	28	62	62	46
$Lind(0.5)$	10	11	12	8	10	10	11	9	9	9
	10	11	12	8	10	10	12	9	9	10
	10	10	12	8	9	10	11	9	9	9
$Lind(2)$	10	10	10	7	9	9	10	9	9	8
	10	10	10	6	9	9	10	9	9	9
	10	10	11	7	9	9	9	9	9	8

Table 7 Estimated powers for 20% censoring for a sample size of $n = 50$ with three different censoring distributions

F	KS_n	CM_n	LS_n	KR_n	$S_{n,0.75}^{(1)}$	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,0.75}^{(2)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$
$W(0.5)$	7	6	8	6	7	8	7	7	7	7
	8	6	8	6	8	8	8	8	8	8
	9	9	9	6	7	7	8	9	8	7
$W(1)$	8	8	9	6	7	7	7	8	8	7
	8	8	8	7	8	8	8	8	8	8
	8	8	9	6	7	7	7	8	8	7
$W(1.5)$	9	9	10	10	8	8	9	9	8	8
	9	9	9	9	8	8	9	9	8	8
	8	8	9	6	7	7	7	9	8	7
$W(2)$	9	9	10	8	8	8	9	9	9	8
	9	9	9	10	8	8	9	9	9	8
	8	8	9	7	7	7	8	8	8	7
$\Gamma(2)$	12	13	12	12	13	13	13	11	12	12
	11	12	11	11	12	13	14	10	11	13
	11	12	11	9	11	11	12	11	11	11
$\Gamma(3)$	15	17	16	12	19	19	19	14	15	18
	15	17	16	15	18	18	18	13	14	18
	13	16	15	10	15	15	14	13	14	15

Table 7 continued

F	KS_n	CM_n	LS_n	KR_n	$S_{n,0.75}^{(1)}$	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,0.75}^{(2)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$
$LN(0.5)$	40	50	52	24	57	59	58	32	41	55
	40	50	50	20	56	57	53	31	40	53
	35	45	48	16	49	49	42	29	37	46
$LN(1)$	34	45	44	17	45	45	37	27	34	42
	32	42	34	21	37	37	33	21	28	35
	35	45	48	16	50	50	42	29	36	46
$\chi^2(8)$	17	20	20	12	23	24	23	15	18	22
	17	20	19	13	22	23	22	15	17	22
	15	18	18	12	19	19	17	14	16	18
$\chi^2(10)$	20	23	23	14	26	27	27	16	19	26
	19	22	22	14	26	27	26	16	19	25
	17	21	20	12	21	21	19	16	18	20
$\beta(1, 1)$	55	67	74	34	70	71	73	59	64	71
	55	66	72	32	68	69	69	58	63	68
	51	51	64	0	19	9	0	49	44	6
$\beta(0.5, 1)$	50	59	64	19	56	54	47	52	55	53
	48	54	61	3	39	32	12	48	48	29
	51	51	64	0	19	9	0	48	43	6
$Lind(0.5)$	9	10	12	9	9	9	9	9	8	8
	9	9	11	6	9	9	8	8	8	8
	9	9	12	6	8	8	8	9	8	8
$Lind(2)$	8	8	10	6	7	7	7	8	8	7
	8	8	9	6	8	8	7	8	8	7
	9	9	10	6	7	7	7	8	8	7

Table 8 Estimated powers for 10% censoring for a sample size of $n = 100$ with three different censoring distributions

F	KS_n	CM_n	LS_n	KR_n	$S_{n,0.75}^{(1)}$	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,0.75}^{(2)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$
$W(0.5)$	10	10	9	6	7	7	7	8	8	7
	9	9	9	7	8	8	8	8	8	8
	10	10	10	10	9	8	9	10	9	8
$W(1)$	10	10	10	10	8	8	9	9	9	8
	10	10	10	11	8	8	10	9	9	8
	10	10	10	10	9	9	9	10	9	9
$W(1.5)$	10	10	10	9	9	9	9	10	10	9
	10	10	10	9	9	9	9	9	9	8
	10	10	10	11	8	8	9	9	9	8
$W(2)$	10	10	10	8	9	9	9	10	10	9

Table 8 continued

F	KS_n	CM_n	LS_n	KR_n	$S_{n,0.75}^{(1)}$	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,0.75}^{(2)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$
$\Gamma(2)$	10	10	10	9	9	9	10	10	10	9
	10	10	10	11	9	9	9	9	9	8
	17	19	19	18	21	22	23	14	16	20
	17	19	18	18	20	21	22	14	16	20
$\Gamma(3)$	17	18	18	15	19	20	21	14	16	19
	25	29	30	30	34	36	38	19	23	33
	24	29	29	29	33	35	37	19	23	32
	23	28	29	23	32	33	34	19	23	31
$LN(0.5)$	74	85	87	84	91	92	95	61	74	90
	72	84	87	83	90	92	94	61	74	90
	71	83	86	79	89	91	93	60	72	89
$LN(1)$	71	83	85	42	88	90	90	58	71	87
	70	82	83	26	87	88	87	54	68	85
	71	83	86	78	89	91	93	60	73	88
$\chi^2(8)$	30	36	38	38	43	45	49	22	28	41
	30	36	38	37	42	45	48	23	28	41
	29	34	36	28	39	42	43	22	27	38
$\chi^2(10)$	34	41	43	44	48	51	56	26	33	47
	34	41	44	44	49	52	56	26	33	48
	33	40	42	33	46	48	50	25	31	45
$\beta(1, 1)$	91	97	99	98	98	98	99	96	98	99
	91	97	99	98	98	98	98	96	97	99
	90	95	97	0	84	79	50	91	91	77
$\beta(0.5, 1)$	90	96	98	96	97	98	98	95	97	98
	90	96	98	95	97	98	98	95	97	98
	91	95	98	0	83	77	45	91	90	74
$Lind(0.5)$	11	11	13	10	11	11	13	10	10	10
	11	11	13	11	10	11	12	10	10	10
	11	11	13	11	10	11	12	10	10	10
$Lind(2)$	10	10	11	10	9	9	10	9	9	9
	10	10	11	12	9	9	10	9	9	8
	10	10	11	10	9	9	10	9	9	9

Comparing the results associated with a sample size of 50, we see that $S_{n,1}^{(1)}$, KR_n and LS_n generally tend to provide the highest powers. However, it should be noted that KR_n achieves very low power against certain alternatives; notably against the beta distributions considered when the censoring distribution is the Kozoil-Green model. When considering the empirical powers associated with samples of size 100, $S_{n,2}^{(1)}$ exhibits the highest powers, followed by $S_{n,1}^{(1)}$ and LS_n .

Table 9 Estimated powers for 20% censoring for a sample size of $n = 100$ with three different censoring distributions

F	KS_n	CM_n	LS_n	KR_n	$S_{n,0.75}^{(1)}$	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,0.75}^{(2)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$
$W(0.5)$	8	7	9	6	7	7	7	7	7	7
	9	7	9	6	8	8	8	7	7	8
	9	9	9	6	7	7	7	8	8	7
$W(1)$	9	9	9	6	7	7	7	8	8	7
	8	8	8	7	7	8	8	7	7	8
	9	9	9	6	7	7	7	8	8	7
$W(1.5)$	9	10	10	10	8	8	9	9	9	8
	9	10	9	12	9	9	10	9	9	9
	9	9	9	11	8	8	9	9	8	8
$W(2)$	10	10	10	8	9	9	9	9	9	8
	10	10	10	10	8	8	9	9	9	8
	9	9	9	6	7	7	7	9	8	7
$\Gamma(2)$	15	17	17	12	18	18	18	14	15	17
	15	17	15	12	15	16	16	12	13	15
	15	16	16	10	15	15	14	13	13	14
$\Gamma(3)$	23	27	28	17	30	32	32	18	22	29
	22	26	26	15	28	29	29	17	20	27
	20	24	25	12	23	24	21	17	19	22
$LN(0.5)$	67	80	83	57	86	88	90	55	68	85
	65	79	82	37	85	87	87	53	66	83
	61	76	80	18	81	82	79	51	63	78
$LN(1)$	61	75	77	20	76	76	67	47	59	72
	43	50	49	26	45	45	37	30	37	42
	44	52	54	38	55	56	56	39	45	54
$\chi^2(8)$	27	32	34	22	38	40	42	21	26	37
	27	32	34	18	37	38	39	21	26	36
	24	30	31	13	30	31	27	20	23	28
$\chi^2(10)$	31	38	40	26	44	47	49	24	30	43
	23	26	27	22	29	30	31	19	22	28
	21	24	25	19	26	27	27	18	20	26
$\beta(1, 1)$	87	95	96	84	96	97	97	93	95	97
	87	94	96	81	96	96	97	92	94	97
	85	85	91	0	23	8	0	78	68	5
$\beta(0.5, 1)$	86	91	92	73	93	93	93	89	92	94
	86	89	89	49	87	86	80	86	87	86
	85	85	91	0	24	8	0	78	68	5
$Lind(0.5)$	10	11	13	9	9	9	9	9	8	8
	10	11	13	6	10	9	9	9	9	9
	10	11	13	6	9	8	8	8	8	8

Table 9 continued

F	KS_n	CM_n	LS_n	KR_n	$S_{n,0.75}^{(1)}$	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,0.75}^{(2)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$
$Lind(0.5)$	9	10	10	7	7	7	7	8	8	7
	9	9	10	6	7	7	7	8	7	7
	9	10	10	6	7	7	7	8	8	6

When compiling the numerical results, we also considered a wider range of values for the tuning parameter a than those reported in the table. Although some power variation is evident when varying a , the powers achieved by the newly proposed classes of tests are not particularly sensitive to the choice of the tuning parameter a . However, based on the observed numerical powers, we recommend using $S_{n,2}^{(1)}$ when testing the hypothesis in question in the presence of censoring.

4 Practical applications and conclusion

In this section, we apply the tests used in Sect. 3 to test the hypothesis in (1) based on two real-world data sets. The first data set, reported in Table 10, contains the survival

Table 10 Survival times after leukemia diagnosis, in days

7, 47, 58, 74, 177, 232, 273, 285, 317, 429, 440, 445, 455, 468, 495, 497, 532, 571, 579, 581, 650, 702, 715, 779, 881, 900, 930, 968, 1077, 1109, 1314, 1334, 1367, 1534, 1712, 1784, 1877, 1886, 2045, 2056, 2260, 2429, 2509

Table 11 p -values associated with the various tests used in the full sample case

$Test$	KS_n	CM_n	LS_n	KR_n	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,5}^{(1)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$	$S_{n,5}^{(2)}$
p -value	0.49	0.62	0.41	0.34	0.68	0.25	0.18	0.32	0.36	0.19

Table 12 Initial remission times of leukemia patients, in days

4, 5, 8, 8, 9, 10, 10, 10, 10, 10, 11, 12, 12, 12*, 13, 14, 20, 20*, 23, 23, 25, 25, 25, 28, 28, 28, 28, 29, 31, 31, 31, 32, 37, 40, 41, 41, 48, 48, 57, 62, 70, 74, 75, 89, 99, 100, 103, 124, 139, 143, 159*, 161*, 162, 169, 190*, 195, 196*, 197*, 199*, 205*, 217*, 219*, 220, 245*, 258*, 269*

Table 13 p -values associated with the various tests used in the censored case

$Test$	KS_n	CM_n	LS_n	KR_n	$S_{n,1}^{(1)}$	$S_{n,2}^{(1)}$	$S_{n,5}^{(1)}$	$S_{n,1}^{(2)}$	$S_{n,2}^{(2)}$	$S_{n,5}^{(2)}$
p -value	0.03	0.04	0.01	0.28	0.13	0.28	0.41	0.66	0.82	0.73

times, in days, of 43 Leukemia patients. For a discussion of the original data set see Kotze and Johnson (1983) as well as Allison et al. (2017). This data set is not subject to censoring, i.e. all lifetimes are observed. The second data set contains the initial remission times of leukemia patients, in days; for more details see Lee and Wang (2003), this data set can be found in Table 12. These data contains censored observations, indicated using an asterisk. The original data were segmented into three treatment groups. However, Lee and Wang (2003) showed that the data do not display significant differences among the various treatments. As a result we treat the data as i.i.d. realisations from a single, censored, lifetime distribution. All reported p-values are estimated using one hundred thousand bootstrap replications; these results are displayed in Tables 11 and 13, respectively.

From the results of the practical example in Table 11 it is clear that none of the tests reject the null hypothesis that the survival times after a leukemia diagnosis are Weibull distributed at the 5% or 10% levels of significance. As a result, we conclude that the Weibull distribution is an appropriate model for these data.

The results associated with the initial remission times, in Table 13, indicate that KS_n , CM_n and LS_n reject the hypothesis in (1) at a 5% significance level. However, none of the remaining 7 tests considered result in a rejection of the null hypothesis at the 5% or 10% levels. We conclude that the Weibull distribution is likely to be an appropriate model for the observed times. The data set under consideration was also analysed in Bothma et al. (2020), where the null hypothesis of exponentiality of the remission time was strongly rejected. The mentioned paper recommended that a more flexible distribution be used when modelling these data. The results above indicate that the additional flexibility of the Weibull (compared to the exponential) distribution indeed ensures that the Weibull distribution is a more appropriate model than the exponential for the initial remission times considered.

A number of interesting numerical phenomena are evident when considering the powers of the various tests. It is clear that the achieved powers and, therefore, the null distribution of the test statistic, is influenced by the shape of the censoring distribution. The effect of the censoring distribution on the critical values of the tests seem not to have been investigated in the literature to date. Some authors perform goodness-of-fit testing by enforcing a parametric assumption on the censoring distribution, see e.g., Kim (2017). An additional consideration that seems to have been neglected in the literature is the effect on the null distribution of the test statistic, and hence the power of the test, of a specific assumption made in the Kaplan-Meier estimate of the distribution function. Some authors, in order to ensure that G_n satisfies the requirements of a distribution, defines $G_n(t) = 1$ for all $t > X_{(n)}$ regardless of whether or not the sample maximum is censored. We are currently investigating these open questions.

Funding Not applicable.

Availability of data and material All data used are in the public domain.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Code availability Custom code was written and are available upon request.

References

- Allison J, Milošević B, Obradović M, Smuts M (2021) Distribution-free goodness-of-fit tests for the pareto distribution based on a characterization. *Comput Stat* pp 1–16
- Allison J, Santana L (2015) On a data-dependent choice of the tuning parameter appearing in certain goodness-of-fit tests. *J Stat Comput Simul* 85(16):3276–3288
- Allison JS, Betsch S, Ebner B, Visagie IJH (2019) New weighted L^2 -type tests for the inverse Gaussian distribution. *arXiv preprint* [arXiv:1910.14119](https://arxiv.org/abs/1910.14119)
- Allison JS, Huskova M, Meintanis SG (2017) Testing the adequacy of semiparametric transformation models. *Test* 27:1–25
- Allison JS, Santana L, Smit N, Visagie IJH (2017) An “apples-to-apples” comparison of various tests for exponentiality. *Comput Stat* 32(4):1241–1283
- Balakrishnan N, Chimitova E, Vedernikova M (2015) An empirical analysis of some nonparametric goodness-of-fit tests for censored data. *Commun Stat Simul Comput* 44(4):1101–1115
- Baringhaus L, Ebner B, Henze N (2017) The limit distribution of weighted L^2 -goodness-of-fit statistics under fixed alternatives, with applications. *Ann Inst Stat Math* 69(5):969–995
- Baringhaus L, Henze N (1988) A consistent test for multivariate normality based on the empirical characteristic function. *Metrika* 35(1):339–348
- Baringhaus L, Henze N (1991) A class of consistent tests for exponentiality based on the empirical laplace transform. *Ann Inst Stat Math* 43(3):551–564
- Betsch S, Ebner B (2018) Testing normality via a distributional fixed point property in the Stein characterization. *TEST* pp. 1–34. <https://doi.org/10.1007/s11749-019-00630-0>
- Betsch S, Ebner B (2019) A new characterization of the gamma distribution and associated goodness-of-fit tests. *Metrika* 82(7):779–806
- Bothma E, Allison JS, Cockeran M, Visagie IJH (2020) Kaplan-Meier based tests for exponentiality in the presence of censoring. *arXiv preprint* [arXiv:2011.04519](https://arxiv.org/abs/2011.04519)
- Breslow N, Crowley J (1974) A large sample study of the life table and product limit estimates under random censorship. *Ann Stat* pp. 437–453
- Cabaña A, Quiroz AJ (2005) Using the empirical moment generating function in testing for the Weibull and the type I extreme value distributions. *Test* 14(2):417–431
- Cox DR, Oakes D (1984) *Analysis of survival data*, vol 21. CRC Press, Boca Raton
- Cuparić M, Milošević B (2021) New characterization-based exponentiality tests for randomly censored data. *TEST*. <https://doi.org/10.1007/s11749-021-00787-7>
- D’Agostino RB, Stephens MA (1986) *Goodness-of-fit techniques*, vol 68. CRC Press, Boca Raton
- Döring M, Cramer E (2019) On the power of goodness-of-fit tests for the exponential distribution under progressive Type-II censoring. *J Stat Comput Simul* 89:2997–3034
- Efron B (1967) The two sample problem with censored data. In: *Proceedings of the fifth berkeley symposium on mathematical statistics and probability*, vol 4, pp 831–853
- Fernández T, Rivera N (2020) Kaplan-Meier V-and U-statistics. *Electron J Stat* 14(1):1872–1916
- Fernández VA, Jiménez Gamero MD, García M (2008) A test for the two-sample problem based on empirical characteristic functions. *Comput Stat Data Anal* 52:3730–3748
- Feuerverger A, Mureika RA (1977) The empirical characteristic function and its applications. *Ann Stat* 5(1):88–97
- Giacomini R, Politis DN, White H (2013) A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econom Theory* 29(3):567–589
- Gupta RD, Richards DSP (1997) Invariance properties of some classical tests for exponentiality. *J Stat Plann Inference* 63(2):203–213
- Henze N, Visagie IJH (2020) Testing for normality in any dimension based on a partial differential equation involving the moment generating function. *Ann Inst Stat Math* 72:1109–1136
- Hlavac M (2018) stargazer: well-formatted regression and summary statistics tables. <https://CRAN.R-project.org/package=stargazer>
- Jacobson A, Wilson V, Pileggi S (2018) *parmsurvfit: Parametric Models for Survival Data*. <https://CRAN.R-project.org/package=parmsurvfit>. R package version 0.1.0

- Jiang R, Murthy D (2011) A study of Weibull shape parameter: properties and significance. *Reliab Eng Syst Saf* 96(12):1619–1626
- Kalbfleisch JD, Prentice RL (2011) *The statistical analysis of failure time data*, vol 360. Wiley, New York
- Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53(282):457–481
- Kim N (2017) Goodness-of-fit tests for randomly censored Weibull distributions with estimated parameters. *Commun Stat Appl Methods* 24(5):519–531
- Klar B, Meintanis SG (2005) Tests for normal mixtures based on the empirical characteristic function. *Comput Stat Data Anal* 49(1):227–242
- Kotz S, Nadarajah S (2000) *Extreme value distributions: theory and applications*. World Scientific, Singapore
- Kotze S, Johnson N (1983) *Encyclopedia of statistical sciences*, vol 3. Wiley, New York
- Koziol JA, Green SB (1976) A Cramér-von Mises statistic for randomly censored data. *Biometrika* 63(3):465–474
- Krit M (2014) Goodness-of-fit tests for the Weibull distribution based on the Laplace transform. *Journal de la Société Française de Statistique* 155(3):135–151
- Lee ET, Wang J (2003) *Statistical methods for survival data analysis*, vol 476. Wiley, New York
- Liao M, Shimokawa T (1999) A new goodness-of-fit test for type-i extreme-value and 2-parameter Weibull distributions with estimated parameters. *Optimization* 64(1):23–48
- Mann NR, Scneuer EM, Fertig KW (1973) A new goodness-of-fit test for the two-parameter Weibull or extreme-value distribution with unknown parameters. *Commun Stat Theory Methods* 2(5):383–400
- Mazucheli J, Fernandes LB, de Oliveira RP (2016) *LindleyR: The Lindley Distribution and Its Modifications*. <https://CRAN.R-project.org/package=LindleyR>. R package version 1.1.0
- Meintanis SG, Iliopoulos G (2003) Tests of fit for the Rayleigh distribution based on the empirical Laplace transform. *Ann Inst Stat Math* 55(1):137–151
- Meintanis SG, Ngatchou-Wandji J, Allison JS (2018) Testing for serial independence in vector autoregressive models. *Stat Papers* 59(4):1379–1410
- Mijburgh PA, Visagie IJH (2020) An overview of goodness-of-fit tests for the Poisson distribution. *South African Stat J* 54(2):207–230
- R Core Team (2019) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Tiku ML, Singh M (1981) Testing the two parameter Weibull distribution. *Commun Stat Theory Methods* 10(9):907–918

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

3.1 Addendum to Article 2

Note that on page 3 of the article there is a typographical error in the formula for the log-likelihood. The log-likelihood of the Weibull distribution is

$$\mathcal{L}(\theta, \lambda | T_1, \dots, T_n) = d \log(\theta) - d\theta \log(\lambda) + (\theta - 1) \sum_{j=1}^n \delta_j \log(T_j) - \lambda^{-\theta} \sum_{j=1}^n T_j^\theta,$$

where $d = \sum_{j=1}^n \delta_j$. In the full sample case $d = n$.

Note that on page 8 there are typographical errors in the formulae for KS_n , CM_n and LS_n they should be

$$KS_n = \max \left[\max_{1 \leq j \leq n} \left\{ G_n(Y_{(j)}) - \left(1 - e^{-e^{Y_{(j)}}} \right) \right\}, \right. \\ \left. \max_{1 \leq j \leq n} \left\{ \left(1 - e^{-e^{Y_{(j)}}} \right) - G_n^-(Y_{(j)}) \right\} \right],$$

where $G_n^-(Y_{(j)}) = \lim_{t \uparrow Y_{(j)}} G_n(t)$,

$$CM_n = \frac{n}{3} + n \sum_{j=1}^{d+1} \left\{ G_n \left(X_{(j-1)}^{(t)} \right) \left(X_{(j)}^{(t)} - X_{(j-1)}^{(t)} \right) \right. \\ \left. \times \left[G_n \left(X_{(j-1)}^{(t)} \right) - \left(X_{(j)}^{(t)} + X_{(-1)}^{(t)} \right) \right] \right\},$$

and

$$LS_n = \frac{1}{\sqrt{n}} \sum_{j=1}^n \left(\frac{\max [j/n - G_n(Y_{(j)}), G_n(Y_{(j)}) - (j-1)/n]}{\sqrt{G_n(Y_{(j)}) [1 - G_n(Y_{(j)})]}} \right).$$

Below we include some of the derivations omitted in the main text of the paper. The specific derivations included relate to the quantity η and the function φ which are required for the derivation of the computational form of the test statistic.

We propose the following integral form of the test statistic based on (3):

$$S_{n,a} = n \int_{-\infty}^{\infty} \varphi_n(t) w(t) dt, \quad (3.1)$$

where $w_a(t)$ is a weight function containing a user-defined tuning parameter $a > 0$. Now, let $v_j = (1 - e^{Y_j})$,

and define $\varphi(t)_n$;

$$\begin{aligned}
\varphi_n(t) &= \left| \sum_{j=1}^n \Delta_j \left[ite^{itY_j} + v_j e^{itY_j} \right] \right|^2 \\
&= \left| \sum_{j=1}^n \Delta_j \left\{ it \left[\cos(tY_j) + i \sin(tY_j) \right] + v_j \left[\cos(tY_j) + i \sin(tY_j) \right] \right\} \right|^2 \\
&= \left| \sum_{j=1}^n \Delta_j \left\{ -t \sin(tY_j) + v_j \cos(tY_j) + i \left[t \cos(tY_j) + v_j \sin(tY_j) \right] \right\} \right|^2 \\
&= \left(\sum_{j=1}^n \Delta_j \left[v_j \cos(tY_j) - t \sin(tY_j) \right] \right)^2 + \left(\sum_{j=1}^n \Delta_j \left[t \cos(tY_j) + v_j \sin(tY_j) \right] \right)^2 \\
&= \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k (v_j \cos(tY_j) - t \sin(tY_j)) (v_k \cos(tY_k) - t \sin(tY_k)) \\
&\quad + \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k (t \cos(tY_j) + v_j \sin(tY_j)) (t \cos(tY_k) + v_k \sin(tY_k)) \\
&= \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[v_j v_k \cos(tY_j) \cos(tY_k) - 2tv_j \cos(tY_j) \sin(tY_k) + t^2 \sin(tY_j) \sin(tY_k) \right. \\
&\quad \left. + t^2 \cos(tY_j) \cos(tY_k) + 2tv_j \sin(tY_j) \cos(tY_k) + v_j v_k \sin(tY_j) \sin(tY_k) \right]
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left\{ t^2 \left[\cos(tY_j) \cos(tY_k) + \sin(tY_j) \sin(tY_k) \right] \right. \\
&\quad + 2tv_j \left[\sin(tY_j) \cos(tY_k) - \cos(tY_j) \sin(tY_k) \right] \\
&\quad \left. + v_j v_k \left[\cos(tY_j) \cos(tY_k) + \sin(tY_j) \sin(tY_k) \right] \right\} \\
&= \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[t^2 \cos(t(Y_j - Y_k)) + 2tv_j \sin(t(Y_j - Y_k)) + v_j v_k \cos(t(Y_j - Y_k)) \right].
\end{aligned}$$

CHAPTER 4

Article 3: On the effect of the Kaplan-Meier estimator's assumed tail behaviour on goodness-of-fit testing

The third article, *On the effect of the Kaplan-Meier estimator's assumed tail behaviour on goodness-of-fit testing*, is a draft article and has therefore not yet been submitted to any journal. We plan to submit this paper to *Computational Statistics and Data Analysis*.

ON THE EFFECT OF THE KAPLAN-MEIER ESTIMATOR'S ASSUMED TAIL BEHAVIOUR ON GOODNESS-OF-FIT TESTING

A PREPRINT

E. Bothma
Subject Group Statistics
North-West University
South Africa

J.S. Allison
Subject Group Statistics
North-West University
South Africa

I.J.H. Visagie
Subject Group Statistics
North-West University
South Africa
jaco.visagie@nwu.ac.za

January 24, 2022

ABSTRACT

When analysing lifetime data in the presence of censoring one is often required to estimate the distribution function of the lifetimes non-parametrically. The most popular estimator used for this purpose is the Kaplan-Meier estimator. Interestingly, in its initial formulation this estimator is only defined up to the observed sample maximum. For values larger than the sample maximum two different assumptions are commonly used in the statistical literature. The first is to set the value of the estimate to one while the second is to use the value of the estimate at the sample maximum when estimating the tail of the distribution function. This paper illustrates the profound effect of these assumptions on the sizes and powers of goodness-of-fit tests for three classes of distributions often used in survival analysis. These differences are illustrated using observed remission time data. The considered classes of distributions are the exponential, Weibull and gamma. As a result of independent interest, we amend two classes of tests developed for the gamma distribution in the full sample case for use with censored data.

Keywords Exponential distribution · Gamma distribution · Goodness-of-fit testing · Random right censoring · Weibull distribution.

1 Introduction

Statistical practitioners and other researchers are often interested in the parametric modelling of observed lifetime data in survival analysis and reliability theory as well as many other fields. In order to perform inference, it is necessary to test the goodness-of-fit hypothesis that the observed lifetimes are realised from a specified class of distributions. This hypothesis is complicated by the fact that random right censoring is often present in the applications of interest. Testing the goodness-of-fit of a specific class of distributions is a topic of continued interest; recent papers on this topic include [1] and [2] (in which the assumption of exponentiality is tested) as well as [3] (in which the the assumption of the class of Weibull distributions is tested).

In the full sample case, the classical goodness-of-fit tests are based on comparisons between non-parametric and parametric estimates of the distribution function (obtained by estimating the parameters of the specified class of distributions under the null hypothesis). Consider, for example, the Cramér-von Mises test, see [4]. The idea underlying this test is to compare the empirical distribution function to a fitted parametric distribution function obtained by estimating the parameters (using consistent estimators) of the specified class of distributions. If the distance between the two estimates, as measured by an integral involving the squared difference between the empirical distribution function and the fitted distribution, is “too large” (to be made precise below), then the null hypothesis is rejected. In fact, this approach also underlies other classical goodness-of-fit tests such as the Kolmogorov-Smirnov test, which is obtained upon replacing the distance measure above by the supremum difference.

A natural approach to the goodness-of-fit problem in the presence of censoring is to extend the techniques described above to account for censoring. That is, in the presence of censoring, the Cramér-von Mises test for a class of distributions is based on a specified distance between a non-parametric estimate of the distribution function, similar to the empirical distribution function, and a parametric estimate of the distribution function under the null hypothesis. Parameter estimation in the presence of random right censoring has been studied extensively in the literature, see e.g. [5] as well as [6]. The most popular estimation technique used is the method of maximum likelihood; this method is used for parameter estimation throughout this paper. It remains to find a non-parametric estimate of the distribution function. This can be achieved using the well-known Kaplan-Meier estimator, see [7]. In order to proceed we introduce some notation below.

Let $X_1, \dots, X_n$ denote independent and identically distributed (i.i.d.) lifetimes with continuous distribution and density functions F and f , respectively. Let $C_1, \dots, C_n$ be i.i.d. censoring variables with continuous distribution G , independent of $X_1, \dots, X_n$. We assume that the censoring mechanism used is non-informative. The observed times are $T_j = \min(X_j, C_j)$ and we define the indicator variables $\delta_j = \mathbb{I}(X_j \leq C_j)$ and the order statistics of T_j as $T_{(1)} \leq \dots \leq T_{(n)}$, with $\delta_{(j)}$ corresponding to $T_{(j)}$. Let $\mathcal{F}_\theta$ denote a parametric class of distribution functions indexed by some, possibly vector valued, parameter θ . Based on the observed pairs (T_j, δ_j) , $j = 1, \dots, n$ we wish to test the composite goodness-of-fit hypothesis

$$H_0 : F \in \mathcal{F}_\theta,$$

against general alternatives. In this paper, we present numerical results relating to three classes of distributions, the exponential, Weibull and gamma. References to recent papers relating to testing for the first two of these are provided above. However, we have been unable to obtain any papers relating to goodness-of-fit testing for the gamma distribution in the presence of censoring. In Section 2, we adapt several existing tests for the gamma distribution in the full sample case for use with random right censored data.

Using the notation introduced above, the Kaplan-Meier estimator, $\tilde{F}_n$, is a step function given by

$$\tilde{F}_n(t) = \begin{cases} 0, & 0 < t \leq T_{(1)} \\ 1 - \prod_{j=1}^{k-1} \left(\frac{n-j}{n-j+1} \right)^{\delta_{(j)}}, & T_{(k-1)} < t \leq T_{(k)}, \quad k = 2, \dots, n. \end{cases} \quad (1)$$

Interestingly, in its initial formulation, the Kaplan-Meier estimator of the distribution function is only defined up to the sample maximum, say $\tau = T_{(n)}$. In cases where the maximum observation is censored, two different assumptions or conventions are commonly used in the statistical literature to define $\tilde{F}_n(t)$ for $t > \tau$. The first is to simply set the value of $\tilde{F}_n(t) = 1$, $\forall t > \tau$, see e.g. [8], while the second is to set $\tilde{F}_n(t) = \tilde{F}_n(\tau)$, $\forall t > \tau$, see e.g. [5]. For the sake of brevity, we refer to this convention as the “set-to-1” convention; i.e., we refer to the first and second conventions mentioned above as cases where the set-to-1 assumption is made and not made, respectively.

Regardless of the set-to-1 assumption, all jumps in $\tilde{F}_n$ occurring before the sample maximum occur at values of T_j for which $\delta_j = 1$. If the sample maximum τ is uncensored, then the final jump in $\tilde{F}_n$ occurs at τ regardless of the set-to-1 assumption. However, if the sample maximum is censored, then the set-to-1 assumption influences the tail behaviour of the value of $\tilde{F}_n(t)$ when $t > \tau$. If the set-to-1 assumption is made the final jump occurs immediately after τ ; however, if this assumption is not made there is no jump immediately following τ . The latter case results in an improper distribution with positive probability of realising the value ∞ . In summary, the set-to-1 assumption only influences the estimate $\tilde{F}_n$ if the sample maximum is censored, and then it only affects $\tilde{F}_n(t)$, $t > \tau$. This paper illustrates the profound effect that this seemingly small difference has on the powers of goodness-of-fit tests by conducting an extensive Monte Carlo simulation study. Before we proceed, the effect of the assumption used in the Kaplan-Meier estimator is demonstrated using an example based on observed data below.

A data set containing 66 initial remission times of leukemia patients, 14 of which are censored, is presented in [9]. The data are shown in Table 1, censored observations are indicated using an asterisk in the superscript. Initially, the data were segmented into three groups based on the treatments that the patients received. However, [9] found no statistically significant differences between the groups, so we treat the data as realisations from a single distribution. Figure 1 shows the Kaplan-Meier estimate of the distribution function using both of the assumptions discussed above. The two Kaplan-Meier step functions considered coincide up to the sample maximum value of 269 (indicated in Figure 1 using a dotted vertical line). However, since the last observation is censored, the two versions of this estimate differ when $t > 269$. Under the set-to-1 assumption, $\tilde{F}_n(t) = 1$, $\forall t > 269$. If this assumption is not made, then $\tilde{F}_n(t) = 0.853$, $\forall t > 269$. As can be seen in the figure, the set-to-1 assumption makes a substantial difference in the estimated value of the distribution function for large values of t . Additionally, Figure 1 shows two parametric estimates of the distribution function obtained using exponential and Weibull distributions respectively. The techniques used to obtain these parametric estimates are discussed below.

Table 1: Initial remission times of leukemia patients in days.

4, 5, 8, 8, 9, 10, 10, 10, 10, 10, 11, 12, 12, 12*, 13, 14, 20, 20*, 23, 23, 25, 25, 25, 28, 28, 28, 28, 29, 31, 31, 31, 32, 37, 40, 41, 41, 48, 48, 57, 62, 70, 74, 75, 89, 99, 100, 103, 124, 139, 143, 159*, 161*, 162, 169, 190*, 195, 196*, 197*, 199*, 205*, 217*, 219*, 220, 245*, 258*, 269*
--

As a first parametric model, consider the class of exponential distributions, with density $f(t) = \lambda e^{-\lambda t}$, $\lambda > 0$, $t > 0$. We denote a random variable having this density by $\exp(\lambda)$. The maximum likelihood estimate of the rate parameter of the exponential distribution is

$$\hat{\lambda} = \frac{\sum_{j=1}^n \delta_j}{\sum_{j=1}^n T_j}.$$

For the example at hand, $\hat{\lambda} = 0.01$. Figure 1 shows the fitted exponential distribution function using a solid line. Additionally, we fit a Weibull distribution to the observed data. The density of the Weibull distribution with parameters κ and λ is $f(t) = \kappa \lambda^\kappa t^{\kappa-1} e^{-(\lambda t)^\kappa}$, $\kappa > 0$, $\lambda > 0$, $t > 0$. A random variable with this density is denoted by $W(\kappa, \lambda)$. The log-likelihood of the $W(\kappa, \lambda)$ distribution given (T_j, δ_j) , $j = 1, \dots, n$ is

$$\mathcal{L}(\kappa, \lambda | X_1, \dots, X_n) = d \log(\kappa) + d\kappa \log(\lambda) + (\kappa - 1) \sum_{j=1}^n \delta_j \log(X_j) - \lambda^\kappa \sum_{j=1}^n X_j^\kappa,$$

where $d = \sum_{j=1}^n \delta_j$. Closed form formulae for the maximum likelihood estimators of κ and λ are not available. The required estimates are obtained using numerical optimisation. All calculations presented are performed in R, see [10]. We estimate the parameters of the Weibull distribution using the `fit_data` function in the `parmsurvfit` package, see [11]. For the current example, the parameter estimates obtained are $\hat{\kappa} = 0.812$ and $\hat{\lambda} = 0.01$. The fitted Weibull distribution function is shown in Figure 1 using a dotted line.

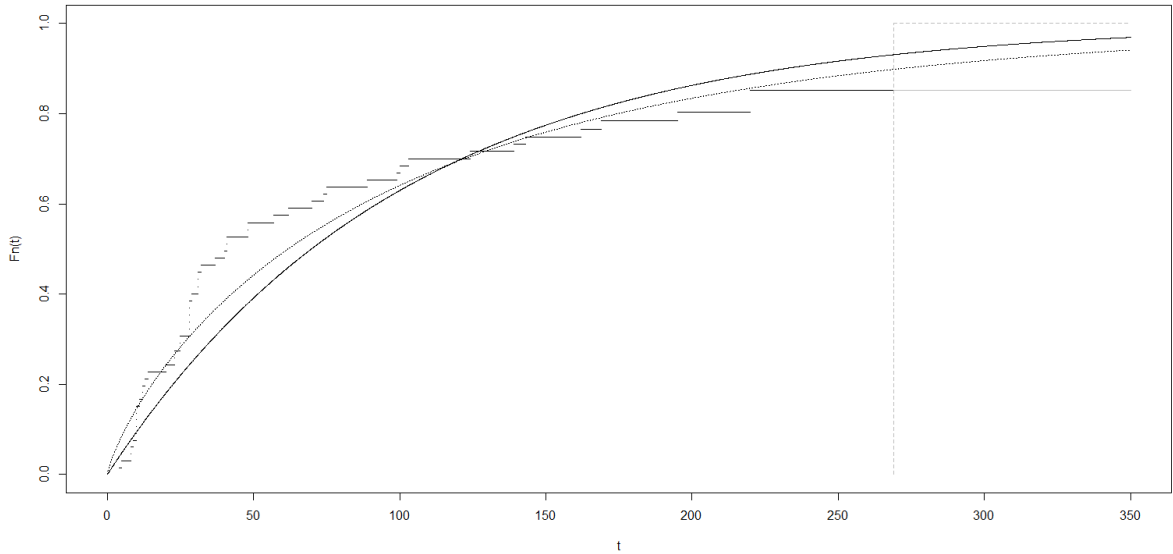


Figure 1: Kaplan-Meier estimate of the lifetimes with fitted exponential and Weibull distributions.

Figure 1 provides us with a visual test of fit for both the exponential and Weibull distributions; the fitted Weibull seems to resemble the non-parametric estimates more closely than does the fitted exponential. This is to be expected as the exponential distribution is a special case of the Weibull. In order to formally test the goodness-of-fit hypotheses of interest, we use the Cramér-von Mises test for both classes of distributions in turn. Goodness-of-fit test statistics for the exponential and Weibull distributions are typically performed based on transformed data. In the case of the exponential distribution, we rescale the data by multiplication with the estimated rate parameter. This essentially reduces the null hypothesis to testing for standard exponentiality. In the same way, we may use a transformation to reduce the

test for the Weibull class of distributions to a test for the standard extreme value distribution; i.e., $X \sim W(\kappa, \lambda)$ if, and only if, $Y = \kappa(\log(X_j) - \log(\lambda))$, $j = 1, \dots, n$ follows a standard extreme value distribution. That is, Y has density $h(x) = e^{x-e^x}$, $x \in \mathbb{R}$. We denote the variables resulting from the sample version of this transformation by $Y_j = \hat{\lambda}[\log(T_j) - \log(\hat{\kappa})]$, $j = 1, \dots, n$ for both the exponential and Weibull classes (as well as for the gamma class later in the paper) as we believe that there is little possibility of confusion. Using transformations of this kind is standard practice in the goodness-of-fit literature, see e.g., [12], [13] as well as [14]. Below, we use $F_{\hat{\theta}}$ and $\tilde{F}_n$ to denote the fitted distribution function and the Kaplan-Meier estimate of the distribution function, respectively.

Let CM_n denote the Cramér-von Mises test statistic based on (Y_j, δ_j) , $j = 1, \dots, n$. As was mentioned above, CM_n is a distance measure between the parametric estimate $F_{\hat{\theta}}$ and the non-parametric estimate $\tilde{F}_n$;

$$CM_n = \int_{-\infty}^{\infty} (F_{\hat{\theta}}(t) - \tilde{F}_n)^2 dF_{\hat{\theta}}(t), \quad t \geq 0. \quad (2)$$

We can exploit the fact that $\tilde{F}_n$ is a step function in order to arrive at an easily calculable expression for CM_n . Using notation similar to that of [4], denote by y_j , $j = 1, \dots, d$ the set of uncensored lifetimes and let $\tilde{\tau}$ be the transformed value of the sample maximum. The Cramér-von Mises test statistic defined in (2) can be expressed as

$$\begin{aligned} CM_n = & \frac{n}{3} + n \sum_{j=1}^d \tilde{F}_n(y_{j-1}) [F_{\hat{\theta}}(y_j) - F_{\hat{\theta}}(y_{j-1})] [\tilde{F}_n(y_{j-1}) - (F_{\hat{\theta}}(y_j) - F_{\hat{\theta}}(y_{j-1}))] \\ & + n \tilde{F}_n(y_d) [F_{\hat{\theta}}(\tilde{\tau}) - F_{\hat{\theta}}(y_d)] [\tilde{F}_n(y_d) - (F_{\hat{\theta}}(\tilde{\tau}) - F_{\hat{\theta}}(y_d))] \\ & + n \tilde{F}_n^+(\tilde{\tau}) [1 - F_{\hat{\theta}}(\tilde{\tau})] [\tilde{F}_n^+(\tilde{\tau}) - (1 - F_{\hat{\theta}}(\tilde{\tau}))], \end{aligned}$$

where $g^+(t)$ denotes the right limit of the function g . In the case where the set-to-1 assumption is made the formula above coincides with the formula given in [4].

For the example under consideration, the sample maximum is censored, meaning that the calculated value of $\tilde{F}_n$, and therefore the value of CM_n , depends on the set-to-1 assumption. Consider first the case where $\mathcal{F}_{\theta}$ is the class of exponential distributions. When making the set-to-1 assumption, CM_n is calculated to be 0.543. If this assumption is not made, then CM_n equals 0.596.

In order to formally test the hypothesis that the lifetimes are realised from an exponential distribution, we need to approximate the distribution of the test statistic under the null hypothesis. In the full sample case, the fact that the exponential class of distributions is a scale family ensures that the distribution of the CM_n under the null hypothesis is independent of the value of λ . However, in the setup considered in this paper, we are unable to exploit this feature of the exponential class as even the hypothesis that lifetimes are standard exponential is a composite hypothesis as the censoring distribution remains unspecified. As a result, we use a bootstrap algorithm in order to approximate the null distribution of CM_n . Appropriate bootstrap algorithms for use with the exponential and Weibull classes can be found in [2] and [13], respectively. However, the provided algorithms specifically assume that the set-to-1 convention is not used. Below we provide a more general algorithm in which the user can specify whether or not to make the set-to-1 assumption.

In order to proceed, we need to be able to obtain bootstrap samples from both the lifetime and censoring distributions. Generating samples from the lifetime distribution under the null hypothesis is a simple matter; we estimate the parameters of the specified class of distributions and we simulate from this fitted distribution (using, e.g. the inverse transform method). Since the null hypothesis does not specify the censoring distribution, we need to estimate this distribution from the observed data. We use the Kaplan-Meier estimator of the distribution of the censored observations; this is done by calculating $\tilde{F}_n$ in (1), but replacing δ_j by $1 - \delta_j$ in the calculations. Of course, since we are using the Kaplan-Meier estimator to estimate the distribution of the censoring times, we are confronted with the set-to-1 assumption once again. Immediately, one is tempted to argue that whatever convention was used to estimate the distribution function of the lifetimes should be used in order to estimate the distribution of the censoring times. However, there is no contradiction in, e.g. making the set-to-1 assumption for the censoring times and not for the lifetimes. In order to ease notation in the remainder of the paper we use the notation “TF” in order to indicate this specific combination of conventions used since it means that we assume that the set-to-1 assumption is “true” for the censoring distribution and “false” for the lifetime distribution. As a result, we are left with four distinct possibilities:

- Case 1: The set-to-1 assumption is not made for either of the estimated distributions (indicated by “FF”).
- Case 2: The set-to-1 assumption is made for the lifetime distribution but not for the censoring distribution (indicated by “FT”).

- Case 3: The set-to-1 assumption is made for the censoring distribution but not for the lifetime distribution (indicated by “TF”).
- Case 4: The set-to-1 assumption is made for both of the estimated distributions (indicated by “TT”).

The Kaplan-Meier estimate of the censoring distribution in the practical example is shown in Figure 2. Since the sample maximum is censored, the Kaplan-Meier estimate achieves the value of 1 at $t = 269$ regardless of the set-to-1 assumption.

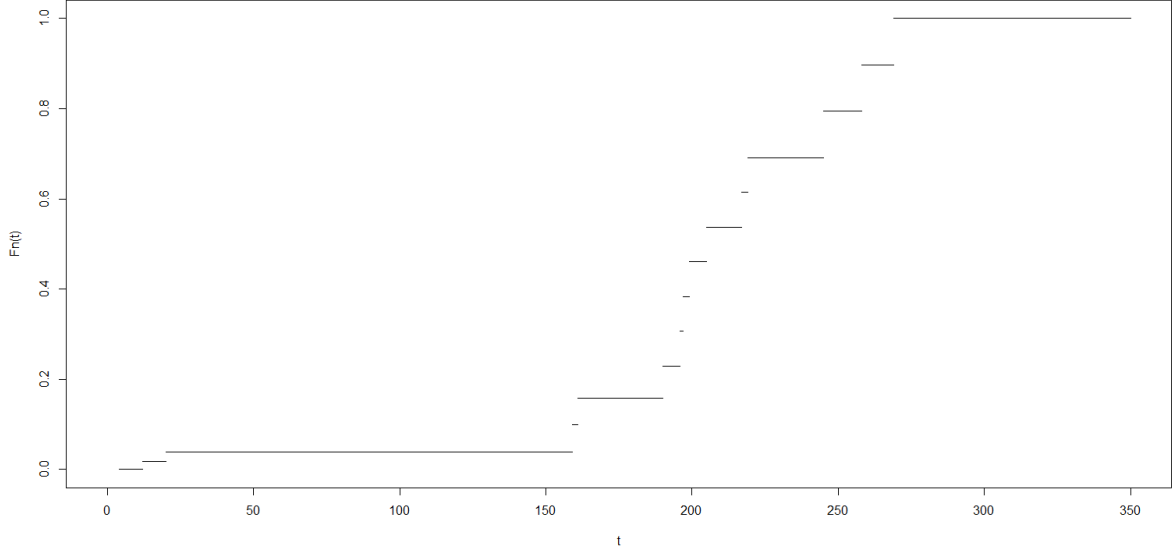


Figure 2: Kaplan-Meier estimate of the censored times.

We now provide the bootstrap algorithm required in order to approximate critical values for goodness-of-fit tests in the presence of random right censoring.

1. Using $(T_j, \delta_j), j = 1, \dots, n$ estimate the parameters of the class of distributions as specified in H_0 .
2. Obtain a parametric bootstrap sample $X_1^*, \dots, X_n^*$ by sampling from the fitted distribution.
3. Obtain a non-parametric bootstrap sample $C_1^*, \dots, C_n^*$ by sampling from the Kaplan-Meier estimate of the distribution of the censoring times (setting the set-to-1 assumption to true or false).
4. Set

$$T_j^* = \min(X_j^*, C_j^*) \text{ and } \delta_j^* = \begin{cases} 1, & \text{if } X_j^* \leq C_j^* \\ 0, & \text{if } X_j^* > C_j^*. \end{cases}$$

5. Using (T_j^*, δ_j^*) estimate the parameters of the class of distributions and use the transformation to obtain $Y_j^*, j = 1, \dots, n$.
6. Based on the data $(Y_j^*, \delta_j^*), j = 1, \dots, n$ calculate the value of the test statistic, say $S^* = S((Y_1^*, \delta_1^*), \dots, (Y_n^*, \delta_n^*))$.
7. Repeat steps 2-6 B times, resulting in $S_1^*, \dots, S_B^*$. Denote the corresponding order statistics by $S_{(1)}^* \leq \dots \leq S_{(B)}^*$. The approximate critical value is $\hat{c}_n(\alpha) = S_{(\lfloor B(1-\alpha) \rfloor)}^*$ where $\lfloor \cdot \rfloor$ denotes the floor function.

The algorithm presented above is general and not only intended for use with the exponential null hypothesis. Using this algorithm, and setting $B = 50\,000$, we obtain an approximation to the null distribution of the CM_n statistic when testing for the exponential class. In the current example, the set-to-1 assumption regarding the Kaplan-Meier estimate of the censoring distribution is unimportant since it does not influence the resulting estimated distribution function. However, the set-to-1 convention used influences the value of the test statistic obtained and therefore the p -value of the test. In this case, an approximate p -value of 0.00072 is obtained when making the set-to-1 assumption, while the

p -value is estimated to be 0.00032 when not using this assumption. In both cases, we reject the hypothesis that the lifetimes are realised from an exponential distribution at a significance level of 1%.

We now turn our attention to the hypothesis that the lifetimes are realised from the Weibull class of distributions. The values of CM_n is calculated to be 0.244 when making the set-to-1 assumption, which corresponds to a p -value of 0.019. If this assumption is not made, the realised value of CM_n is 0.290, corresponding to a p -value of 0.011. As a result, we do not reject the hypothesis that the lifetimes follow a Weibull distribution at a 1% significance level.

The final class of distributions considered is the gamma class. A random variable with density $f(t) = \lambda^\kappa \Lambda^{-1}(\kappa) t^{\kappa-1} e^{-\lambda t}$, $\kappa > 0$, $\lambda > 0$, $t > 0$, with $\Lambda(\kappa) = \int_0^\infty x^{\kappa-1} e^{-x} dx$, is said to have a gamma distribution with shape parameter κ and rate parameter λ . We denote a random variable with this density by $\Gamma(\kappa, \lambda)$. When testing the hypothesis that the remission times considered are realised from a gamma distribution, we calculate CM_n to be 0.376 and 0.326 respectively when making the set-to-1 assumption and not making this assumption. In both cases, the corresponding p -value is below 1%, meaning that we reject the class of gamma distributions as a suitable model for the lifetimes. Based on the p -values obtained using the three classes of distributions considered, we recommend using a Weibull distribution when modelling the observed lifetimes for this data set.

2 Test statistics for the various classes of lifetime distributions

Below, we provide the details of a number of goodness-of-fit tests for the exponential, Weibull and gamma classes. The performances of the tests in this section are evaluated in a Monte Carlo simulation study presented in Section 3. We begin by specifying the classical Kolmogorov-Smirnov test. As is the case for the Cramér-von Mises test, discussed in Section 1, the Kolmogorov-Smirnov test is general in the sense that it can be used to test for any of the classes of distributions considered.

After considering the Kolmogorov-Smirnov test, we turn our attention to tests developed for each of the considered classes of distributions separately. The remaining tests are based on the characteristic function and Laplace transform. These functions and their empirical counterparts have been studied extensively in the full sample case, see e.g. [15]. However, estimates for the empirical versions of these functions are less well known in the presence of censoring. As a result, we discuss these functions before defining the relevant test statistics.

2.1 The Kolmogorov-Smirnov test

The Kolmogorov-Smirnov (KS_n) test measures the supremum distance between $\tilde{F}_n$ and $F_{\hat{\theta}}$:

$$KS_n = \sup_{t \geq 0} |F_{\hat{\theta}}(t) - \tilde{F}_n(t)|.$$

Similar to the full sample case, the differences need only be evaluated in the jump points of the piecewise constant $\tilde{F}_n$. The value of KS_n can be calculated as

$$KS_n = \max\{KS_n^+, KS_n^-, 1 - \lim_{t \rightarrow \infty} \tilde{F}_n(t)\}, \quad (3)$$

where

$$KS_n^+ = \max_{1 \leq j \leq n} \left\{ \tilde{F}_n(Y_{(j)}) - F_{\hat{\theta}}(Y_{(j)}) \right\}$$

and

$$KS_n^- = \max_{1 \leq j \leq n} \left\{ F_{\hat{\theta}}(Y_{(j)}) - \tilde{F}_n(Y_{(j-1)}) \right\},$$

with $Y_{(0)} := 0$. Note that, if the set-to-1 assumption is made, then $\lim_{t \rightarrow \infty} \tilde{F}_n(t) = 1$ for all samples and the final term in (3) can be omitted.

2.2 Empirical characteristic function and Laplace transforms

Recall that the characteristic function of a random variable, Y , with distribution function, F , is defined to be

$$\phi(s) = E[e^{isY}] = \int e^{isy} dF(y),$$

with $i = \sqrt{-1}$. Using the notation introduced above and using $\tilde{F}_n$ to estimate F , we may estimate the characteristic function based on a censored sample by

$$\tilde{\phi}_n(s) = \int e^{isy} d\tilde{F}_n(y) = \sum_{j=1}^n \Delta_j e^{isY_{(j)}},$$

see e.g. [16], where Δ_j denotes the size of the jump in $\tilde{F}_n(Y_{(j)})$, $j = 1, \dots, n$, given by

$$\Delta_1 = \frac{\delta_{(1)}}{n}, \quad \Delta_n = \prod_{l=1}^{n-1} \left(\frac{n-l}{n-l+1} \right)^{\delta_{(l)}}, \quad \Delta_j = \frac{\delta_{(j)}}{n-j+1} \prod_{k=1}^{j-1} \left(\frac{n-k}{n-k+1} \right)^{\delta_{(k)}}, \quad j = 2, \dots, n-1.$$

Additionally, recall that the Laplace transform of a random variable Y is defined as

$$\psi(s) = \int e^{-sy} dF(y).$$

Upon estimating F by $\tilde{F}_n$, the resulting estimate of the Laplace transform is

$$\tilde{\psi}_n(s) = \sum_{j=1}^n \Delta_j e^{-sY_{(j)}}.$$

2.3 Tests for the exponential distribution

We now consider four tests for the exponential distribution in the presence of random right censoring. An exponentially distributed random variable is defined by the following density function; $f(t) = \lambda e^{-\lambda t}$, $\lambda > 0$, $t > 0$. Each of the tests below contains a user defined tuning parameter $a > 0$. The first two tests are based on the empirical characteristic function while the last two are based on the empirical Laplace transform.

2.3.1 Epps-Pulley (1986)

A test statistic based on the difference between the characteristic function and the empirical characteristic function is proposed in [17]. In the current setting, the test statistic is

$$EP_n = \frac{1}{2\pi} \int \left[\tilde{\phi}_n(t) - \phi(t) \right] \phi(-t) dt.$$

Straightforward calculations result in a calculable form of this test statistic given by

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

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

2.3.2 Henze-Meintanis (2002a)

A test statistic based on a characterisation of the exponential distribution via the characteristic function is introduced in [18]. In the presence of random censoring, the modified test is given by

$$H_{n,a} = n \int_0^\infty [S_n(t) - tC_n(t)]^2 e^{-at} dt,$$

where $S_n(t) = \sum_{j=1}^n \Delta_j \sin(tY_j)$ and $C_n(t) = \sum_{j=1}^n \Delta_j \cos(tY_j)$. An easily computable form of the test statistic is

$$H_{n,a} = \frac{an}{2} \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{1}{a^2 + (Y_j - Y_k)^2} - \frac{1}{a^2 + (Y_j + Y_k)^2} - \frac{4(Y_j + Y_k)}{(a^2 + (Y_j + Y_k)^2)^2} \right. \\ \left. + \frac{2a^2 - 6(Y_j - Y_k)^2}{(a^2 + (Y_j - Y_k)^2)^3} + \frac{2a^2 - 6(Y_j + Y_k)^2}{(a^2 + (Y_j + Y_k)^2)^3} \right].$$

In the simulation study in Section 3, we set $a = 1$ in accordance with the recommendation made in [18]. The null hypothesis of exponentiality is rejected for large values of $H_{n,a}$.

2.3.3 Baringhaus-Henze (1991)

[19] introduces a test statistic based on a partial differential equation involving the Laplace transform. In the presence of random censoring the test statistic becomes

$$B_{n,a} = n \int_0^\infty \left[(1+t) \tilde{\psi}'_n(t) + \tilde{\psi}_n(t) \right]^2 e^{-at} dt,$$

where $\tilde{\psi}'_n$ is the derivative of $\tilde{\psi}_n$. In the simulation study below we use $a = 0.25$. The calculable form of the modified test is given by

$$B_{n,a} = n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{(1-Y_j)(1-Y_k)}{Y_j + Y_k + a} - \frac{Y_j + Y_k}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^3} \right].$$

The null hypothesis of exponentiality is rejected for large values of $B_{n,a}$.

2.3.4 Henze-Meintanis (2002b)

A test statistic based on the squared difference between the Laplace transform and its empirical counterpart is proposed in [20]. In the presence of censoring, the modified test can be expressed as

$$L_{n,a} = n \int_0^\infty \left[\tilde{\psi}_n(t) - \psi(t) \right]^2 (1+t)^2 e^{-at} dt.$$

This test statistic lends itself to the following easily calculable form;

$$L_{n,a} = n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\frac{1 + (Y_j + Y_k + a + 1)^2}{(Y_j + Y_k + a)^3} \right] - 2n \sum_{j=1}^n \Delta_j \left[\frac{1 + Y_j + a}{(Y_j + a)^2} \right] + \frac{n}{a}.$$

We obtain the results in the simulation study by setting $a = 0.25$ as is recommended in [20]. The null hypothesis is rejected for large values of $L_{n,a}$.

2.4 Tests for the Weibull distribution

Below we consider three tests for the Weibull distribution in the presence of random right censoring. The density of the Weibull distribution with parameters κ and λ is $f(t) = \kappa \lambda^\kappa t^{\kappa-1} e^{-(\lambda t)^\kappa}$, $\kappa > 0$, $\lambda > 0$, $t > 0$. The first two tests are based on Stein's method for the approximation of integrals while the final test is based on the Laplace transform of the standard extreme value distribution.

2.4.1 Bothma-Allison-Visagie (2021)

[13] proposes new classes of tests for the Weibull distribution based on Stein's method. This test statistic is defined as

$$S_{n,a} = n \int_{-\infty}^\infty \left| \sum_{j=1}^n \Delta_j \left[i t e^{i t Y_j} + (1 - e^{Y_j}) e^{i t Y_j} \right] \right|^2 w_a(t) dt, \quad (4)$$

where $w_a(t)$ is a weight function containing a user-defined tuning parameter $a > 0$. The use of two weight functions resulted in two different classes of test statistics. For the first weight function, e^{-at^2} , the test statistic is given by

$$S_{n,a}^{(1)} = n \sqrt{\frac{\pi}{a}} \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k e^{-(Y_j - Y_k)^2 / 4a} \left\{ -\frac{1}{4a^2} ((Y_j - Y_k)^2 - 2a) + 2(1 - e^{Y_j}) \left(\frac{1}{2a} \right) (Y_j - Y_k) + (1 - e^{Y_j}) (1 - e^{Y_k}) \right\}.$$

In the simulation study below we use $a = 1$ as well as $a = 2$. The second weight function used is $e^{-a|t|}$. In this case, the test statistic simplifies to

$$S_{n,a}^{(2)} = n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left\{ \frac{-4a(3(Y_j - Y_k)^2 - a^2)}{((Y_j - Y_k)^2 + a^2)^3} + \frac{8a(Y_j - Y_k)(1 - e^{Y_j})}{((Y_j - Y_k)^2 + a^2)^2} + \frac{2a(1 - e^{Y_j})(1 - e^{Y_k})}{(Y_j - Y_k)^2 + a^2} \right\}.$$

We obtain the results in the simulation study by setting $a = 2$. The null hypothesis is rejected for large values of both $S_{n,a}^{(1)}$ and $S_{n,a}^{(2)}$.

2.4.2 Krit (2014)

A test that compares the empirical Laplace transform of a random variable Y to the Laplace transform of an $EV(0, 1)$ random variable was introduced in [21]. The resulting test statistic is

$$KR_n^* = n \int_{-1}^0 [\psi_n(t) - \Gamma(1-t)]^2 e^{\gamma t - e^{\gamma t}} dt, \quad (5)$$

where $\gamma \in \mathbb{R}$ is a user-specified tuning parameter. The following approximation provides an easily implementable form of KR_n^* ;

$$KR_n = n \sum_{k=-m}^{-1} \left[\sum_{j=1}^n \Delta_j e^{-Y_j k/m} - \Gamma(1 - k/m) \right]^2 e^{\gamma k/m - e^{\gamma k/m}},$$

where m is the number of points at which the integrand is evaluated and γ is a user-defined tuning parameter. In the numerical results shown below, we use $\gamma = -5$ and $m = 100$ as recommended in [21].

For a more detailed discussion on the tests above, see [2] as well as [13].

2.5 Tests for the gamma distribution

A random variable with density $f(t) = \lambda^\kappa \Lambda^{-1}(\kappa) t^{\kappa-1} e^{-\lambda t}$, $\kappa > 0$, $\lambda > 0$, $t > 0$, with $\Lambda(\kappa) = \int_0^\infty x^{\kappa-1} e^{-x} dx$, is said to have a gamma distribution with shape parameter κ and rate parameter λ . As was mentioned in the introduction, we have been unable to find published papers on testing for the gamma distribution in the presence of random right censoring. Below we amend three existing tests for this distribution developed for use with full samples, using techniques similar to those used above. The first and second tests below are based on the empirical Laplace transform while the third uses the Stein characterisation. Below, $Y_j = \hat{\lambda} T_j$, $j = 1, \dots, n$ where $\hat{\lambda}$ denotes the estimated rate parameter of the gamma distribution.

2.5.1 Henze-Meintanis-Ebner (2012)

New classes of tests based on a differential equation involving the Laplace transform are proposed in [22] for the full sample case. Two tests are introduced based on two different weight functions. The modified test is defined as

$$M = n \int_0^\infty \left[(1+t) \tilde{\psi}'_n(t) + \kappa \tilde{\psi}_n(t) \right]^2 w_a(t) dt,$$

where $w_a(t)$ is a weight function with a user-defined tuning parameter $a > 0$. If the weight function is, e^{-at} , then the modified version of the test is given by

$$M_{n,a}^{(1)} = n \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left(\frac{Y_j Y_k - 2\kappa Y_j + \kappa^2}{Y_j + Y_k + a} + \frac{2Y_j Y_k - 2\kappa Y_j}{(Y_j + Y_k + a)^2} + \frac{2Y_j Y_k}{(Y_j + Y_k + a)^3} \right)$$

and if the weight function is, e^{-at^2} , then the modified test simplifies to

$$\begin{aligned} M_{n,a}^{(2)} &= \frac{n}{2} \sqrt{\frac{\pi}{a}} \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k (Y_j Y_k - 2\kappa Y_j + \kappa^2) \varphi_{jk}(a) \\ &\quad + \frac{n}{4a} \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k (2Y_j Y_k - 2\kappa Y_j) \left(2 - \sqrt{\frac{\pi}{a}} (Y_j + Y_k) \varphi_{jk}(a) \right) \\ &\quad + \frac{n}{8a^2} \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \left[\left\{ \sqrt{\frac{\pi}{a}} (Y_j + Y_k)^2 + 2\sqrt{\pi a} \right\} \varphi_{jk}(a) - 2(Y_j + Y_k) \right], \end{aligned}$$

where $\varphi_{jk}(a) = \left[1 - \operatorname{erf} \left(\frac{Y_j + Y_k}{2\sqrt{a}} \right) \right] \exp \left(\frac{(Y_j + Y_k)^2}{4a} \right)$ and $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$ denotes the error function. The null hypothesis is rejected for large values for both $M_{n,a}^{(1)}$ and $M_{n,a}^{(2)}$.

2.5.2 Betsch-Ebner (2019)

[23] shows that F is the gamma distribution if, and only if,

$$F(t) = \mathbb{E} \left[\left(-\frac{\kappa - 1}{X} + \frac{1}{\lambda} \right) \min(X, t) \right], t > 0.$$

As a result,

$$\Lambda_n(t) = \sum_{j=1}^n \Delta_j \left(1 + \frac{1 - \hat{\kappa}}{Y_{(j)}} \right) \min(Y_{(j)}, t) - \tilde{F}_n(t)$$

should be approximately zero for all $t > 0$ if the hypothesis of the gamma class is true. We propose the following statistic based on $\Lambda_n(t)$;

$$\begin{aligned} R_{n,a} &= \int_0^\infty \Lambda_n^2(t) e^{-at} d\tilde{F}_n(t) \\ &= \sum_{j=1}^n \Delta_j \Lambda_n^2(Y_{(j)}) e^{-aY_{(j)}}, \end{aligned}$$

where $a > 0$ is a user defined tuning parameter. In the simulations in Section 3, we use $a = 0.5$ and $a = 1$. The hypothesis of the gamma class of distributions is rejected for large values of $R_{n,a}$.

3 Monte Carlo simulation

This section presents empirical sizes and powers of the tests considered the exponential, Weibull and gamma classes of distributions; we consider each of these classes in turn. The results presented include two different censoring proportions; 10% and 30% respectively. All of the results presented in the main text are for samples of size $n = 50$. However, Appendix A shows the results corresponding to a sample size of $n = 100$. Other configurations of sample size and censoring proportion were also considered when performing the numerical analyses. However, for the sake of brevity, we do not present all of these results below.

The numerical results presented below are based on a nominal significance level of 5% throughout. For each lifetime distribution considered, we report the powers obtained using three different censoring distributions; the exponential, uniform and Lindley distributions. The alternative lifetime distributions considered are listed in Table 2. For each of these censoring distributions, the parameters used are chosen so as to ensure that the specified censoring proportion is achieved. In the case of the uniform censoring distribution, this entails choosing the value of the maximum of the distribution while the minimum remains fixed at 0.

For every test considered, we report the empirical powers for the four possible combinations of whether or not the set-to-1 assumption is made. The four variations for each of the tests are defined in Section 1 and indicated by FF , FT , TF and TT . The empirical powers obtained for the various distributions are reported in Tables 20 to 8. Tables 20 and 4 show the results pertaining to testing for exponentiality, while Tables 5 and 6 show the results associated with the Weibull class. The results associated with the gamma class can be found in Tables 7 and 8. The reported powers represent the percentages (rounded to the nearest integer) of 10 000 independent Monte Carlo samples for which the null hypothesis is rejected. For each alternative considered, the empirical powers associated with the three censoring distributions mentioned above are shown in consecutive rows of the table. The order in which the censoring distributions are arranged is first the exponential, then the uniform and finally the Lindley distributions. In order to ease comparisons, the highest power, including ties, in each row is shown in bold. The empirical sizes and powers shown in Tables 9 to 14 in Appendix A make use of the same conventions listed above.

Samples from censored distributions were generated using the *LindleyR* package, see [24], while parameter estimation for the results pertaining to the Weibull and gamma classes were performed using the *parmsurvfit* package, see [11]. The tables were produced using the *Stargazer* package, see [25].

3.1 Numerical results

In this section we discuss the performance of the tests against all of the alternative lifetime distributions considered. We start with a general discussion of the results in Tables 20-8, thereafter we discuss the estimated powers for the three lifetime distributions under consideration separately. Furthermore, we will make recommendations regarding the set-to-1 assumption for the best performing test for each of the three lifetime distributions.

Table 2: Density functions of the alternative distributions.

Alternative	Density	Notation
Lognormal	$(\theta x \sqrt{2\pi})^{-1} \exp(-\log^2(x) (2\theta^2)^{-1})$	$LN(\theta)$
Chi square	$(2^{\theta/2} \Gamma(\theta/2))^{-1} x^{\theta/2-1} \exp(-x/2)$	$\chi^2(\theta)$
Beta	$x^{\alpha-1} (1-x)^{\theta-1} \Gamma(\alpha+\theta) (\Gamma(\alpha)\Gamma(\theta))^{-1}$	$\beta(\alpha, \theta)$
Lindley	$\frac{\theta^2}{\theta+1} (1+x) \exp(-\theta x)$	$\mathcal{L}(\theta)$

First, consider the performance of the tests for all three of the lifetime distributions against all alternatives. In general, the sizes as well as the estimated powers are highly dependent on whether the set-to-1 assumption is made or not. It is important to also note that with an increase in the censoring proportion this dependence becomes more prominent. In some cases, we observe that the specified sizes are achieved for a specific test under certain assumptions, but under a different set of assumptions the same test achieves a size of approximately zero. The same phenomenon is observed for the estimated powers where fluctuating powers are observed for the same test, depending on the assumptions made. Overall, it seems that the tests for the exponential distribution are less sensitive to the set-to-1 assumption. The influence of the censoring distribution becomes more pronounced when the censoring proportion is increased. Again, we observe that the tests for the exponential distribution is less sensitive to the choice of censoring distribution. The empirical powers are observed to increase with sample size and decrease with the censoring proportion.

3.1.1 Power comparison for the exponential distribution

For the exponential distribution, the nominal significance level is achieved for the majority of the tests for a censoring proportion of 10%. For a censoring proportion of 30% these sizes are attained for fewer of the tests considered. $L_{0.25,n}$ is the best performing test against the majority of alternatives, regardless of the set-to-1 assumptions or choice of the censoring distribution. Therefore we recommend using this test when working with a censored sample that is thought to be exponentially distributed.

3.1.2 Power comparison for the Weibull distribution

For the Weibull distribution, the specified sizes are attained for most of the tests for smaller censoring proportions, but as the censoring proportion increases the sizes as well as the estimated powers are more variable. The choice of the censoring distribution strongly influences the estimated powers for a censoring proportion of 30%, especially against alternatives such as the chi-squared or beta distributions. The largest differences in power occur when the censoring distribution is specified to be the uniform distribution. For a censoring proportion of 10% the $S_{n,2}^{(1)}$ test outperforms the other tests for the majority of the alternatives, but only for the FT assumption, i.e. the case where the set-to-1 assumption is made for the lifetime distribution but not for the censoring distribution. The $S_{n,2}^{(2)}$ test performs the best overall against the alternatives considered. Once again, this is only for a specific set of assumptions, in this case TT; i.e., the set-to-1 assumption is made for both of the estimated distributions. Even though these two tests perform well under a specific set of assumptions their powers are noticeably erratic for some of the alternatives. The $S_{n,1}^{(1)}$ test performs well with competitive estimated powers regardless of the censoring proportion and set-to-1 assumptions. As a result, we recommend using the $S_{n,2}^{(1)}$ test in general. However, $S_{n,2}^{(2)}$ tests can also be used for both censoring proportions considered.

3.1.3 Power comparison for the gamma distribution

For the gamma distribution the specified sizes are mostly attained for the 10% censoring proportion, but is rarely attained for the 30% censoring proportion. For the majority of alternatives most of the tests perform poorly, exhibiting low estimated powers, especially when the censoring proportion is high. The influence of the censoring distribution increases with an increase in the censoring proportion, specifically for alternatives such as the lognormal distribution.

When working with a censored sample with a low censoring proportion we recommend using the $M_{n,1}^{(1)}$ test along with the FT assumption, i.e. the case where the set-to-1 assumption is made for the lifetime distribution but not for the censoring distribution. If the censoring proportion is high, then we recommend using the CV_n with the FT assumption. Overall, the CV_n test exhibit the most stable powers as it is the least affected by the assumptions made.

Table 3: Estimated powers for the exponential distribution for 10% censoring for a sample size of $n = 50$ with three different censoring distributions.

F	KS_n				CV_n				EP_n				$L_{n,0.25}$				$B_{n,0.25}$				$H_{n,1}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$exp(0.5)$	5	5	4	5	5	5	5	5	5	5	6	5	5	5	5	5	4	5	2	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	6	5	5	5	5	5	4	5	2	5	5	5	5	4
	5	5	4	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	2	5	5	5	5	4
$exp(1)$	5	5	5	5	5	5	5	5	5	5	6	5	5	5	5	5	5	5	2	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	2	5	5	5	5	4
	5	5	4	5	5	5	5	5	5	4	5	5	5	5	5	5	4	5	2	5	4	4	5	4
$exp(2)$	5	5	5	5	5	5	5	5	5	5	6	5	5	5	5	5	4	5	2	5	5	5	5	4
	5	5	5	5	5	5	5	5	5	5	6	5	5	5	5	5	5	5	2	5	5	5	5	5
	5	5	4	5	5	5	5	5	5	5	6	5	5	5	5	5	4	5	2	5	5	4	5	5
$\Gamma(0.6)$	55	55	55	55	62	62	62	62	62	62	62	62	82	82	82	82	79	79	76	79	56	56	56	56
	54	54	55	55	62	62	62	62	63	63	64	63	82	82	82	82	79	79	76	79	56	56	56	55
	55	55	55	56	62	62	62	62	63	63	63	63	82	82	82	82	79	79	76	80	56	56	56	56
$\Gamma(0.8)$	14	14	14	14	16	16	16	16	18	18	18	18	27	27	27	27	24	25	19	25	14	14	15	14
	13	14	13	14	16	16	16	17	18	18	19	18	27	27	27	27	24	25	19	25	14	14	15	14
	14	14	14	14	16	16	17	17	18	18	19	18	27	27	28	28	24	25	19	26	14	14	15	14
$\Gamma(1.2)$	10	10	9	11	11	12	12	12	11	11	13	12	11	11	11	11	11	12	4	13	11	11	12	10
	10	10	9	11	12	12	11	12	12	11	13	11	12	12	11	11	10	13	3	13	11	10	11	10
	10	10	9	11	11	11	11	12	11	11	14	12	11	11	11	11	9	12	3	13	10	10	11	10
$W(0.8)$	32	32	32	33	38	38	38	38	43	43	43	43	50	50	49	49	49	49	44	50	32	32	32	32
	31	32	32	32	37	37	37	37	42	42	43	42	49	49	48	48	49	49	44	49	31	31	32	32
	31	31	31	31	37	37	37	37	42	42	42	42	49	49	48	48	48	49	43	49	31	32	32	32
$W(1.2)$	20	20	17	21	25	25	24	25	26	26	30	27	23	23	23	23	23	27	7	27	23	22	23	20
	20	20	17	22	25	25	24	26	26	25	31	28	22	22	23	23	22	26	7	28	22	22	23	21
	19	20	16	21	24	25	24	25	26	25	30	27	23	23	23	23	21	26	6	27	22	22	23	20
$LN(1)$	20	21	20	21	24	24	25	25	12	12	14	13	27	27	26	27	24	24	15	24	16	16	16	15
	19	19	19	20	23	23	23	24	11	11	11	11	27	27	27	27	24	23	15	23	15	16	16	15
	19	19	18	19	23	23	23	23	11	11	11	10	28	28	28	29	23	24	14	24	15	15	16	15
$LN(1.5)$	82	82	82	82	86	86	86	86	90	90	90	90	70	70	70	70	85	84	84	84	77	78	77	77
	81	81	81	81	86	86	86	86	89	89	90	90	68	68	68	68	84	83	84	83	76	77	77	77
	80	80	80	80	85	85	85	85	89	89	88	88	66	66	66	66	83	82	82	81	76	77	75	76
$\chi^2(1)$	82	82	82	82	87	87	87	87	86	86	86	86	97	97	97	97	96	96	95	96	82	82	82	82
	82	82	82	82	87	87	87	87	86	86	87	87	97	97	97	97	96	96	95	96	82	82	82	82
	82	82	82	82	87	87	87	87	86	87	86	86	97	97	96	96	96	96	95	96	82	82	82	82
$\chi^2(3)$	32	33	27	34	40	40	39	40	40	39	45	41	45	45	47	47	44	49	17	50	37	37	38	34
	32	33	27	35	40	40	39	41	40	39	46	42	46	46	46	46	43	48	16	50	37	37	38	34
	31	33	25	35	39	39	38	40	39	38	45	40	45	45	46	46	37	48	13	49	36	35	37	33
$\beta(1,1)$	88	89	5	93	97	97	80	85	96	95	93	87	54	54	56	57	82	83	5	76	97	97	95	82
	88	89	4	93	97	97	79	84	95	95	93	87	54	54	56	57	82	83	5	76	97	97	94	81
	88	89	4	93	97	97	78	83	95	95	93	85	54	54	56	57	81	83	4	75	97	97	95	82
$\beta(0.5,1)$	35	35	25	37	47	47	44	46	7	7	10	8	69	69	69	70	60	61	33	61	48	48	48	43
	35	35	25	38	46	46	43	46	8	8	10	8	70	70	70	70	60	61	33	62	48	48	48	43
	35	35	24	37	46	46	43	45	8	8	10	8	69	69	69	69	60	61	33	61	48	48	48	42
$\beta(1,1.5)$	59	60	12	66	76	76	63	68	78	78	83	76	33	33	34	34	53	56	4	54	70	70	67	52
	57	59	11	65	75	75	62	67	77	77	82	74	33	33	33	34	51	54	3	53	70	69	67	52
	59	60	11	64	75	75	62	66	77	76	82	74	32	32	33	33	51	55	3	52	70	70	66	51

Table 4: Estimated powers for the exponential distribution for 30% censoring for a sample size of $n = 50$ with three different censoring distributions.

F	KS_n				CV_n				EP_n				$L_{n,0.25}$				$B_{n,0.25}$				$H_{n,1}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$exp(0.5)$	2	3	1	4	3	4	3	3	3	3	4	3	5	5	5	5	3	5	1	5	4	4	5	4
	2	3	1	4	3	3	2	3	2	2	2	2	5	5	5	5	3	4	2	4	3	3	4	3
	2	3	1	4	3	3	2	3	3	2	2	2	5	5	5	5	3	4	1	4	4	4	5	4
$exp(1)$	2	3	1	4	3	4	3	3	3	3	4	3	5	5	5	5	3	4	1	5	4	4	5	4
	2	3	1	4	3	3	2	3	2	2	2	2	5	5	5	5	4	4	2	4	4	3	4	4
	2	3	1	4	3	3	2	3	3	3	3	2	5	5	5	5	3	4	1	5	4	4	5	4
$exp(2)$	2	3	1	4	3	4	2	3	3	3	4	3	5	5	5	5	3	4	1	5	4	4	5	4
	2	3	1	4	3	3	2	2	2	2	2	2	5	5	5	5	4	4	2	4	3	3	4	3
	2	3	1	4	3	3	2	3	3	3	3	3	5	5	5	5	3	4	1	5	4	3	5	3
$\Gamma(0.6)$	25	32	22	34	46	46	44	46	42	44	43	44	77	77	76	76	62	71	54	72	39	41	41	40
	16	33	14	34	43	43	41	42	32	46	32	46	76	76	76	76	58	69	55	69	38	41	39	41
	24	33	20	34	46	46	44	46	41	44	41	44	76	76	77	77	61	71	54	71	39	40	41	41
$\Gamma(0.8)$	5	7	3	8	11	11	10	11	11	12	12	11	24	24	25	25	15	21	9	21	10	11	12	11
	4	7	3	8	9	8	8	8	7	11	6	11	24	24	24	24	16	18	13	18	9	11	10	11
	4	7	3	8	10	11	9	11	11	12	11	11	25	25	25	25	15	20	9	21	11	11	12	11
$\Gamma(1.2)$	2	5	1	6	7	7	4	7	6	5	7	4	10	10	10	10	2	10	0	11	7	6	8	5
	1	5	0	6	4	6	3	5	5	3	4	2	10	10	10	10	1	11	1	10	6	4	8	4
	1	5	0	6	5	7	4	6	5	4	5	3	10	10	10	11	1	10	0	11	6	5	8	5
$W(0.8)$	11	16	9	17	23	23	22	23	25	27	25	26	42	42	42	42	32	39	25	40	21	21	22	22
	8	14	6	15	20	19	19	19	15	24	15	23	39	39	39	39	30	33	27	33	18	21	19	21
	9	15	7	16	22	22	20	22	22	25	22	24	41	41	40	40	29	37	23	38	20	21	21	21
$W(1.2)$	4	9	1	12	13	15	8	12	12	10	12	8	19	19	19	19	4	20	0	20	12	11	15	9
	1	9	0	12	8	12	5	10	9	6	7	4	18	18	18	19	1	20	0	20	12	8	14	7
	2	8	0	11	11	13	7	11	11	9	10	6	19	19	19	20	3	20	0	20	11	9	15	9
$LN(1)$	5	9	3	11	13	14	12	14	8	7	8	6	34	34	34	34	10	25	4	26	10	10	12	10
	2	7	2	9	9	10	7	10	6	4	6	3	38	38	38	38	4	27	3	28	9	8	11	8
	3	8	1	9	11	12	9	12	6	5	7	5	37	37	37	37	7	26	3	27	10	9	12	9
$LN(1.5)$	35	41	33	42	53	52	53	55	58	55	58	36	36	36	36	52	53	49	53	42	46	43	46	
	21	33	20	33	42	40	43	41	29	45	29	46	25	25	25	25	46	34	45	36	35	42	36	43
	25	35	23	36	47	47	46	47	42	48	42	48	30	30	29	29	47	46	43	45	38	44	39	44
$\chi^2(1)$	51	59	50	61	74	74	73	74	65	68	65	67	95	94	94	94	88	92	84	92	65	67	67	67
	36	61	35	63	71	72	71	72	55	71	57	72	94	94	94	94	83	91	82	91	64	67	66	68
	47	59	44	60	73	74	72	74	63	68	63	67	94	94	94	94	86	92	82	92	64	66	66	66
$\chi^2(3)$	6	15	1	19	23	25	15	21	18	16	18	13	40	40	41	42	12	39	1	40	21	18	24	16
	1	15	0	20	15	21	9	18	16	10	13	8	40	40	40	40	1	39	0	38	21	15	24	14
	2	12	0	18	17	21	9	16	14	11	11	7	38	38	39	39	3	37	0	37	19	14	23	13
$\beta(1,1)$	66	77	0	90	86	86	18	30	54	53	8	5	42	42	43	45	56	67	0	43	87	86	85	57
	46	74	0	90	78	79	11	23	40	38	3	1	39	39	40	42	36	61	0	38	83	79	82	53
	65	77	0	90	85	85	16	29	51	50	7	4	42	42	43	44	53	66	0	43	87	85	84	56
$\beta(0.5,1)$	16	24	4	29	36	37	27	35	9	9	10	8	72	72	73	73	47	61	23	62	34	32	39	30
	6	22	2	29	29	32	21	30	13	16	12	13	75	75	75	75	37	62	24	63	30	28	36	29
	15	24	4	28	36	37	26	34	9	10	9	9	72	72	72	73	46	61	23	62	33	32	39	30
$\beta(1,1.5)$	20	35	0	49	46	47	14	25	32	30	18	9	24	24	25	25	18	37	0	30	45	41	48	27
	6	31	0	47	34	37	8	17	21	18	6	4	22	22	21	22	4	32	0	26	38	32	43	23
	17	34	0	48	44	46	12	23	30	28	14	8	23	23	23	24	15	37	0	29	44	40	46	25

Table 5: Estimated powers for the Weibull distribution for 10% censoring for a sample size of $n = 50$ with three different censoring distributions.

F	KS_n				CV_n				KR_n				$S_{n,1}^{(1)}$				$S_{n,2}^{(1)}$				$S_{n,2}^{(2)}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$W(0.5)$	3	4	2	4	4	5	4	4	2	1	0	1	3	3	3	3	3	2	1	1	3	3	1	3
	3	4	2	5	4	4	4	5	2	1	1	0	4	4	4	4	3	1	2	1	3	2	2	2
	3	4	2	5	4	4	4	4	2	1	1	0	3	3	3	3	3	1	2	1	3	2	2	2
$W(1)$	5	5	4	4	5	5	5	5	2	2	0	1	4	4	4	4	4	0	3	4	5	0	5	5
	4	4	4	4	4	4	4	4	2	1	0	0	3	3	3	3	3	4	0	2	3	4	0	4
	4	5	4	5	5	5	5	5	2	2	0	0	3	3	3	3	3	3	0	2	3	4	0	4
$W(2)$	5	5	4	5	5	5	4	4	4	3	0	0	4	4	4	4	5	5	0	3	4	5	0	5
	5	5	5	5	5	5	5	5	4	2	0	0	5	5	5	5	4	5	0	3	4	5	0	5
	4	4	4	5	5	5	5	5	4	2	0	0	4	4	4	4	4	4	0	3	4	5	0	5
$\Gamma(3)$	9	9	10	10	11	11	11	12	8	9	0	6	13	13	13	13	14	14	1	13	12	13	3	14
	9	9	8	8	11	11	11	11	7	9	0	5	13	13	13	13	13	14	1	12	13	13	3	13
	9	10	9	9	11	11	11	11	6	7	0	4	12	12	12	12	11	13	1	11	11	13	3	13
$\Gamma(5)$	13	13	12	13	16	16	16	16	14	15	0	10	21	21	21	21	21	23	1	22	19	20	6	20
	13	13	13	13	16	16	16	17	13	14	0	10	21	21	21	21	21	23	2	22	18	19	7	21
	13	13	12	12	16	16	16	16	6	13	0	9	19	19	19	19	18	22	2	20	17	20	6	20
$\Gamma(10)$	17	17	17	17	23	23	22	22	22	23	0	18	32	32	32	32	33	34	4	32	28	29	12	28
	18	18	18	18	23	23	23	23	21	24	0	19	31	31	31	31	32	33	4	32	28	28	13	29
	16	16	17	17	21	21	22	22	16	20	0	14	29	29	29	29	29	32	3	30	26	27	10	28
$LN(0.5)$	31	31	32	32	45	45	44	44	33	40	0	37	55	55	55	55	56	60	20	60	51	52	35	53
	32	32	33	33	43	43	44	44	28	40	0	35	55	55	55	55	56	60	17	59	50	52	34	53
	30	30	30	31	41	41	42	42	15	37	1	35	52	52	52	52	50	57	17	56	47	50	31	52
$LN(1)$	30	30	29	29	42	42	40	40	9	31	2	29	50	50	50	50	41	56	18	54	45	49	31	50
	29	29	30	31	41	41	41	41	8	28	3	25	45	45	45	45	33	53	18	52	41	48	27	49
	27	27	27	28	37	38	39	39	8	24	4	23	40	40	40	40	30	47	16	45	37	43	23	46
$LN(2)$	24	25	24	25	37	36	36	36	9	15	6	14	29	29	29	29	23	37	18	36	26	38	21	39
	22	25	23	26	35	34	35	35	11	4	9	4	26	26	26	26	25	25	23	26	24	30	22	32
	21	23	19	23	32	32	33	33	9	6	7	5	24	24	24	24	23	20	19	18	22	28	18	28
$\chi^2(6)$	9	9	10	10	11	11	12	12	8	9	0	6	13	13	13	13	14	14	1	13	12	13	3	14
	9	9	8	8	11	11	11	11	7	9	0	5	13	13	13	13	13	14	1	12	13	13	3	13
	9	9	9	9	11	11	10	10	5	7	0	4	12	12	12	12	10	13	1	11	11	13	2	13
$\chi^2(8)$	11	11	11	11	13	13	13	13	11	12	0	8	17	17	17	17	16	19	1	17	16	16	5	16
	11	11	11	11	14	14	14	15	9	12	0	8	17	17	17	17	16	19	1	17	15	17	5	17
	11	11	11	11	14	14	14	14	7	10	0	7	16	16	16	16	15	18	1	17	15	16	4	18
$\chi^2(10)$	13	13	12	13	16	16	16	16	14	15	0	10	21	21	21	21	21	23	1	22	18	20	6	20
	13	13	13	13	16	16	16	17	13	14	0	10	21	21	21	21	21	23	2	22	18	19	7	21
	13	13	13	13	16	16	16	16	7	13	0	9	19	19	19	19	18	21	2	20	17	19	6	20
$\beta(1,1)$	49	50	10	52	63	63	49	52	46	47	0	0	66	66	66	66	69	70	0	5	66	68	0	24
	48	49	8	53	64	64	47	51	43	44	0	0	67	67	67	67	70	71	0	5	67	69	0	24
	48	48	9	52	61	61	47	52	45	46	0	0	66	66	66	66	68	69	0	5	65	67	0	23
$\beta(0.5,1)$	44	46	4	51	60	60	40	45	24	33	0	0	61	61	61	61	59	65	0	3	61	65	0	18
	43	46	3	51	60	60	38	44	19	28	0	0	58	58	58	58	56	63	0	3	57	64	0	17
	45	47	3	52	60	61	40	45	24	33	0	0	58	58	58	58	58	64	0	3	57	65	0	18
$\beta(1,2)$	15	16	10	16	19	19	18	19	6	9	0	0	19	19	19	19	18	24	0	7	18	21	0	14
	15	15	10	16	19	19	17	18	4	7	0	0	17	17	17	17	16	23	0	5	15	21	0	13
	14	14	9	16	17	17	17	19	4	7	0	0	17	17	17	17	15	22	0	5	15	20	0	12

Table 6: Estimated powers for the Weibull distribution for 30% censoring for a sample size of $n = 50$ with three different censoring distributions.

F	KS_n				CV_n				KR_n				$S_{n,1}^{(1)}$				$S_{n,2}^{(1)}$				$S_{n,2}^{(2)}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$W(0.5)$	2	3	1	5	2	1	1	1	2	1	1	1	4	4	4	4	4	1	3	1	4	1	3	1
	2	4	2	5	2	1	2	1	2	1	2	1	5	5	5	5	5	1	5	1	5	1	5	1
	2	4	2	6	2	1	2	1	2	1	2	1	4	4	4	4	4	1	3	1	4	1	3	1
$W(1)$	2	3	1	3	3	3	2	2	2	1	0	0	3	3	3	3	3	1	1	1	3	2	1	2
	2	3	1	3	2	1	2	1	2	1	1	0	4	4	4	4	4	1	3	1	4	1	3	1
	1	3	1	4	2	2	1	2	2	1	1	1	3	3	3	3	3	1	1	1	3	2	1	2
$W(2)$	3	4	2	4	4	4	3	4	2	1	0	0	3	3	3	3	2	3	0	2	3	4	0	4
	3	4	2	4	4	4	3	4	1	1	0	0	3	3	3	3	3	2	0	1	3	3	1	3
	3	4	2	4	4	4	3	4	2	1	0	0	3	3	3	3	3	3	0	1	3	3	0	3
$\Gamma(3)$	6	7	5	7	8	8	8	8	3	2	0	1	8	8	8	8	7	8	2	7	7	9	2	10
	5	6	4	7	7	8	7	7	4	1	1	1	7	7	7	7	7	4	4	4	7	7	3	7
	5	6	4	6	8	8	7	7	2	2	1	1	6	6	6	6	5	6	2	5	6	8	3	8
$\Gamma(5)$	8	9	8	9	11	11	11	12	5	3	0	1	13	13	13	13	11	14	2	14	12	14	4	15
	8	8	7	8	11	11	10	11	4	2	1	1	11	11	11	11	10	12	3	10	10	13	4	13
	8	9	7	9	11	11	11	11	3	3	1	2	12	12	12	12	9	13	3	11	11	14	4	14
$\Gamma(10)$	13	13	12	13	17	17	17	17	7	4	0	2	21	21	21	21	19	23	2	21	19	21	7	22
	12	13	12	13	17	17	17	17	7	4	0	2	20	20	20	20	16	22	3	22	18	21	7	22
	10	11	11	12	15	15	16	16	6	3	0	2	18	18	18	18	14	20	3	19	16	20	6	21
$LN(0.5)$	22	22	21	22	31	31	31	31	6	6	1	5	35	35	35	35	26	40	10	39	31	39	19	40
	17	18	16	19	27	27	28	29	8	5	3	4	28	28	28	28	21	30	12	30	24	31	15	34
	18	20	19	21	28	28	28	29	5	6	2	5	29	29	29	29	20	33	10	34	26	34	16	36
$LN(1)$	11	14	9	15	23	22	20	22	7	5	4	5	19	19	19	19	16	17	12	17	17	23	13	25
	6	10	6	12	16	13	15	15	8	0	7	1	17	17	17	17	19	1	17	1	16	5	14	5
	6	12	6	12	17	18	16	18	6	4	4	4	16	16	16	16	15	11	12	11	14	16	12	18
$LN(2)$	6	8	5	9	10	7	8	6	8	1	5	1	16	16	16	16	17	2	15	2	14	4	13	4
	6	5	5	6	7	2	6	1	6	1	5	1	19	19	19	19	21	1	21	1	18	1	17	1
	5	6	6	9	7	2	7	3	6	0	6	0	16	16	16	16	17	1	17	1	15	1	14	1
$\chi^2(6)$	6	7	6	7	8	8	8	8	3	2	0	1	8	8	8	8	7	8	2	7	7	9	2	10
	5	6	4	7	7	8	7	7	4	1	1	1	7	7	7	7	7	4	4	4	7	7	3	7
	4	6	3	6	7	7	6	7	3	2	1	1	6	6	6	6	5	6	3	5	6	8	3	8
$\chi^2(8)$	7	8	7	8	10	10	10	10	4	2	0	1	11	11	11	11	9	12	2	11	10	13	3	13
	7	8	6	8	9	9	9	9	3	2	1	1	9	9	9	9	8	8	3	7	9	10	4	11
	5	7	5	7	9	9	9	9	3	3	1	2	8	8	8	8	7	9	2	8	8	12	3	12
$\chi^2(10)$	8	9	9	9	11	11	11	12	5	3	0	1	13	13	13	13	11	14	2	14	12	14	4	15
	8	8	7	8	11	11	10	11	4	2	1	1	11	11	11	11	10	12	3	10	10	13	4	13
	8	9	7	9	11	11	10	11	4	3	1	2	11	11	11	11	8	12	2	10	9	13	4	14
$\beta(1,1)$	26	30	1	39	42	42	17	23	3	7	0	0	30	30	30	30	23	39	0	1	29	43	0	9
	21	27	0	39	36	37	8	16	0	5	0	0	17	17	17	17	3	26	0	1	14	34	0	5
	24	30	0	38	40	40	15	21	2	6	0	0	27	27	27	27	13	34	0	1	25	39	0	7
$\beta(0.5,1)$	6	21	0	39	21	24	1	4	0	5	0	0	3	3	3	3	0	8	0	1	3	17	0	3
	1	19	0	41	2	7	0	2	0	2	0	0	1	1	1	1	0	2	0	1	1	5	0	2
	4	20	0	41	16	19	1	4	0	3	0	0	3	3	3	3	0	6	0	1	2	12	0	3
$\beta(1,2)$	4	8	1	11	9	10	5	7	1	2	0	0	2	2	2	2	1	6	0	1	2	8	0	4
	2	7	0	10	5	7	2	4	0	1	0	0	1	1	1	1	0	3	0	1	1	5	0	3
	4	7	0	10	8	9	4	6	0	1	0	0	2	2	2	2	0	5	0	1	1	7	0	4

Table 7: Estimated powers for the gamma distribution for 10% censoring for a sample size of $n = 50$ with three different censoring distributions.

F	KS_n				CV_n				$M_{n,1}^{(1)}$				$M_{n,4}^{(2)}$				$R_{n,0.5}$				$R_{n,1}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$\Gamma(0.5)$	4	5	4	5	5	5	5	5	3	4	0	4	2	2	0	1	2	1	5	1	2	1	5	1
	4	5	4	5	5	5	5	5	4	4	1	4	4	2	1	1	2	2	5	1	2	2	5	1
	4	4	3	5	5	5	4	5	3	4	0	4	3	2	0	1	3	2	6	1	3	2	6	1
$\Gamma(1)$	4	5	5	5	5	5	5	5	3	4	0	4	4	4	0	2	2	1	4	0	2	1	4	0
	4	5	4	5	5	5	5	5	4	4	0	4	4	3	0	2	2	1	4	0	2	1	4	0
	5	5	4	5	5	5	5	5	3	4	0	4	3	4	0	2	2	1	4	1	2	1	4	1
$\Gamma(2)$	5	5	4	5	5	5	5	5	4	4	0	3	5	4	0	2	2	1	3	0	2	1	3	0
	5	5	5	5	5	5	5	5	5	4	0	3	5	4	0	2	2	1	4	0	2	1	4	0
	5	5	4	5	5	5	5	5	4	5	0	3	4	4	0	2	2	1	4	0	2	1	4	0
$W(2)$	10	10	9	11	12	13	12	12	14	18	0	7	8	14	0	3	1	1	7	0	1	1	7	0
	11	11	9	12	13	13	13	13	14	19	0	7	6	14	0	3	1	1	6	0	1	1	6	0
	11	12	9	11	13	13	12	12	10	18	0	7	4	12	0	2	1	1	6	0	1	1	6	0
$W(3)$	17	17	13	18	21	21	19	20	28	30	0	4	18	21	1	3	1	1	8	0	1	1	8	0
	17	17	13	18	21	21	21	21	27	30	0	5	15	20	1	3	2	1	9	0	2	1	9	0
	17	17	13	19	21	21	20	21	24	29	0	4	13	19	1	4	1	1	7	0	1	1	7	0
$LN(1)$	34	34	33	33	41	41	41	41	39	53	14	52	31	53	10	51	2	2	3	2	2	2	3	2
	31	31	32	33	39	39	40	40	34	50	16	50	28	50	14	51	3	3	4	3	3	3	4	3
	31	31	31	31	38	38	38	38	29	49	14	49	24	49	12	49	4	4	5	3	4	4	5	3
$LN(1.5)$	54	54	53	53	66	66	64	64	58	75	40	75	39	67	25	67	3	4	3	4	3	4	3	4
	50	51	51	52	62	62	63	63	51	73	41	74	40	64	32	64	4	6	4	6	4	6	4	6
	49	50	48	49	61	61	61	61	42	71	34	71	32	58	25	58	5	7	5	6	5	7	5	6
$LN(2.5)$	77	78	77	78	87	87	88	88	60	84	54	86	41	53	37	53	4	6	4	6	4	6	4	6
	75	77	75	77	87	87	86	86	59	82	56	82	46	42	43	41	3	9	4	9	3	9	4	9
	70	75	69	74	84	84	84	84	51	77	48	77	40	32	37	31	5	9	5	9	5	9	5	9
$\chi^2(0.5)$	4	5	3	5	5	5	4	4	3	3	1	2	2	1	1	0	3	2	6	1	3	2	6	1
	3	4	3	5	4	4	4	4	4	2	2	2	4	1	2	0	3	2	7	1	3	2	7	1
	4	5	3	5	5	5	4	4	3	2	1	2	3	1	1	0	3	2	7	1	3	2	7	1
$\chi^2(0.75)$	4	4	3	5	5	5	4	5	3	4	1	4	3	1	1	1	3	2	6	1	3	2	6	1
	4	5	3	5	4	4	4	5	3	3	1	3	3	1	1	0	2	2	6	1	2	2	6	1
	4	5	4	6	5	5	5	5	3	3	1	3	3	1	1	0	3	2	7	1	3	2	7	1
$\beta(1,1)$	66	66	4	71	81	81	54	60	75	76	0	5	74	75	0	0	2	1	93	0	2	1	93	0
	62	64	3	70	80	80	49	58	74	75	0	4	74	74	0	0	1	1	93	0	1	1	93	0
	66	66	3	69	81	81	53	59	76	77	0	4	75	76	0	0	1	1	93	0	1	1	93	0
$\beta(0.5,1)$	37	38	7	42	50	50	37	40	44	50	0	15	33	41	0	0	1	0	55	0	1	0	55	0
	37	39	6	42	51	52	35	40	41	49	0	12	27	37	0	0	1	0	54	0	1	0	54	0
	36	38	6	42	50	51	34	41	42	50	0	13	29	37	0	0	1	1	55	0	1	1	55	0
$\beta(1,2)$	21	22	11	25	29	29	24	27	25	32	0	12	23	32	0	2	1	1	20	0	1	1	20	0
	22	23	11	25	29	29	24	27	23	31	0	10	21	31	0	2	1	0	22	0	1	0	22	0
	21	22	10	24	27	27	23	25	21	31	0	10	20	30	0	2	1	1	20	0	1	1	20	0
$\mathcal{L}(0.25)$	7	7	6	8	7	7	7	8	7	11	0	7	5	9	0	4	1	1	4	0	1	1	4	0
	7	7	6	7	8	8	8	8	5	11	0	8	4	8	0	4	1	1	4	0	1	1	4	0
	7	7	5	7	7	7	7	7	2	9	0	7	1	6	0	3	2	1	5	0	2	1	5	0
$\mathcal{L}(0.5)$	7	7	6	8	8	8	7	8	5	9	0	7	4	7	0	4	1	1	4	0	1	1	4	0
	6	6	5	6	7	7	7	7	4	10	0	7	3	7	0	3	2	1	4	0	2	1	4	0
	6	7	5	7	7	7	7	7	2	9	0	7	1	6	0	3	2	1	4	0	2	1	4	0

Table 8: Estimated powers for the gamma distribution for 30% censoring for a sample size of $n = 50$ with three different censoring distributions.

F	KS_n				CV_n				$M_{n,1}^{(1)}$				$M_{n,4}^{(2)}$				$R_{n,0.5}$				$R_{n,1}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$\Gamma(0.5)$	1	3	1	4	1	2	1	1	2	1	2	1	3	1	2	1	4	2	8	2	4	2	8	2
	2	4	1	5	2	1	1	1	3	1	3	1	3	1	3	1	4	3	8	3	4	3	8	3
	1	3	1	5	2	2	1	1	3	1	1	1	3	1	1	0	4	2	9	2	4	2	9	2
$\Gamma(1)$	2	3	1	4	3	3	2	3	2	2	1	1	2	1	1	1	3	2	6	1	3	2	6	1
	2	3	1	3	2	2	1	1	3	1	2	0	3	1	2	0	4	2	8	2	4	2	8	2
	1	3	1	4	2	2	1	2	2	1	1	1	2	1	1	0	4	2	7	2	4	2	7	2
$\Gamma(2)$	3	4	2	4	4	4	3	4	2	2	0	1	2	1	0	1	3	2	5	1	3	2	5	1
	2	3	1	4	3	4	2	3	2	1	1	0	2	0	1	0	3	2	6	1	3	2	6	1
	2	3	1	4	3	3	3	3	2	1	1	1	2	1	1	0	3	2	6	1	3	2	6	1
$W(2)$	5	7	2	8	8	9	6	7	1	7	0	1	1	3	0	0	1	1	5	0	1	1	5	0
	4	7	1	8	8	9	6	8	1	3	0	1	1	1	0	0	2	1	6	0	2	1	6	0
	5	7	2	8	8	9	6	8	1	5	0	1	1	1	0	0	2	1	6	0	2	1	6	0
$W(3)$	12	13	6	14	16	16	13	15	3	13	0	1	3	7	1	3	1	1	5	0	1	1	5	0
	10	12	4	13	15	15	11	13	0	7	0	0	2	5	1	3	1	0	5	0	1	0	5	0
	10	11	3	13	14	15	10	13	1	10	0	0	2	5	1	3	1	1	5	0	1	1	5	0
$LN(1)$	12	14	10	15	20	21	20	21	13	20	8	20	12	14	8	13	5	5	6	5	5	5	6	5
	8	11	7	12	17	15	15	15	16	2	14	2	16	1	14	1	5	8	6	8	5	8	6	8
	9	13	7	14	18	18	16	18	13	13	10	14	13	8	9	8	6	6	7	5	6	6	7	5
$LN(1.5)$	10	19	9	19	29	29	25	27	19	22	16	22	18	9	15	9	6	8	8	8	6	8	8	8
	8	13	7	15	17	13	15	13	22	1	20	1	21	1	19	1	5	10	6	10	5	10	6	10
	7	14	7	16	19	18	19	20	19	8	18	9	18	3	17	4	7	9	8	10	7	9	8	10
$LN(2.5)$	9	18	8	19	25	25	21	22	27	5	25	4	24	2	21	2	6	12	7	12	6	12	7	12
	7	17	7	17	13	4	11	4	27	1	28	1	19	1	19	1	4	12	4	12	4	12	4	12
	8	14	7	14	15	10	12	9	28	1	26	1	21	1	19	1	6	12	7	13	6	12	7	13
$\chi^2(0.5)$	2	4	1	5	2	1	1	1	3	1	3	1	3	1	2	1	5	3	10	3	5	3	10	3
	2	5	1	6	2	1	1	1	3	1	3	1	3	1	2	1	4	2	7	3	4	2	7	3
	2	4	1	6	2	1	2	1	3	1	2	1	2	1	2	1	5	3	10	3	5	3	10	3
$\chi^2(0.75)$	2	3	1	5	2	1	1	1	4	1	2	1	4	1	2	1	4	3	10	3	4	3	10	3
	2	4	1	5	2	1	1	1	4	1	3	1	3	1	3	1	4	3	7	2	4	3	7	2
	2	3	1	5	2	1	1	1	3	1	2	1	3	1	2	1	5	3	9	3	5	3	9	3
$\beta(1,1)$	36	43	0	58	57	58	12	19	20	40	0	0	6	25	0	0	0	0	34	0	0	0	34	0
	24	42	0	57	50	51	7	15	0	21	0	0	0	7	0	0	1	0	32	0	1	0	32	0
	35	44	0	56	55	56	9	18	12	38	0	0	2	22	0	0	0	0	31	0	0	0	31	0
$\beta(0.5,1)$	4	16	0	28	15	17	2	6	0	4	0	0	0	2	0	0	1	0	19	0	1	0	19	0
	1	13	0	29	3	6	0	2	0	1	0	0	0	1	0	0	1	0	24	0	1	0	24	0
	3	15	0	27	15	17	2	5	0	4	0	0	0	2	0	0	1	0	19	0	1	0	19	0
$\beta(1,2)$	6	10	0	14	13	14	5	10	0	8	0	1	0	2	0	0	1	1	11	0	1	1	11	0
	1	9	0	13	7	9	2	7	0	2	0	1	0	1	0	0	2	0	12	0	2	0	12	0
	4	9	0	12	11	12	4	8	0	6	0	1	0	1	0	0	1	1	9	0	1	1	9	0
$\mathcal{L}(0.25)$	2	4	1	5	5	5	3	4	1	4	0	2	1	1	0	0	2	1	6	1	2	1	6	1
	1	4	1	5	3	4	2	4	2	2	1	1	2	0	1	0	3	2	6	1	3	2	6	1
	1	3	0	4	3	4	2	3	1	2	0	1	1	1	0	0	3	2	7	1	3	2	7	1
$\mathcal{L}(0.5)$	2	4	1	5	5	5	3	5	1	3	0	2	1	1	0	0	2	1	5	1	2	1	5	1
	1	4	0	5	3	4	2	3	2	1	1	1	2	1	1	0	3	1	7	1	3	1	7	1
	1	3	0	4	3	3	1	3	1	2	0	1	1	1	0	0	3	2	8	1	3	2	8	1

4 Conclusions and further research

The numerical results above indicate that the set-to-1 assumption required when using the Kaplan-Meier estimator of the distribution function plays an unexpectedly important role when testing goodness-of-fit hypotheses. It should also be noted that this assumption is important not only when estimating the lifetime distribution, but also when estimating the censoring distribution. Furthermore, it should be noted that the highest powers are often achieved when the set-to-1 assumption is made for one of the two distributions being estimated but not the other.

Further research is currently underway in which we consider the effect of using kernel distribution function estimation for the lifetime and censoring distributions. The effect of this assumption, as opposed to the dichotomous approaches used at the sample maximum, will be considered in the goodness-of-fit context. Furthermore, we are testing the effect of the set-to-1 assumption in other settings, such as the estimation of the moments of the lifetime distribution. The effect of kernel distribution function estimations will be considered in this work as well.

References

- [1] M. Cuparić and B. Milošević. New characterization-based exponentiality tests for randomly censored data. *TEST*, 2021.
- [2] E. Bothma, J. S. Allison, M. Cockeran, and I. J. H. Visagie. Characteristic function and Laplace transform-based tests for exponentiality in the presence of random right censoring. *Stat*, 10(1):e394, 2021.
- [3] N. Kim. Goodness-of-fit tests for randomly censored Weibull distributions with estimated parameters. *Communications for Statistical Applications and Methods*, 24(5):519–531, 2017.
- [4] R. B. D’Agostino and M. A. Stephens. *Goodness-of-fit Techniques*, volume 68. CRC press, 1986.
- [5] D. R. Cox and D. Oakes. *Analysis of Survival Data*, volume 21. CRC Press, 1984.
- [6] S. Datta. Estimating the mean life time using right censored data. *Statistical Methodology*, 2(1):65–69, 2005.
- [7] E. L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481, 1958.
- [8] P. Martínez-Cambor. Testing the equality among distribution functions from independent and right censored samples via Cramér–von Mises criterion. *Journal of Applied Statistics*, 38(6):1117–1131, 2011.
- [9] E. T. Lee and J. Wang. *Statistical Methods for Survival Data Analysis*, volume 476. John Wiley & Sons, 2003.
- [10] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.
- [11] A. Jacobson, V. Wilson, and S. Pileggi. *parmsurvfit: Parametric Models for Survival Data*, 2018. R package version 0.1.0.
- [12] J. S. Allison, L. Santana, N. Smit, and I. J. H. Visagie. An "apples-to-apples" comparison of various tests for exponentiality. *Computational Statistics*, 32(4):1241–1283, 2017.
- [13] E. Bothma, J. S. Allison, and I. J. H. Visagie. New classes of tests for the Weibull distribution using Stein’s method in the presence of random right censoring. *Computational Statistics*, 2022.
- [14] N. Henze and I. J. H. Visagie. Testing for normality in any dimension based on a partial differential equation involving the moment generating function. *Annals of the Institute of Statistical Mathematics*, 72:1109–1136, 2020.
- [15] S. G. Meintanis. A review of testing procedures based on the empirical characteristic function. *South African Statistical Journal*, 50(1):1–14, 2016.
- [16] S. Csörgö. Estimating characteristic functions under random censorship. *Theory of Probability & Its Applications*, 28(3):615–623, 1984.
- [17] T. W. Epps and L. B. Pulley. A test of exponentiality vs. monotone-hazard alternatives derived from the empirical characteristic function. *Journal of the Royal Statistical Society: Series B (Methodological)*, 48(2):206–213, 1986.
- [18] N. Henze and S. G. Meintanis. Goodness-of-fit tests based on a new characterization of the exponential distribution. *Communications in Statistics-Theory and Methods*, 31(9):1479–1497, 2002.
- [19] L. Baringhaus and N. Henze. A class of consistent tests for exponentiality based on the empirical Laplace transform. *Annals of the Institute of Statistical Mathematics*, 43(3):551–564, 1991.
- [20] N. Henze and S. G. Meintanis. Tests of fit for exponentiality based on the empirical Laplace transform. *Statistics*, 36(2):147–161, 2002.
- [21] M. Krit. Goodness-of-fit tests for the Weibull distribution based on the Laplace transform. *Journal de la Société Française de Statistique*, 155(3):135–151, 2014.
- [22] N. Henze, S. G. Meintanis, and B. Ebner. Goodness-of-fit tests for the gamma distribution based on the empirical Laplace transform. *Communications in Statistics - Theory and Methods*, 41(9):1543–1556, 2012.
- [23] S. Betsch and B. Ebner. A new characterization of the gamma distribution and associated goodness-of-fit tests. *Metrika*, 82(7):779–806, 2019.
- [24] J. Mazucheli, L. B. Fernandes, and R. P. de Oliveira. *LindleyR: The Lindley distribution and its modifications*, 2016. R package version 1.1.0.
- [25] M. Hlavac. *stargazer: Well-formatted regression and summary statistics tables*, 2018.

Appendix A

Tables 9 to 14 show the empirical sizes and powers associated with samples of size $n = 100$. Tables 9 and 10 pertain to the hypothesis of exponentiality while Tables 11 and 12 contain results associated with testing for the Weibull class. The results for the gamma class are shown in Tables 13 and 14.

Table 9: Estimated powers for the exponential distribution for 10% censoring for a sample size of $n = 100$ with three different censoring distributions.

F	KS_n				CV_n				EP_n				$L_{n,0.25}$				$B_{n,0.25}$				$H_{n,1}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$exp(0.5)$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	2	5	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	2	5	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	6	5	5	5	5	5	4	5	2	5	5	5	5	5
$exp(1)$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	2	5	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	2	5	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	6	5	5	5	5	5	4	5	2	5	5	5	5	4
$exp(2)$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	3	5	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	6	6	5	5	5	5	5	3	5	5	5	5	5	5
	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	5	2	5	5	5	5	5
$\Gamma(0.6)$	85	85	85	85	90	90	90	90	89	89	89	89	98	98	98	98	97	97	97	97	85	85	86	86
	85	85	85	85	90	90	90	90	89	89	89	89	97	97	98	98	97	97	96	97	86	86	86	86
	85	85	85	85	90	90	90	90	89	89	89	89	98	98	97	97	97	97	96	97	85	85	86	85
$\Gamma(0.8)$	24	24	24	24	29	29	29	29	30	30	30	30	44	44	44	44	41	42	36	42	25	25	25	25
	24	25	24	24	29	29	29	29	30	30	31	31	44	44	44	44	42	42	36	42	26	26	26	26
	23	24	24	24	28	28	28	28	30	30	31	30	44	44	44	44	41	41	35	42	25	25	26	25
$\Gamma(1.2)$	16	16	16	17	19	19	20	20	19	19	21	21	23	23	23	23	22	24	12	25	18	18	18	17
	16	16	16	17	19	19	20	20	20	19	22	21	22	22	23	23	20	23	11	25	18	17	18	17
	17	17	16	17	20	20	19	19	20	20	21	20	23	23	23	23	20	24	10	24	18	18	18	17
$W(0.8)$	57	57	56	56	65	65	65	65	69	69	69	69	75	75	74	74	76	76	73	76	58	58	57	57
	56	56	55	55	64	64	64	64	69	69	69	69	75	75	74	74	76	75	73	75	57	57	57	57
	56	56	55	56	64	64	64	64	69	69	69	69	74	74	74	74	75	75	72	75	57	57	57	57
$W(1.2)$	37	38	36	39	46	46	47	48	51	50	54	52	46	46	47	47	48	52	25	52	41	41	42	40
	38	39	36	39	47	47	47	47	51	51	54	52	45	45	46	46	47	52	23	53	42	42	42	39
	37	38	35	39	46	46	46	47	50	50	54	51	45	45	46	46	45	51	22	53	41	41	42	38
$LN(1)$	41	41	40	40	50	49	49	49	14	14	14	14	60	58	58	58	58	57	49	56	31	31	29	29
	38	38	39	39	47	47	48	48	11	11	12	11	61	61	62	62	58	55	51	56	29	30	30	30
	38	38	38	38	47	47	47	47	11	11	12	11	61	61	62	62	55	55	47	56	30	31	31	30
$LN(1.5)$	98	98	98	98	99	99	99	99	99	99	99	99	93	93	93	93	98	98	98	98	97	97	97	97
	98	98	98	98	99	99	99	99	99	99	99	99	92	92	92	92	98	98	98	98	97	97	97	97
	98	98	98	98	99	99	99	99	99	99	99	99	91	91	91	91	98	98	98	98	97	97	96	97
$\chi^2(1)$	99	99	99	99	99	99	99	99	99	99	99	99	100	100	100	100	100	100	100	100	98	98	98	98
	98	98	98	98	99	99	99	99	99	99	99	99	100	100	100	100	100	100	100	100	98	98	98	98
	98	98	98	98	99	99	99	99	99	99	99	99	100	100	100	100	100	100	100	100	98	98	98	98
$\chi^2(3)$	62	62	60	63	71	71	71	71	71	71	74	73	81	81	81	81	80	82	56	83	67	67	67	64
	62	62	58	62	72	72	71	71	72	72	74	72	81	81	82	82	79	83	53	83	67	67	67	63
	61	61	57	62	71	71	71	72	72	72	75	73	81	81	81	81	76	82	48	83	66	65	66	63
$\beta(1,1)$	100	100	5	100	100	100	99	99	100	100	100	100	86	86	87	88	99	99	9	98	100	100	100	98
	100	100	5	100	100	100	99	99	100	100	100	100	87	87	87	88	99	99	9	98	100	100	100	98
	100	100	4	100	100	100	99	99	100	100	100	100	86	86	86	87	99	99	8	97	100	100	100	98
$\beta(0.5,1)$	68	68	51	70	83	83	81	83	9	9	11	10	91	91	91	91	88	88	55	88	82	82	81	75
	67	68	49	70	83	83	81	83	10	10	12	10	91	91	91	91	87	88	55	88	82	82	80	74
	67	67	50	70	83	83	81	83	9	9	12	10	91	91	91	91	87	88	55	88	82	82	81	74
$\beta(1,1.5)$	91	91	19	93	98	98	94	96	98	98	99	98	61	61	62	63	88	89	7	86	97	97	94	83
	90	91	18	93	98	98	94	95	98	98	99	98	61	61	62	62	87	88	7	86	97	97	94	82
	90	92	17	93	98	98	94	96	98	98	99	98	62	62	62	62	87	89	7	86	97	97	94	83

Table 10: Estimated powers for the exponential distribution for 30% censoring for a sample size of $n = 100$ with three different censoring distributions.

F	KS_n				CV_n				EP_n				$L_{n,0.25}$				$B_{n,0.25}$				$H_{n,1}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$exp(0.5)$	2	3	1	4	4	4	4	4	4	4	5	4	5	5	5	5	3	5	1	5	4	4	5	4
	2	3	2	4	4	3	4	3	4	3	3	3	3	5	5	5	5	4	4	3	5	4	3	4
	2	3	1	4	4	4	4	3	4	4	3	4	3	5	5	5	5	3	5	1	5	4	4	5
$exp(1)$	2	3	1	4	4	4	3	4	4	4	5	4	5	5	5	5	3	5	1	5	4	4	5	3
	2	3	2	4	3	4	3	3	3	3	3	3	3	5	5	5	5	3	4	2	5	4	3	4
	2	3	1	4	4	4	3	4	4	4	4	4	4	5	5	5	5	2	5	1	5	4	4	5
$exp(2)$	3	4	1	4	4	5	4	5	4	4	5	4	5	5	5	5	3	5	1	5	4	4	5	4
	2	3	2	4	3	4	3	4	3	3	3	3	3	5	5	5	5	4	5	3	5	4	4	4
	2	3	1	4	4	4	3	4	4	4	5	4	5	5	5	5	3	5	1	5	4	4	5	4
$\Gamma(0.6)$	63	66	61	67	81	81	80	81	74	76	75	76	96	96	96	96	90	94	87	94	71	72	72	72
	38	65	36	66	78	79	78	79	66	81	65	80	95	95	95	95	82	93	81	93	69	72	70	72
	60	66	59	67	80	80	80	80	74	76	75	76	96	96	96	96	88	94	85	94	70	72	72	72
$\Gamma(0.8)$	11	14	10	15	21	21	20	21	22	22	23	22	40	40	40	40	26	36	17	37	18	18	19	18
	7	12	6	13	20	19	18	19	14	24	15	24	39	39	39	39	23	32	20	32	16	19	18	19
	9	13	7	14	20	20	19	21	20	22	21	22	39	39	39	39	23	34	16	36	17	18	19	17
$\Gamma(1.2)$	5	9	2	10	13	14	12	14	12	11	15	11	20	20	20	20	21	6	20	1	21	11	9	12
	1	7	1	9	9	12	7	12	13	7	12	7	20	20	20	20	20	1	21	0	21	9	6	11
	3	7	1	8	12	13	10	13	12	10	13	9	20	20	20	20	20	3	20	0	21	10	8	11
$W(0.8)$	30	34	27	34	47	48	47	47	49	50	49	50	65	65	65	65	55	64	46	64	40	41	41	41
	15	31	14	31	43	43	42	42	35	51	35	50	63	63	63	63	46	58	44	58	36	41	36	40
	25	32	22	33	46	46	45	46	46	49	46	49	65	64	64	64	50	63	43	63	38	40	41	40
$W(1.2)$	12	20	3	22	30	32	26	31	30	27	33	25	38	38	39	39	14	42	1	43	25	22	27	19
	1	15	0	20	21	28	16	26	28	17	26	15	36	37	37	37	1	41	0	42	22	14	25	13
	7	18	1	21	28	30	23	29	29	25	30	22	38	38	38	39	8	41	1	43	23	19	26	16
$LN(1)$	19	25	14	24	38	39	35	37	10	10	11	9	71	71	70	71	34	59	18	59	22	21	22	20
	4	18	3	20	26	30	24	30	14	7	14	7	77	77	78	78	10	62	8	63	19	16	21	16
	9	18	6	21	33	34	31	35	11	9	12	9	74	74	75	75	23	58	12	61	19	19	22	19
$LN(1.5)$	73	75	73	76	85	85	85	85	86	88	87	88	61	61	61	61	80	83	79	83	75	79	76	79
	48	65	47	66	77	76	78	76	62	80	63	80	45	45	46	46	70	65	69	65	65	74	67	75
	61	69	60	68	81	80	80	80	77	81	77	81	52	52	52	52	72	76	69	77	69	75	70	75
$\chi^2(1)$	92	93	91	93	97	97	97	97	93	94	93	94	100	100	100	100	99	100	99	100	93	94	94	94
	77	93	75	93	97	97	97	97	88	96	88	96	100	100	100	100	98	100	98	100	93	94	93	94
	89	92	89	92	97	97	97	97	92	94	93	94	100	100	100	100	99	100	99	100	93	93	93	94
$\chi^2(3)$	25	38	7	41	54	56	47	53	48	44	51	41	74	74	75	75	39	74	7	74	45	41	47	36
	2	31	0	37	41	51	33	47	47	31	44	27	74	74	74	74	4	74	1	74	41	29	45	28
	7	28	1	35	47	52	37	48	44	37	43	31	73	74	74	74	15	73	2	73	40	33	45	29
$\beta(1,1)$	100	100	0	100	100	100	59	74	90	89	40	25	74	74	75	76	95	96	0	80	100	100	99	85
	99	100	0	100	99	100	46	64	79	78	17	10	71	72	71	73	89	95	0	75	100	100	99	82
	99	100	0	100	100	100	55	71	87	86	33	19	73	73	73	75	94	96	0	78	100	100	99	84
$\beta(0.5,1)$	51	57	17	62	76	76	67	73	14	14	17	13	93	93	93	93	81	88	43	88	67	66	70	58
	17	49	5	58	66	70	56	68	28	32	28	28	94	94	94	94	59	89	39	89	59	56	66	54
	48	55	13	60	75	75	66	73	16	17	19	15	93	93	93	93	80	88	41	89	65	64	69	57
$\beta(1,1.5)$	65	79	0	87	87	87	48	64	74	72	61	41	45	45	46	47	53	73	0	62	85	82	79	51
	20	72	0	86	76	79	32	51	55	50	30	19	42	42	42	44	11	66	0	56	77	69	73	44
	58	78	0	87	86	86	45	61	70	68	54	35	46	46	45	47	47	73	0	62	84	81	79	50

Table 11: Estimated powers for the Weibull distribution for 10% censoring for a sample size of $n = 100$ with three different censoring distributions.

F	KS_n				CV_n				KR_n				$S_{n,1}^{(1)}$				$S_{n,2}^{(1)}$				$S_{n,2}^{(2)}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$W(0.5)$	4	4	3	4	4	4	4	4	2	2	0	1	2	2	2	2	2	2	1	2	2	3	1	4
	3	4	3	5	4	5	4	4	2	1	1	1	3	3	3	3	3	2	2	1	3	3	2	3
	3	4	2	5	4	4	5	5	2	1	1	1	3	3	3	3	3	1	1	1	3	3	1	3
$W(1)$	5	5	4	5	5	5	4	4	3	3	0	1	4	4	4	4	4	4	0	4	4	5	0	5
	5	5	5	5	4	4	5	5	2	3	0	1	3	3	3	3	3	4	0	3	3	4	0	5
	5	5	5	5	5	5	5	5	2	3	0	2	3	3	3	3	2	4	0	3	3	4	0	5
$W(2)$	5	5	5	5	5	5	4	4	4	3	0	1	4	4	4	4	4	5	0	4	4	5	0	5
	5	5	5	5	5	5	5	5	4	3	0	2	4	4	4	4	4	5	0	4	5	5	0	5
	5	5	5	5	5	5	5	5	5	3	0	1	4	4	4	4	4	4	0	4	4	4	0	5
$\Gamma(3)$	15	15	15	15	20	20	18	18	18	23	0	21	26	26	26	26	28	29	3	29	23	24	9	24
	15	15	15	15	19	19	18	18	14	21	0	20	24	24	24	24	26	28	3	29	21	23	9	24
	14	14	15	15	18	18	18	18	6	19	0	19	22	22	22	22	21	27	3	28	20	23	8	24
$\Gamma(5)$	23	23	23	23	30	30	30	30	31	34	0	33	39	39	39	39	43	45	11	45	36	37	21	39
	22	22	24	24	30	30	31	31	29	36	0	35	40	40	40	40	43	45	10	47	36	37	21	39
	22	22	23	23	29	29	30	30	15	34	0	33	40	40	40	40	42	46	9	44	35	38	20	38
$\Gamma(10)$	35	35	33	33	44	44	45	45	48	51	0	50	58	58	58	58	62	64	30	62	53	53	40	53
	31	31	32	32	42	42	41	41	46	49	0	49	56	56	56	56	60	62	29	62	51	52	38	52
	32	32	33	33	44	44	44	44	41	50	0	50	56	56	56	56	60	63	28	64	51	52	38	54
$LN(0.5)$	60	60	60	60	76	76	75	75	70	79	1	78	88	88	88	88	90	92	75	92	83	84	77	84
	60	60	60	60	76	76	77	77	68	79	1	79	87	87	87	87	90	91	75	91	83	83	78	84
	60	60	59	59	75	75	75	75	49	77	1	76	86	86	86	86	88	91	70	90	82	83	75	83
$LN(1)$	56	56	56	56	73	73	73	73	9	72	4	73	82	82	82	82	79	89	56	90	78	82	69	83
	56	56	56	56	72	72	73	73	13	70	8	70	78	78	78	78	70	88	52	88	73	81	65	82
	55	55	54	54	72	72	71	71	11	66	6	67	77	77	77	77	64	87	42	87	71	80	59	80
$LN(2)$	49	50	50	51	67	67	67	67	14	49	10	49	57	57	57	57	38	78	29	78	50	74	40	74
	49	50	48	50	66	66	68	67	19	22	16	22	42	42	42	42	38	69	35	69	40	66	35	67
	47	50	44	49	67	66	65	65	15	25	13	26	38	38	38	38	32	63	29	63	34	65	30	66
$\chi^2(6)$	15	15	15	15	20	20	18	18	18	23	0	21	26	26	26	26	28	29	3	29	23	24	9	24
	15	15	14	14	18	18	18	18	15	22	0	20	23	23	23	23	25	27	3	28	21	23	8	24
	15	15	15	15	19	19	18	18	6	18	0	18	22	22	22	22	21	27	3	28	19	23	7	24
$\chi^2(8)$	19	19	19	19	24	24	25	25	25	29	0	27	32	32	32	32	35	37	6	38	29	30	16	31
	19	19	18	18	24	24	24	24	21	28	0	27	32	32	32	32	35	38	6	37	29	30	14	31
	19	19	21	21	25	25	26	26	7	27	0	25	32	32	32	32	32	37	6	37	29	31	13	32
$\chi^2(10)$	21	21	22	22	28	28	30	30	31	35	0	35	40	40	40	40	43	45	11	45	36	37	21	38
	22	22	21	21	30	30	29	29	28	34	0	33	38	38	38	38	42	44	12	45	35	36	21	38
	22	22	22	22	29	29	29	29	14	33	0	33	38	38	38	38	39	45	10	45	34	36	20	37
$\beta(1,1)$	83	83	12	86	94	94	86	88	93	93	0	0	96	96	96	96	96	97	0	13	96	96	0	50
	83	83	12	85	94	94	85	87	92	93	0	0	96	96	96	96	96	96	0	12	96	96	0	49
	83	84	11	85	94	94	85	87	91	91	0	0	96	96	96	96	96	96	0	11	96	96	0	47
$\beta(0.5,1)$	82	83	4	85	93	93	78	83	87	88	0	0	95	95	95	95	96	96	0	6	96	96	0	39
	81	82	4	85	93	93	78	82	85	86	0	0	95	95	95	95	95	95	0	6	96	96	0	37
	81	82	3	86	93	93	76	80	82	82	0	0	94	94	94	94	94	95	0	5	95	96	0	35
$\beta(1,2)$	27	27	20	29	35	35	36	38	24	32	0	0	41	41	41	41	42	49	0	16	39	45	0	30
	29	29	19	31	39	39	36	38	23	31	0	0	41	41	41	41	44	50	0	16	39	46	0	30
	29	30	18	31	38	38	36	37	24	31	0	0	43	43	43	43	44	50	0	15	41	46	0	30

Table 12: Estimated powers for the Weibull distribution for 30% censoring for a sample size of $n = 100$ with three different censoring distributions.

F	KS_n				CV_n				KR_n				$S_{n,1}^{(1)}$				$S_{n,2}^{(1)}$				$S_{n,2}^{(2)}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$W(0.5)$	2	4	1	5	2	1	1	1	2	1	1	1	3	3	3	3	3	1	2	1	3	1	2	1
	2	5	2	6	2	1	2	1	2	1	2	1	4	4	4	4	4	1	4	1	4	1	4	1
	2	3	1	5	2	1	1	1	2	1	1	1	3	3	3	3	3	1	2	1	3	1	2	1
$W(1)$	2	3	1	4	3	4	3	4	2	1	0	1	2	2	2	2	2	2	0	2	2	3	1	3
	2	3	2	4	3	2	2	2	3	1	2	1	3	3	3	3	3	1	3	1	3	1	3	1
	1	3	1	3	3	3	2	3	2	1	1	1	2	2	2	2	2	2	1	2	2	2	1	2
$W(2)$	4	4	4	4	4	4	4	4	2	1	0	0	3	3	3	3	3	3	0	3	3	4	0	5
	4	4	3	4	5	5	4	5	1	1	0	0	2	2	2	2	2	3	0	2	2	3	0	3
	4	4	3	4	4	4	4	5	1	1	0	0	2	2	2	2	2	3	0	3	2	4	0	5
$\Gamma(3)$	11	12	11	12	15	15	14	15	4	5	1	3	15	15	15	15	11	21	2	20	12	19	5	20
	10	10	8	10	13	13	12	13	5	4	2	3	9	9	9	9	9	12	4	13	9	14	4	15
	9	10	8	9	13	13	12	13	4	4	1	4	10	10	10	10	7	15	3	16	9	16	4	17
$\Gamma(5)$	17	17	17	18	23	23	23	23	6	6	0	5	27	27	27	27	24	34	4	35	24	29	10	31
	17	17	16	17	22	22	23	23	6	6	1	5	23	23	23	23	18	30	5	31	20	27	8	29
	15	16	15	16	22	22	22	22	5	7	1	6	24	24	24	24	17	30	4	32	20	28	8	31
$\Gamma(10)$	27	27	25	25	36	36	35	35	8	12	0	9	45	45	45	45	46	51	13	51	41	44	25	44
	26	26	25	25	34	34	33	33	8	11	1	8	41	41	41	41	38	49	11	50	37	41	22	43
	26	26	26	26	33	33	35	35	6	9	1	7	39	39	39	39	34	48	10	50	34	41	20	45
$LN(0.5)$	44	44	47	47	61	61	63	63	10	21	3	19	70	70	70	70	62	79	38	80	65	72	52	73
	43	44	42	43	59	59	59	59	12	17	7	16	57	57	57	57	39	71	26	72	50	67	39	69
	44	44	42	43	59	59	58	58	8	21	3	18	65	65	65	65	53	75	30	75	58	69	44	69
$LN(1)$	28	31	28	32	50	50	50	51	9	20	7	20	33	33	33	33	23	52	18	53	29	56	24	58
	17	27	13	28	42	41	41	42	15	1	12	1	22	22	22	22	25	5	22	5	22	18	19	19
	21	28	19	28	45	46	44	46	10	18	8	18	25	25	25	25	21	37	17	40	22	46	18	48
$LN(2)$	11	16	10	18	28	28	27	27	12	4	11	4	17	17	17	17	19	7	17	7	16	12	15	13
	10	9	9	12	15	5	14	4	10	1	9	1	22	22	22	22	25	1	25	0	20	0	20	0
	10	15	9	17	17	15	17	16	10	1	10	1	17	17	17	17	20	2	19	2	16	4	16	4
$\chi^2(6)$	11	12	11	12	15	15	14	15	4	5	1	3	15	15	15	15	11	21	2	20	12	19	5	20
	10	10	8	10	13	13	12	13	5	4	2	3	9	9	9	9	9	12	4	13	9	14	4	15
	9	9	7	9	13	13	12	13	3	5	1	4	9	9	9	9	7	16	3	16	8	16	4	17
$\chi^2(8)$	15	15	15	15	19	19	19	19	5	5	0	3	22	22	22	22	19	28	3	28	19	24	7	26
	13	13	12	14	18	18	17	18	6	5	1	4	16	16	16	16	12	21	4	22	14	21	5	23
	12	13	11	13	18	18	16	17	4	6	1	5	16	16	16	16	11	22	4	24	14	22	6	24
$\chi^2(10)$	17	17	17	17	23	23	23	23	6	6	0	5	27	27	27	27	24	34	4	35	24	29	10	31
	17	18	16	16	23	23	21	21	5	7	1	5	22	22	22	22	17	30	5	30	20	28	8	28
	16	16	15	16	22	22	21	21	5	7	1	6	22	22	22	22	16	29	5	31	19	27	9	30
$\beta(1,1)$	64	67	0	75	82	82	46	54	25	26	0	0	83	83	83	83	81	83	0	2	84	87	0	17
	60	67	0	76	77	77	32	44	12	22	0	0	66	66	66	66	54	71	0	1	66	78	0	13
	64	67	0	74	81	81	43	54	24	26	0	0	80	80	80	80	78	82	0	2	81	85	0	19
$\beta(0.5,1)$	32	60	0	80	64	65	7	18	0	17	0	0	14	14	14	14	1	38	0	1	11	56	0	8
	1	44	0	77	10	22	1	6	0	2	0	0	0	0	0	0	4	0	1	0	12	0	5	
	20	59	0	81	60	63	7	16	0	15	0	0	7	7	7	7	0	30	0	1	5	51	0	8
$\beta(1,2)$	12	16	2	19	21	21	14	19	0	4	0	0	4	4	4	4	0	18	0	3	3	21	0	11
	3	14	0	18	15	18	8	14	0	2	0	0	1	1	1	1	0	7	0	2	0	14	0	8
	12	16	1	19	22	23	14	19	0	4	0	0	4	4	4	4	0	17	0	4	3	21	0	12

Table 13: Estimated powers for the gamma distribution for 10% censoring for a sample size of $n = 100$ with three different censoring distributions.

F	KS_n				CV_n				$M_{n,1}^{(1)}$				$M_{n,4}^{(2)}$				$R_{n,0.5}$				$R_{n,1}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$\Gamma(0.5)$	5	5	4	5	5	5	5	5	3	5	0	5	2	3	0	3	2	1	4	0	2	1	4	0
	5	5	5	5	5	5	5	5	4	4	1	5	3	2	0	2	2	2	5	1	2	2	5	1
	4	5	5	5	4	4	5	5	2	4	0	5	2	2	0	2	3	1	5	1	3	1	5	1
$\Gamma(1)$	5	5	5	6	6	6	5	5	4	6	0	5	4	5	0	4	2	1	3	0	2	1	3	0
	5	5	5	5	5	5	5	5	3	5	0	4	3	4	0	3	2	1	4	0	2	1	4	0
	5	5	5	5	5	5	5	5	3	5	0	5	3	4	0	3	2	1	5	0	2	1	5	0
$\Gamma(2)$	4	4	4	4	5	5	5	5	4	5	0	4	4	4	0	3	1	1	2	0	1	1	2	0
	5	5	5	5	5	5	5	5	4	5	0	4	5	4	0	3	1	1	3	0	1	1	3	0
	5	5	5	5	5	5	5	5	3	4	0	4	4	4	0	3	2	1	4	0	2	1	4	0
$W(2)$	19	19	17	19	23	23	23	24	27	34	0	22	18	28	0	13	1	1	8	0	1	1	8	0
	18	18	18	20	22	22	23	23	25	33	0	22	14	27	0	12	1	1	7	0	1	1	7	0
	18	18	17	19	22	22	23	23	21	33	0	20	12	27	0	11	1	1	6	0	1	1	6	0
$W(3)$	31	31	28	32	39	39	39	40	52	56	0	16	34	35	0	4	2	1	14	0	2	1	14	0
	32	32	27	34	40	40	40	41	51	55	0	15	33	34	0	3	1	1	14	0	1	1	14	0
	33	33	28	32	41	41	39	40	49	55	0	16	33	34	0	4	1	1	13	0	1	1	13	0
$LN(1)$	58	58	59	59	70	70	71	71	75	83	53	83	66	79	35	79	3	3	3	3	3	3	3	3
	58	58	59	59	69	68	71	71	70	82	52	84	60	83	39	84	4	5	4	5	4	5	4	5
	57	57	56	57	70	70	69	69	62	83	41	82	47	81	30	81	5	6	5	5	5	6	5	5
$LN(1.5)$	85	85	85	85	92	92	93	93	93	97	88	97	74	93	61	93	4	6	4	6	4	6	4	6
	83	83	84	84	91	91	92	92	86	96	82	97	62	94	58	95	5	12	5	12	5	12	5	12
	82	82	82	83	91	91	91	91	80	96	70	96	54	93	46	93	7	12	8	12	7	12	8	12
$LN(2.5)$	98	98	98	98	100	100	100	100	94	99	92	99	61	92	57	92	6	13	7	13	6	13	7	13
	98	98	98	98	99	99	99	99	88	99	86	99	66	88	64	87	5	23	5	22	5	23	5	22
	97	97	97	97	99	99	99	99	79	99	77	99	56	79	54	80	7	20	8	20	7	20	8	20
$\chi^2(0.5)$	4	5	4	4	5	5	4	5	2	4	1	4	2	1	1	3	1	6	1	3	1	6	1	1
	4	5	4	5	5	5	5	5	3	2	2	3	3	0	2	0	4	2	6	1	4	2	6	1
	4	5	3	5	4	5	4	5	2	3	1	3	3	1	1	0	3	2	8	1	3	2	8	1
$\chi^2(0.75)$	4	4	4	5	4	4	4	5	2	4	0	4	2	2	0	1	2	2	4	1	2	2	4	1
	5	5	4	5	4	4	4	5	4	4	1	4	3	1	1	0	3	2	6	1	3	2	6	1
	4	5	4	5	5	5	4	4	3	4	1	4	2	1	1	3	2	6	1	3	2	6	1	
$\beta(1,1)$	96	96	3	96	99	99	88	91	98	98	0	11	98	98	0	1	3	3	100	0	3	3	100	0
	95	95	3	97	99	99	88	92	98	98	0	10	98	98	0	1	2	2	100	0	2	2	100	0
	96	96	3	97	99	99	89	91	98	98	0	12	97	97	0	1	2	1	100	0	2	1	100	0
$\beta(0.5,1)$	70	71	8	75	85	85	73	77	85	87	0	37	88	89	0	0	1	1	99	0	1	1	99	0
	71	72	6	74	86	86	70	75	85	86	0	34	87	88	0	0	1	1	98	0	1	1	98	0
	70	71	7	76	86	86	71	75	85	87	0	35	87	88	0	0	1	1	98	0	1	1	98	0
$\beta(1,2)$	41	42	21	46	54	54	52	54	55	62	0	29	58	64	0	11	1	0	52	0	1	0	52	0
	43	44	20	47	56	56	51	54	53	62	0	29	54	64	0	10	1	0	52	0	1	0	52	0
	43	43	22	43	56	56	50	53	54	62	0	29	56	64	0	10	1	0	48	0	1	0	48	0
$\mathcal{L}(0.25)$	9	9	9	9	9	9	10	10	9	15	0	15	6	13	0	11	1	1	4	0	1	1	4	0
	9	9	8	9	9	9	10	10	7	16	0	15	4	14	0	11	1	1	4	0	1	1	4	0
	9	9	8	9	10	10	10	10	2	15	0	14	1	13	0	9	1	1	4	0	1	1	4	0
$\mathcal{L}(0.5)$	9	9	9	9	10	10	10	11	8	14	0	13	4	12	0	9	1	1	3	0	1	1	3	0
	9	9	9	10	10	10	11	11	6	15	0	13	3	13	0	9	1	1	3	0	1	1	3	0
	9	9	8	9	10	10	10	10	2	14	0	13	1	11	0	8	1	1	4	0	1	1	4	0

Table 14: Estimated powers for the gamma distribution for 30% censoring for a sample size of $n = 100$ with three different censoring distributions.

F	KS_n				CV_n				$M_{n,1}^{(1)}$				$M_{n,4}^{(2)}$				$R_{n,0.5}$				$R_{n,1}$			
	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT	FF	FT	TF	TT
$\Gamma(0.5)$	2	3	1	4	2	3	2	3	3	1	1	1	3	1	1	1	4	2	8	2	4	2	8	2
	2	5	2	6	2	1	2	1	3	1	3	1	3	1	3	1	5	3	7	3	5	3	7	3
	2	3	1	5	2	2	2	2	3	1	2	1	3	1	2	1	4	2	9	2	4	2	9	2
$\Gamma(1)$	2	3	1	3	3	4	3	4	2	2	1	2	2	1	1	1	3	2	7	1	3	2	7	1
	3	3	2	4	3	3	2	2	4	1	2	1	4	1	2	0	4	3	8	2	4	3	8	2
	2	3	1	4	4	4	2	3	2	2	1	2	2	1	1	1	3	2	7	2	3	2	7	2
$\Gamma(2)$	4	4	3	4	5	5	4	5	2	3	0	2	2	3	0	1	2	1	4	1	2	1	4	1
	3	4	1	4	4	4	3	3	3	1	1	1	3	1	1	0	3	2	6	1	3	2	6	1
	3	4	2	4	4	4	4	5	2	2	0	2	2	2	0	1	3	2	5	1	3	2	5	1
$W(2)$	12	13	7	14	16	16	15	16	1	18	0	5	0	11	0	2	1	0	5	0	1	0	5	0
	9	12	3	12	16	16	13	15	0	11	0	3	0	5	0	1	1	0	6	0	1	0	6	0
	11	12	5	13	15	15	14	16	0	15	0	5	0	8	0	2	1	0	4	0	1	0	4	0
$W(3)$	22	23	14	24	31	31	29	30	11	34	0	3	3	11	0	3	1	0	5	0	1	0	5	0
	21	23	11	25	30	31	27	30	0	24	0	2	2	8	1	3	1	0	5	0	1	0	5	0
	22	23	13	23	30	30	27	29	1	29	0	2	3	10	0	3	1	0	5	0	1	0	5	0
$LN(1)$	28	31	28	32	46	46	46	46	22	54	15	56	19	46	13	47	7	8	9	7	7	8	9	7
	20	26	17	28	40	38	38	37	28	14	25	15	27	4	23	4	6	16	8	15	6	16	8	15
	22	28	22	28	43	43	41	42	18	41	15	44	17	30	13	32	8	9	9	9	8	9	9	9
$LN(1.5)$	32	44	32	45	65	66	64	65	30	62	25	64	24	33	21	35	9	14	10	14	9	14	10	14
	15	31	16	34	43	41	46	43	33	2	33	3	30	1	31	1	5	18	6	19	5	18	6	19
	19	34	18	34	54	55	53	57	27	37	25	39	24	16	22	17	9	17	9	16	9	17	9	16
$LN(2.5)$	20	38	18	46	70	71	67	72	37	25	34	25	32	11	29	11	8	21	9	21	8	21	9	21
	14	35	14	36	31	19	32	20	43	1	42	1	36	1	36	1	3	20	4	21	3	20	4	21
	17	30	17	31	46	53	46	51	37	7	36	7	32	3	32	3	7	23	8	23	7	23	8	23
$\chi^2(0.5)$	2	4	1	6	2	1	1	1	3	1	2	1	3	1	2	1	5	3	10	3	5	3	10	3
	2	4	2	6	2	1	2	1	3	1	3	1	2	1	2	1	4	3	6	2	4	3	6	2
	2	3	1	5	2	1	1	1	3	1	2	1	3	1	2	1	5	3	10	3	5	3	10	3
$\chi^2(0.75)$	2	3	1	4	2	2	1	1	2	1	2	1	2	1	2	1	5	3	8	2	5	3	8	2
	2	5	2	6	2	1	2	1	4	1	3	1	3	1	3	1	4	3	6	3	4	3	6	3
	2	3	1	5	2	1	1	1	2	1	2	1	3	1	2	1	4	2	10	2	4	2	10	2
$\beta(1,1)$	86	88	0	94	93	93	40	53	87	88	0	0	87	88	0	0	0	0	87	0	0	0	87	0
	79	88	0	95	90	91	26	41	56	73	0	0	38	60	0	0	0	0	77	0	0	0	77	0
	85	88	0	94	93	93	35	48	84	85	0	0	82	84	0	0	0	0	83	0	0	0	83	0
$\beta(0.5,1)$	20	43	0	65	49	51	10	20	0	31	0	1	0	6	0	0	0	0	34	0	0	0	34	0
	1	29	0	63	12	21	2	8	0	3	0	1	0	1	0	0	1	0	28	0	1	0	28	0
	17	44	0	65	47	49	8	19	0	25	0	1	0	5	0	0	0	0	30	0	0	0	30	0
$\beta(1,2)$	18	25	1	29	34	35	20	27	0	28	0	5	0	13	0	0	1	0	13	0	1	0	13	0
	3	19	0	27	22	26	9	17	0	8	0	2	0	1	0	0	1	0	14	0	1	0	14	0
	14	22	0	29	30	31	19	25	0	25	0	4	0	7	0	0	0	0	14	0	0	0	14	0
$\mathcal{L}(0.25)$	4	5	2	6	7	8	7	8	1	9	0	6	1	3	0	1	2	1	6	0	2	1	6	0
	2	5	1	6	6	7	4	7	2	5	1	3	2	1	1	1	3	1	6	1	3	1	6	1
	2	4	1	5	6	7	4	6	1	5	0	4	1	2	0	1	2	1	<i>b</i>	1	2	1	<i>b</i>	1
$\mathcal{L}(0.5)$	4	6	1	6	7	8	6	7	1	8	0	5	1	2	0	1	2	1	5	0	2	1	5	0
	1	5	0	5	4	6	3	5	1	3	1	2	1	1	1	0	3	1	7	1	3	1	7	1
	1	4	1	5	6	6	4	6	1	4	0	3	1	1	0	1	2	1	6	1	2	1	6	1

Table 15: Averages of estimated powers for the exponential distribution for 10% and 30% censoring for a sample size of $n = 50$ with three different censoring distributions.

F	FF	FT	TF	TT	FF	FT	TF	TT
$exp(0.5)$	5	5	4	5	4	3	4	3
	5	5	5	5	3	3	3	3
	5	5	4	5	4	3	4	3
$exp(1)$	5	5	5	5	4	3	4	3
	5	5	4	5	4	3	3	3
	5	5	4	5	4	3	4	3
$exp(2)$	5	5	5	5	4	3	4	3
	5	5	5	5	4	3	3	3
	5	5	4	5	5	3	4	3
$\Gamma(0.6)$	66	66	66	66	47	48	52	47
	66	66	66	66	46	44	51	43
	66	66	66	66	47	48	52	46
$\Gamma(0.8)$	19	19	18	19	15	13	14	12
	19	19	18	19	14	12	13	11
	19	19	19	20	15	13	14	12
$\Gamma(1.2)$	11	11	10	12	9	6	7	5
	11	11	10	11	8	4	6	4
	10	11	10	12	9	5	7	4
$W(0.8)$	41	41	40	41	28	26	28	24
	40	40	39	40	25	22	25	21
	40	40	39	40	27	24	27	22
$W(1.2)$	23	24	21	24	16	11	14	9
	23	23	21	25	15	8	12	7
	22	24	20	24	15	10	13	8
$LN(1)$	20	21	19	21	19	13	16	12
	20	20	18	20	19	11	16	11
	20	20	18	20	20	12	16	12
$LN(1.5)$	82	82	82	82	45	46	48	45
	81	81	81	81	35	33	36	33
	80	80	79	79	39	38	42	37
$\chi^2(1)$	88	88	88	88	70	73	76	72
	88	88	88	88	71	67	76	68
	88	88	88	88	70	71	76	70
$\chi^2(3)$	40	40	36	41	28	20	26	17
	40	40	35	41	26	16	23	14
	38	40	34	40	26	16	22	14
$\beta(1, 1)$	86	86	56	80	42	65	68	26
	86	86	55	80	39	54	62	23
	85	86	55	79	42	64	68	25
$\beta(0.5, 1)$	44	44	38	44	41	36	39	29
	44	45	38	44	42	32	39	28
	44	44	38	44	41	35	39	29
$\beta(1, 1.5)$	62	62	44	58	30	31	36	18
	60	61	43	58	27	21	29	13
	61	61	43	57	30	29	35	16
Average sizes	5	5	4	5	3	4	3	4
Average powers	48	48	43	48	30	34	25	32

Table 16: Averages of estimated powers for the exponential distribution for 10% and 30% censoring for a sample size of $n = 100$ with three different censoring distributions.

F	FF	FT	TF	TT	FF	FT	TF	TT
$exp(0.5)$	5	5	4	5	4	4	4	4
	5	5	4	5	4	4	3	4
	5	5	5	5	4	4	3	4
$exp(1)$	5	5	4	5	4	4	3	4
	5	5	4	5	3	4	3	4
	5	5	5	5	4	4	3	4
$exp(2)$	5	5	5	5	4	4	4	4
	5	5	5	5	4	4	3	4
	5	5	4	5	4	4	3	4
$\Gamma(0.6)$	91	91	91	91	79	81	78	81
	91	91	91	91	71	81	71	81
	91	91	90	90	78	81	78	81
$\Gamma(0.8)$	32	32	31	32	23	25	22	26
	32	33	32	33	20	24	19	24
	32	32	31	32	21	24	20	25
$\Gamma(1.2)$	20	20	18	20	11	14	10	14
	19	19	18	20	9	12	8	12
	20	20	18	20	10	13	9	13
$W(0.8)$	67	67	66	66	48	50	46	50
	66	66	65	66	40	48	39	47
	66	66	65	66	45	49	44	49
$W(1.2)$	45	46	42	46	25	30	22	30
	45	46	41	46	18	25	17	26
	44	45	41	46	22	28	20	28
$LN(1)$	42	42	40	41	32	38	28	37
	41	40	40	41	25	35	25	36
	40	40	40	41	28	35	26	37
$LN(1.5)$	97	97	97	97	77	78	77	79
	97	97	97	97	61	68	62	68
	97	97	97	97	69	72	68	72
$\chi^2(1)$	99	99	99	99	96	96	96	96
	99	99	99	99	92	97	92	97
	99	99	99	99	95	96	95	96
$\chi^2(3)$	72	72	68	72	48	54	39	53
	72	73	68	72	35	48	33	48
	71	72	66	72	38	50	34	48
$\beta(1, 1)$	98	98	67	97	93	93	46	73
	98	98	67	97	90	91	39	67
	98	98	66	97	92	92	43	71
$\beta(0.5, 1)$	70	70	62	70	64	66	51	64
	70	70	61	69	54	65	48	65
	70	70	62	69	63	65	50	64
$\beta(1, 1.5)$	89	89	62	86	68	73	39	59
	88	89	62	86	47	63	30	50
	89	89	62	86	65	72	37	57
Average sizes	5	5	4	5	4	4	3	4
Average powers	68	68	62	68	51	56	43	53

Table 17: Averages of estimated powers for the Weibull distribution for 10% and 30% censoring for a sample size of $n = 50$ with three different censoring distributions.

F	FF	FT	TF	TT	FF	FT	TF	TT
$W(0.5)$	3	3	2	3	3	2	2	2
	3	3	2	3	4	2	4	2
	3	2	2	2	3	2	3	2
$W(1)$	4	4	2	4	3	2	1	2
	3	3	2	3	3	2	2	2
	3	4	2	3	2	2	1	2
$W(2)$	4	4	2	4	3	3	1	3
	4	4	2	4	3	3	2	2
	4	4	2	4	3	3	1	2
$\Gamma(3)$	11	12	6	11	6	7	4	7
	11	12	6	10	6	6	4	6
	10	11	6	10	5	6	4	6
$\Gamma(5)$	17	18	9	17	10	11	6	11
	17	18	10	17	9	10	6	9
	15	17	9	16	9	10	6	10
$\Gamma(10)$	26	26	14	25	16	16	10	16
	26	26	15	25	15	16	10	16
	23	24	14	23	13	14	9	15
$LN(0.5)$	45	47	31	47	25	29	20	29
	44	47	30	46	21	23	17	24
	39	44	29	45	21	25	17	26
$LN(1)$	36	43	28	42	16	17	13	17
	33	41	27	40	14	8	13	8
	30	36	25	37	12	13	11	13
$LN(2)$	25	30	22	30	12	6	10	6
	24	24	23	25	13	5	12	5
	22	22	20	22	11	4	11	5
$\chi^2(6)$	11	12	6	11	6	7	4	7
	11	12	6	10	6	6	4	6
	10	11	6	10	5	6	4	6
$\chi^2(8)$	14	15	8	14	8	9	6	9
	14	15	8	14	8	8	5	8
	13	14	8	14	7	8	5	8
$\chi^2(10)$	17	18	9	17	10	11	6	11
	17	18	10	17	9	10	6	9
	15	17	9	16	8	10	6	10
$\beta(1, 1)$	60	61	21	33	26	32	8	17
	60	61	20	33	15	24	4	13
	59	60	20	33	22	29	7	16
$\beta(0.5, 1)$	52	55	18	30	6	13	1	8
	49	53	16	29	1	6	0	8
	50	55	17	29	4	10	1	9
$\beta(1, 2)$	16	18	8	12	3	6	1	4
	14	17	7	12	2	4	0	3
	14	16	7	12	2	5	1	4
Average sizes	3	3	2	3	3	2	2	2
Average powers	26	28	15	23	11	12	7	11

Table 18: Averages of estimated powers for the Weibull distribution for 10% and 30% censoring for a sample size of $n = 100$ with three different censoring distributions.

F	FF	FT	TF	TT	FF	FT	TF	TT
$W(0.5)$	3	3	2	3	2	2	2	2
	3	3	2	3	3	2	3	2
	3	3	2	3	2	2	2	2
$W(1)$	4	4	2	4	2	2	1	3
	3	4	2	4	3	2	2	2
	3	4	2	4	2	2	1	2
$W(2)$	4	4	2	4	3	3	2	3
	4	4	2	4	3	3	2	3
	4	4	2	4	2	10	2	3
$\Gamma(3)$	22	23	12	22	11	14	8	14
	20	22	12	22	9	10	6	10
	17	20	11	21	9	11	6	12
$\Gamma(5)$	34	35	21	35	20	23	14	23
	33	35	21	36	18	21	13	21
	30	35	20	35	17	21	12	22
$\Gamma(10)$	50	51	34	50	34	36	24	35
	48	49	33	49	31	34	22	33
	47	50	33	50	29	33	22	34
$LN(0.5)$	78	80	63	80	52	58	46	59
	77	79	63	80	43	52	38	53
	73	79	61	78	48	56	40	55
$LN(1)$	63	76	57	76	29	40	27	41
	60	74	55	74	24	19	22	20
	58	73	52	73	24	33	22	34
$LN(2)$	46	62	42	63	17	14	16	14
	42	52	41	53	17	6	16	6
	39	51	36	51	15	9	15	10
$\chi^2(6)$	22	23	12	22	11	14	8	14
	20	21	11	21	9	10	6	10
	17	21	11	21	8	11	6	11
$\chi^2(8)$	27	28	16	29	16	19	11	19
	27	28	16	28	13	16	9	16
	24	28	16	29	12	16	9	16
$\chi^2(10)$	33	34	21	35	20	23	14	23
	32	34	20	34	17	21	12	20
	29	34	20	34	17	20	12	21
$\beta(1, 1)$	93	93	32	56	70	71	22	38
	93	93	32	55	56	64	16	33
	93	93	32	54	68	70	20	38
$\beta(0.5, 1)$	92	92	30	51	20	42	4	20
	91	91	30	51	2	14	0	15
	90	90	29	50	15	38	2	19
$\beta(1, 2)$	35	38	16	26	7	14	3	9
	36	39	16	26	3	9	2	7
	36	40	16	26	7	14	3	10
Average sizes	3	4	2	4	2	3	2	2
Average powers	48	52	29	44	23	27	15	23

Table 19: Averages of estimated powers for the gamma distribution for 10% and 30% censoring for a sample size of $n = 50$ with three different censoring distributions.

F	FF	FT	TF	TT	FF	FT	TF	TT
$\Gamma(0.5)$	3	3	3	3	2	2	4	2
	4	3	4	3	3	2	4	2
	4	3	3	3	3	2	4	2
$\Gamma(1)$	3	3	3	3	2	2	3	2
	4	3	3	3	3	2	4	1
	3	3	3	3	2	2	3	2
$\Gamma(2)$	4	3	2	2	3	2	2	2
	4	3	3	2	2	2	3	2
	4	4	3	2	2	2	3	2
$W(2)$	8	10	6	6	3	5	3	3
	8	10	6	6	3	4	3	3
	7	10	6	5	3	4	3	3
$W(3)$	14	15	8	8	6	8	5	6
	14	15	9	8	5	6	4	5
	13	15	8	8	5	7	4	5
$LN(1)$	25	31	17	30	11	13	10	13
	23	29	18	30	11	8	10	8
	22	29	18	29	11	11	9	11
$LN(1.5)$	37	45	31	44	15	16	14	16
	35	44	32	44	13	8	12	8
	32	42	30	42	13	10	13	12
$LN(2.5)$	46	52	44	53	16	12	15	12
	46	51	45	51	12	8	12	8
	42	48	41	47	14	8	13	8
$\chi^2(0.5)$	4	3	4	3	3	2	4	2
	3	3	4	2	3	2	4	2
	4	3	4	3	3	2	4	2
$\chi^2(0.75)$	4	3	4	3	3	2	4	2
	3	3	4	2	3	2	4	2
	4	3	4	3	3	2	4	2
$\beta(1, 1)$	50	50	41	23	20	28	13	13
	49	49	40	22	13	20	12	12
	50	50	40	22	17	27	12	12
$\beta(0.5, 1)$	28	30	26	16	4	6	7	6
	26	30	25	16	1	4	8	5
	26	30	25	16	3	6	7	5
$\beta(1, 2)$	17	20	12	11	4	6	4	4
	16	19	13	11	2	4	4	4
	15	19	12	10	3	5	4	4
$\mathcal{L}(0.25)$	5	6	4	4	2	3	3	2
	4	6	4	4	2	2	3	2
	4	5	4	4	2	2	3	2
$\mathcal{L}(0.5)$	4	6	4	4	2	2	2	2
	4	5	3	4	2	2	3	2
	3	5	3	4	2	2	3	2
Average sizes	4	3	3	3	2	2	3	2
Average powers	19	22	17	17	7	7	7	6

Table 20: Averages of estimated powers for the gamma distribution for 10% and 30% censoring for a sample size of $n = 100$ with three different censoring distributions.

F	FF	FT	TF	TT	FF	FT	TF	TT
$\Gamma(0.5)$	3	3	3	3	3	2	4	2
	4	3	4	3	3	2	4	2
	3	3	3	3	3	2	4	2
$\Gamma(1)$	4	4	3	3	2	2	3	2
	3	4	3	3	4	2	4	2
	3	4	3	3	3	2	3	2
$\Gamma(2)$	3	3	2	3	3	3	2	2
	4	4	3	3	3	2	3	2
	4	3	3	3	3	3	3	2
$W(2)$	15	18	9	13	5	10	5	6
	14	17	9	13	4	7	5	5
	12	17	9	12	5	8	4	6
$W(3)$	27	27	16	15	12	16	9	10
	26	27	16	16	9	14	8	10
	26	28	16	15	10	15	8	10
$LN(1)$	46	49	37	50	22	32	20	32
	44	50	38	51	21	19	20	19
	41	50	34	50	19	27	18	27
$LN(1.5)$	59	63	56	63	28	39	27	40
	55	65	54	65	22	18	23	20
	54	64	51	64	24	29	23	30
$LN(2.5)$	61	69	60	69	29	31	28	33
	60	72	60	71	22	16	22	17
	58	69	57	69	24	23	24	23
$\chi^2(0.5)$	4	3	4	2	3	2	5	3
	3	2	4	2	2	2	2	2
	3	2	4	2	3	2	4	3
$\chi^2(0.75)$	3	3	3	3	3	2	4	2
	4	3	4	3	3	2	4	2
	3	3	4	3	3	2	4	2
$\beta(1, 1)$	66	66	48	33	59	60	36	24
	66	66	48	33	44	52	30	23
	66	65	49	34	57	58	34	24
$\beta(0.5, 1)$	55	56	46	32	12	22	13	14
	55	56	45	30	2	9	10	12
	55	56	46	31	11	20	11	14
$\beta(1, 2)$	35	37	30	23	9	17	8	10
	35	38	29	23	4	9	6	8
	35	38	28	22	7	14	8	10
$\mathcal{L}(0.25)$	6	8	4	8	3	4	4	4
	5	8	4	8	3	3	3	3
	4	8	4	7	2	3	3	3
$\mathcal{L}(0.5)$	6	8	4	7	3	4	3	3
	5	8	4	7	2	3	3	2
	4	8	4	7	2	3	3	3
Average sizes	3	3	3	3	3	2	3	2
Average powers	31	34	26	27	14	17	12	13

CHAPTER 5

Article 4: On an omnibus test for the parametric proportional hazards model

The conference proceedings, *On an omnibus test for the parametric proportional hazards model*, has been accepted for publication as part of the proceedings for the 2021 SASA conference. A summary of the guidelines to authors from the journal is now presented.

Manuscript	Authors must limit the submission to eight pages or less.
Title	The title should be concise and informative.
Abstract and keywords	The aim of the abstract is to provide a concise description of your article. It should preferably be no more than 300 words and contain a minimum of symbols and references. Keywords should be listed below the abstract and should appear in alphabetical order, each keyword starting with a capital letter and separated from the previous keyword by a comma.
Tables	Table names and descriptions are placed directly above the table. Tables are numbered using Arabic numerals in the order they are cited. Table labels start with the word “Table” and the number is followed by a full stop, all in bold. Then a description follows in normal font, terminated by a full stop.
References	Only references cited in the text should be included. No section number is used. Use the small caps font for the author names. The titles of books are typeset in title case whereas the titles of journal articles employ normal sentence case. Journal titles are written out in full and not abbreviated.
General formatting	A LaTeX template is provided for submission.
Additional information	https://tinyurl.com/3pcwnk7u

ON AN OMNIBUS TEST FOR THE PARAMETRIC PROPORTIONAL HAZARDS MODEL

JS Allison

Subject Group Statistics, North-West University, Potchefstroom, South Africa

E Bothma

Subject Group Statistics, North-West University, Potchefstroom, South Africa

M Smuts

Subject Group Statistics, North-West University, Potchefstroom, South Africa

IJH Visagie¹

Subject Group Statistics, North-West University, Potchefstroom, South Africa
e-mail: *jaco.visagie@nwu.ac.za*

We propose an omnibus test of fit for the parametric Cox proportional hazards model in the presence of random right censoring. The proposed test results from a modification of an existing test for the uniform distribution. This test is demonstrated to be able to detect deviations from the hypothesised model in two cases; first when the baseline distribution is misspecified and second when the regression component of the model is misspecified. Two modified classical tests are considered and a Monte Carlo study shows that the newly proposed test outperforms these tests for the majority of alternatives included. As a result of independent interest, we outline the procedure required to use the newly modified test in the framework of independent and identically distributed random variables.

Key words: Kaplan-Meier estimate, parametric bootstrap, proportional hazard model, uniformity.

1. Introduction

The parametric Cox proportional hazards (CPH) model is widely used in medical research as well as financial applications; see, for example, Bover Hidiroglu, Arellano and Bentolila Chocrón (1996), Klein and Moeschberger (2006) as well as Smuts and Allison (2020). In these fields it is often important to study the effect of explanatory variables on time-to-event outcomes. The parametric CPH model specifies the baseline survival function as well as the form of the regression function of the model. As a result the fitted model includes parameters which allow for a simple interpretation which is a desirable feature when using the model in practice. However, in order to ensure that the inference drawn from the fitted model is valid, one has to test the assumptions of the model. Testing the validity of the model requires simultaneously testing the assumption of the specified baseline distribution and the regression function included in the model. A testing procedure for this

¹Corresponding author.

MSC2010 subject classifications. 62F03, 62F40, 62N01.

goodness-of-fit problem has been proposed for use with full samples in Cockeran, Meintanis and Allison (2019).

Random right censoring frequently occurs in the fields mentioned above, meaning that the assumptions of the parametric CPH model should be tested based on censored data. Consider, for example, a medical study estimating the distribution of lifetimes of patients with a specific disease. For a given patient this distribution will likely depend on several covariates such as the patient's smoking habits. Some of the patients' lifetime's will be observed while other will not; for instance, one of the patients may relocate to another country and leave the study, meaning that this patient's lifetime is censored. In this case testing the goodness-of-fit of the parametric CPH model is complicated by the presence of the censoring mechanism.

Lin and Spiekerman (1996) proposes a test for the baseline distribution of the parametric CPH model. However, there are situations in which one wishes to not only test the assumption of the baseline distribution, but the entire model simultaneously. In this case, it is possible to modify classical tests such as the Kolmogorov-Smirnov and Cramér-von Mises in order to test the fit of the entire model; this procedure is illustrated in this paper. In addition, we also propose a new test for the parametric CPH model based on a test for uniformity found in Meintanis (2009). The main idea of the test rests on fitting the CPH model and applying the inverse of the probability integral transform to obtain a censored sample of variables, the lifetime distribution of which is approximately uniform under the null hypothesis. These statements are made exact in Section 2. Furthermore, we modify the mentioned test to accommodate censoring when testing simple hypothesis or composite hypotheses in the classical goodness-of-fit framework of independent and identically distributed (i.i.d.) random variables; the required procedures are outlined later in the paper.

Some notation is introduced before proceeding. Let $Y_1, \dots, Y_n$ be independent lifetime variables with continuous distribution and survival functions F and S respectively and let $C_1, \dots, C_n$ be i.i.d. censoring variables with distribution function G , independent of $Y_1, \dots, Y_n$. We assume non-informative censoring throughout. Below we denote the transpose of a vector v by $v^\top$. Let

$$T_j = \min(Y_j, C_j), \quad \delta_j = \begin{cases} 1, & \text{if } Y_j \leq C_j, \\ 0, & \text{if } Y_j > C_j, \end{cases} \quad \text{and } \mathbf{x}_j = (x_{j,1}, \dots, x_{j,m})^\top, j = 1, \dots, n,$$

where $\mathbf{x}_j$ is a possible vector valued set of m realised covariates. Further let $\tilde{S}(t|\mathbf{x}_j)$ denote the conditional survival function given an observed set of covariates. Based on the observed triplets $(T_j, \delta_j, \mathbf{x}_j)$, $j = 1, \dots, n$ we wish to test the composite hypothesis that the conditional survival function can be modelled by a parametric Cox model, i.e.

$$H_0 : \tilde{S}_{\theta, \beta}(t|\mathbf{x}_j) = S_\theta(t)e^{\beta^\top \mathbf{x}_j}, \quad (1)$$

where S_θ is a specified parametric baseline survival function indexed by a parameter θ (possibly a vector) and $\beta = (\beta_1, \dots, \beta_m)^\top$ denotes the vector of regression parameters associated with the covariates. Denote the order statistics of $Y_1, \dots, Y_n$ and $T_1, \dots, T_n$ by $Y_{(1)} < \dots < Y_{(n)}$ and $T_{(1)} < \dots < T_{(n)}$, respectively. Note that $\delta_{(j)}$ and $\mathbf{x}_{(j)}$ represents the indicator variable and set of covariates corresponding to $T_{(j)}$, respectively.

The remainder of the paper is structured as follows. In Section 2 we indicate how three existing tests are modified to accommodate random right censoring. Since the null distribution of each of

the test statistics depends on the unknown censoring distribution, we propose a parametric bootstrap procedure in Section 2 in order to compute critical values for the tests under consideration. Section 3 contains the results of a Monte Carlo study where the empirical powers of the newly modified tests are studied and compared. This Monte Carlo study considers various deviations from the null hypothesis. Section 4 concludes and provides some directions for further research.

2. The proposed test

Under the null hypothesis in (1), $\tilde{S}_{\beta, \theta}(Y_i | \mathbf{x}_i)$ follows a standard uniform distribution (i.e. uniform on the interval (0,1)). This is a direct consequence of the probability integral transform. Let $\hat{\beta}$ and $\hat{\theta}$ denote consistent estimators, under H_0 , for β and θ , respectively (throughout the paper we make use of maximum likelihood estimation). Whence $\tilde{S}_{\hat{\beta}, \hat{\theta}}(Y_i | \mathbf{x}_i)$ should approximately follow a standard uniform distribution. Let

$$\hat{U}_j = \tilde{S}_{\hat{\beta}, \hat{\theta}}(T_j | \mathbf{x}_j) = S_{\hat{\theta}}(T_j) e^{\hat{\beta}^\top \mathbf{x}_j}.$$

Under H_0 , $(\hat{U}_j, \delta_j)$ are censored observations for which the lifetime distribution is approximately standard uniform.

Above we have demonstrated that any test for censored uniformity on the basis of $(\hat{U}_j, \delta_j)$ is in effect a test of the fit of the CPH model itself. Note that there are a number of misspecifications of the model in (1) that could lead to the rejection of the null hypothesis. These include an incorrect choice of the baseline distribution, a different form of the parametric regression function (e.g., $1 + \beta^\top \mathbf{x}$ instead of $e^{\beta^\top \mathbf{x}}$) and covariates dependent on time. Due to page restrictions, we investigate only the first two of these types of deviations in the Monte Carlo study in Section 3.

To test for uniformity we will use the Kolmogorov-Smirnov (KS_n) and Cramér-von Mises (CV_n) tests which are modified in Koziol and Green (1976) and Barr and Davidson (1973), respectively, for the random censoring case. In addition, we also modify a complete sample test for uniformity proposed in Meintanis (2009), where a test statistic based on the difference between the characteristic function (CF) of the standard uniform distribution and the empirical characteristic function is introduced. The proposed test statistic in the complete sample case is

$$R_{n,a} = n \int |\hat{\varphi}_n(t) - \varphi_u(t)|^2 e^{-a|t|} dt, \quad (2)$$

where $\hat{\varphi}_n(t) = \int_{-\infty}^{\infty} e^{itx} dF_n(x) = n^{-1} \sum_{j=1}^n e^{itY_j}$ is the empirical CF, $\varphi_u(t) = t^{-1}(\sin t + i(1 - \cos t))$ is the CF of the standard uniform distribution, $F_n(x)$ is the empirical distribution function, and $a > 0$ is a user defined tuning parameter. In the presence of random right censoring, we replace $\hat{\varphi}_n(t)$ in (2) with

$$\tilde{\varphi}_n(t) = \int_{-\infty}^{\infty} e^{itx} d\tilde{F}_n(x),$$

where $\tilde{F}_n(x)$ is the Kaplan-Meier estimate of the distribution F . After some straightforward algebra,

the modified test statistic can be shown to have the calculable form

$$\begin{aligned}\tilde{R}_{n,a} = & \sum_{j=1}^n \sum_{k=1}^n \Delta_j \Delta_k \frac{2a}{(\hat{U}_j - \hat{U}_k)^2 + a^2} + 4 \tan^{-1} \left(\frac{1}{a} \right) - 2a \log \left(1 + \frac{1}{a^2} \right) \\ & - 4 \sum_{j=1}^n \left[\tan^{-1} \left(\frac{\hat{U}_j}{a} \right) + \tan^{-1} \left(\frac{1 - \hat{U}_j}{a} \right) \right],\end{aligned}$$

with

$$\Delta_j = \frac{\delta_{(j)}}{n} \prod_{k=1}^{j-1} \left[1 + \left(1 - \frac{\delta_{(k)}}{n-k} \right) \right], \quad j = 1, \dots, n-1, \quad \text{and} \quad \Delta_n = \prod_{j=1}^{n-1} \left(\frac{n-j}{n-j+1} \right),$$

where an empty product is understood to be 1. Note that Δ_j corresponds to the size of the jump of the Kaplan-Meier estimate of the distribution function at T_j . The test rejects for large values of $\tilde{R}_{n,a}$.

In the current context the null distributions of all test statistics depend on the unknown censoring distribution G . As a result, when approximating the null distribution of a test statistic, we should take care to estimate G and sample censoring times from this estimated distribution. As a result, we propose the bootstrap algorithm below which can be used to estimate the critical value of each of the tests. For ease of notation, let $Q := Q((\hat{U}_1, \delta_1), \dots, (\hat{U}_n, \delta_n))$ be a generic test statistic.

1. Given $(T_j, \delta_j, \mathbf{x}_j), j = 1, \dots, n$, calculate $\hat{\theta}$ and $\hat{\beta}$ under H_0 .
2. Calculate the Kaplan-Meier estimate, $\tilde{G}_n$, of the censored distribution function G .
3. Sample $C_1^*, \dots, C_n^*$ from $\tilde{G}_n$.
4. Generate $Y_1^*, \dots, Y_n^*$ from $S_{\hat{\theta}, \hat{\beta}}^{\sim}$. This is accomplished by generating $U_1, \dots, U_n$ from a standard uniform distribution and setting $Y_j^* = S_{\hat{\theta}, \hat{\beta}}^{-1} \left((U_j) e^{-\hat{\beta}^\top \mathbf{x}_j} \right), j = 1, \dots, n$.
5. Let $T_j^* = \min(Y_j^*, C_j^*)$ and $\delta_j^* = I(Y_j^* \leq C_j^*), j = 1, \dots, n$.
6. Calculate $\hat{\theta}^*$ and $\hat{\beta}^*$ based on $(T_j^*, \delta_j^*, \mathbf{x}_j^*), j = 1, \dots, n$, under H_0 .
7. Let $\hat{U}_j^* = \tilde{S}_{\hat{\theta}^*, \hat{\beta}^*}(T_j^* | \mathbf{x}_j^*), j = 1, \dots, n$.
8. Calculate the test statistic $Q^* = Q((\hat{U}_1^*, \delta_1^*), \dots, (\hat{U}_n^*, \delta_n^*))$.

9. Repeat steps 3-8 B times to obtain $Q_1^*, \dots, Q_B^*$. Let $Q_{(1)}^*, \dots, Q_{(B)}^*$ denote the ordered values of the test statistics.

10. The estimated critical value is $\hat{C}_{n,B} = Q_{(\lfloor B(1-\alpha) \rfloor)}^*$, with $\lfloor \cdot \rfloor$ denoting the floor function.

By simply removing all reference to the covariates $\mathbf{x}_j, j = 1, \dots, n$ from the above algorithm, we obtain an algorithm which can be used to approximate critical values for the newly proposed test in the classical goodness-of-fit testing framework (in which observations are assumed to be realisations from i.i.d. random variables). In this framework, we can test a simple hypothesis by transforming the observed sample using the specified distribution function. In the case where the null hypothesis specifies a parametric family of distributions, we may estimate the necessary parameters and perform the requisite transformation using the estimated parametric distribution function (as is done above). In either case, if the null hypothesis is true, then we obtain censored realisations, the lifetime distribution of which should be approximately standard uniform and the testing procedure proposed above remains valid.

3. Monte Carlo simulation

In this section the finite sample performance of KS_n , CM_n and $\tilde{R}_{n,a}$ is analysed. The KS_n and CM_n tests are ubiquitous in the literature and we do not provide their details here, we merely note that these tests are obtained by replacing the empirical distribution function used in the classical goodness-of-fit testing framework by the Kaplan-Meier estimate.

The methodology described in the previous section is general and can be applied to test for any specified baseline distribution. However, due to the page limitations imposed on this publication, the results shown below are limited to the case where baseline distribution is Weibull. This baseline is chosen due to its flexibility and its popularity in practical applications.

We present empirical sizes and powers of the tests for two scenarios below.

Scenario 1

The null hypothesis H_0 corresponds to the CPH model with a certain baseline distribution against an alternative H_A with a different baseline distribution, i.e., we consider the powers achieved by the various tests when the baseline distribution is misspecified while the regression component of the model is specified correctly.

Scenario 2

Under H_0 and H_A the model has the same baseline distribution, but the regression function under H_A is either $1 + \beta^\top \mathbf{x}_j$ or $\log(1 + \beta^\top \mathbf{x}_j)$, i.e., we consider the powers achieved when the baseline is correctly specified but the regression component of the model is misspecified.

3.1 Setting

The nominal significance level is set to 10%. We employ the so-called “warp-speed” bootstrap discussed in Giacomini, Politis and White (2013) to obtain the empirical powers reported in this section. This methodology is popular for the calculation of empirical powers; see, for example, Allison, Betsch, Ebner and Visagie (2019). All size and power estimates are calculated based on 50 000 independent Monte Carlo samples for sample sizes $n = 100$ and $n = 200$. The reported

powers are calculated in the case of 10% and 20% censoring. As censoring distribution we use a uniform distribution on the interval $(0, b)$, where b is a suitably chosen constant so as to achieve the desired censoring proportion (obtained using a simple optimization algorithm).

Throughout the study we use a single continuous covariate x_j , randomly generated from a standard uniform distribution. The alternative distributions used for the baseline survival function are displayed in Table 1.

Table 1. Density functions of the baseline distributions.

Baseline distribution	Density	Notation
Weibull	$\theta\lambda(x\lambda)^{\theta-1}\exp(-(\lambda x)^\theta)$	$W(\lambda, \theta)$
Gamma	$(\Gamma(\theta))^{-\theta}\lambda^\theta x^{\theta-1}\exp(-\lambda x)$	$\Gamma(\lambda, \theta)$
Lognormal	$(\theta x\sqrt{2\pi})^{-1}\exp(-\log^2(x-\lambda)(2\theta^2)^{-1})$	$LN(\mu, \sigma^2)$
Chi square	$(2^{v/2}\Gamma(v/2))^{-1}x^{v/2-1}\exp(-x/2)$	$\chi^2(v)$

All calculations and simulations are performed in R, see R Core Team (2019). All parameter estimation is performed using maximum likelihood estimation. For this purpose we use Brent optimisation, see Brent (1971), built into R's *optim* function.

For ease of interpretation, we print the highest power (including ties) against each deviation from the null hypothesis in bold in each of the power tables presented in this section.

Scenario 1 above, the misspecification of the baseline distribution, is considered in Tables 2 and 3. Table 2 shows empirical sizes and powers when the censoring proportion is set to 10%, while Table 3 contains sizes and powers in the presence of 20% censoring. When comparing the powers in these two tables a number of remarks are in order. First, when considering the size of the test shown in the first line of these tables, we see that the tests typically reject in fewer instances than the 10% specified by the significance level, especially in the case where the sample size is 100. Second, as expected, the powers of the tests generally increase with sample size and decrease with the censoring proportion. Third, although none of the tests outperform the others uniformly, $\tilde{R}_{n,0.2}$ and CV_n are clearly quite powerful when compared to their competitors.

Table 2. Empirical powers against misspecified baseline distributions with 10% censoring

Distribution	KS_{100}	CV_{100}	$\tilde{R}_{100,.2}$	$\tilde{R}_{100,.5}$	KS_{200}	CV_{200}	$\tilde{R}_{200,.2}$	$\tilde{R}_{200,.5}$
$W(2, 2)$	7	6	6	6	8	7	8	7
$\Gamma(1, 3)$	15	13	15	14	24	22	24	24
$\chi^2(10)$	22	21	23	23	36	37	39	39
$LN(0.5, 4)$	33	27	4	36	56	49	73	62

Next we turn our attention to *Scenario 2* above, the misspecification of the regression component. This scenario is considered in Tables 4 and 5. Again, the results associated with 10% and 20% censoring are shown in two separate tables. As before, the powers of the tests generally increase with sample size and decrease with the censoring proportion. Again, $\tilde{R}_{n,0.2}$ and CV_n prove to be the most powerful tests. In addition to comparing the powers for different specifications of the regression

Table 3. Empirical powers against misspecified baseline distributions with 20% censoring

Distribution	KS_{100}	CV_{100}	$\tilde{R}_{100,.2}$	$\tilde{R}_{100,.5}$	KS_{200}	CV_{200}	$\tilde{R}_{200,.2}$	$\tilde{R}_{200,.5}$
$W(2,2)$	8	7	8	6	10	9	10	8
$\Gamma(1,3)$	11	11	11	11	17	16	17	16
$\chi^2(10)$	15	16	16	15	24	26	27	25
$LN(0.5,4)$	16	8	17	13	28	13	31	24

component of the model, Tables 4 and 5 show powers associated with different choices for the value of β ; specifically, this parameter is set either to 1 or 3. We note that the powers achieved against the different values of β are very similar for the cases considered.

Table 4. Empirical powers against misspecified regression function with 10% censoring

Regression function	KS_{100}	CV_{100}	$\tilde{R}_{100,.2}$	$\tilde{R}_{100,.5}$	KS_{200}	CV_{200}	$\tilde{R}_{200,.2}$	$\tilde{R}_{200,.5}$
$1+x$	18	18	18	16	27	29	29	24
$1+3x$	18	19	19	17	28	30	30	25
$\log(1+x)$	12	16	13	13	15	23	17	18
$\log(1+3x)$	11	15	12	12	13	21	16	16

Table 5. Empirical powers against misspecified regression function with 20% censoring

Regression function	KS_{100}	CV_{100}	$\tilde{R}_{100,.2}$	$\tilde{R}_{100,.5}$	KS_{200}	CV_{200}	$\tilde{R}_{200,.2}$	$\tilde{R}_{200,.5}$
$1+x$	26	23	28	19	45	40	48	32
$1+3x$	26	23	29	19	44	40	48	32
$\log(1+x)$	9	11	11	8	14	19	19	12
$\log(1+3x)$	9	11	12	8	15	18	19	12

4. Conclusion

The Monte Carlo study presented above shows that the CV_n and $\tilde{R}_{n,0.2}$ tests are especially powerful. When considering the empirical powers associated with a sample size of 100, CV_n outperforms the competing tests, while the $\tilde{R}_{n,0.2}$ is more powerful when the sample size is increased to 200. The powers associated with the smaller of these sample sizes do not differ much between the mentioned tests, while there is a pronounced difference in the case of the larger sample size. As a result, we recommend using $\tilde{R}_{n,a}$ for practical purposes.

For practical implementation of the new test one needs a so-called compromise choice of the tuning parameter a . This is a choice of the tuning parameter that provides reasonably high power for a wide range of alternatives. Based on the limited Monte Carlo study presented we suggest using $a = 0.2$. However, a more comprehensive power study is needed including different baseline distributions as well as other forms of misspecification in order to fully investigate the role of the parameter a .

The asymptotic properties of the new tests are unknown and should be derived. Specifically one needs to show that the test is consistent against a wide class of alternatives (or misspecifications). Furthermore, deriving the asymptotic null distributions of the test statistics for given censoring distributions are unknown at present. The above are matters of ongoing research.

References

- ALLISON, J. S., BETSCH, S., EBNER, B., AND VISAGIE, I. J. H. (2019). New weighted L^2 -type tests for the inverse Gaussian distribution. *arXiv:1910.14119*.
- BARR, D. R. AND DAVIDSON, T. (1973). A Kolmogorov-Smirnov test for censored samples. *Technometrics*, **15** (4), 739–757.
- BOVER HIDIROGLU, O., ARELLANO, M., AND BENTOLILA CHOCRÓN, S. (1996). Unemployment duration, benefit duration, and the business cycle. *Estudios económicos/Banco de España*, 57.
- BRENT, R. P. (1971). An algorithm with guaranteed convergence for finding a zero of a function. *The Computer Journal*, **14** (4), 422–425.
- COCKERAN, M., MEINTANIS, S. G., AND ALLISON, J. S. (2019). Goodness-of-fit tests in the Cox proportional hazards model. *Communications in Statistics - Simulation and Computation*, 1–12.
- GIACOMINI, R., POLITIS, D. N., AND WHITE, H. (2013). A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory*, **29** (3), 567–589.
- KLEIN, J. P. AND MOESCHBERGER, M. L. (2006). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer Science & Business Media.
- KOZIOL, J. A. AND GREEN, S. B. (1976). A Cramér-von Mises statistic for randomly censored data. *Biometrika*, **63** (3), 465–474.
- LIN, D. AND SPIEKERMAN, C. (1996). Model checking techniques for parametric regression with censored data. *Scandinavian Journal of Statistics*, 157–177.
- MEINTANIS, S. G. (2009). Goodness-of-fit tests and minimum distance estimation via optimal transformation to uniformity. *Journal of Statistical Planning and Inference*, **139** (2), 100–108.
- R CORE TEAM (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL: <https://www.R-project.org/>
- SMUTS, M. AND ALLISON, J. (2020). An overview of survival analysis with an application in the credit risk environment. *ORiON*, **36** (2).

CHAPTER 6

Conclusion

6.1 Overview of the goals of the papers presented

In this thesis we set out to modify existing tests for the complete sample case to accommodate random right censoring for three classes of distributions; the exponential, Weibull and gamma classes of distributions. We further aimed to propose new classes of tests for the Weibull distribution that were applied when using complete samples as well as censored samples. We tested the finite sample performance of the newly modified or proposed tests by conducting extensive Monte Carlo simulation studies. Additionally, we studied the impact of some of the computational assumptions made when estimating the distribution function with the Kaplan-Meier estimator. The main objectives of this thesis, with their corresponding outcomes can be summarised as follows:

- To explore the existing literature to identify work that has already been done on this topic as well as any open questions relating to this.

A review of existing literature was presented in Chapters 2 and 3.

- Modify existing goodness-of-fit tests for exponentiality to accommodate random right censoring and compare the finite sample performance of these tests to existing tests.

Existing tests for exponentiality were modified in Chapter 2

- Develop new tests for the Weibull distribution and compare the performance of these tests against a broad range of alternatives, in both the complete sample case as well as in the presence of random right censoring.

Two new classes of tests were introduced in Chapter 3 along with a modification of an existing test.

- Propose newly modified tests for the gamma distribution in the presence of random right censoring and compare the performance of these tests to those of existing tests for a broad range of alternatives.

In Chapter 4 we introduced modifications of existing tests for the gamma distribution.

- Explore the computational assumptions, relating to the Kaplan-Meier estimator.

The impact of some of these assumptions was extensively explored in Chapter 4

- Provide new bootstrap algorithms, to calculate critical values when testing the adequacy of fit of survival models are proposed and implemented.

In Chapter 5 we presented bootstrap algorithms that can be employed to calculate critical values.

- Develop a new omnibus test for the parametric proportional hazards model.

A new omnibus test for the parametric proportional hazards model in the presence of random censoring was developed in Chapter 5.

6.2 Overview of results

The results and conclusions of the three articles as well as the proceedings prepared for this thesis are now presented:

Characteristic function and Laplace transform based tests for exponentiality in the presence of random right censoring

In the first article we modified four tests for exponentiality, based on either the characteristic function or the Laplace transform, to accommodate random right censoring. These tests are compared to existing tests using a Monte Carlo study and it is found that the newly modified tests are competitive in terms of their power performance. This is especially true for the test based on the squared difference between the Laplace transform and its empirical counterpart, hence this is the recommended test for use in practise.

In the first article we concluded that a more flexible distribution, such as the Weibull distribution, would be a better fit to model the practical data set. After this we explored goodness-of-fit tests for the Weibull distribution in the presence of random censoring and we found that tests are relatively scarce. Therefore, we developed new classes of goodness-of-fit tests for the Weibull distribution in the second article.

New classes of tests for the Weibull distribution using Stein's method in the presence of random right censoring

In the second article we developed two new classes of tests for the Weibull distribution based on Stein's method. The proposed tests were applied in the full sample case as well as in the presence of random right censoring. We investigated the finite sample performance of the new tests using a comprehensive Monte Carlo study. In both the absence and presence of censoring, it is found that the

newly proposed classes of tests outperform competing tests against the majority of the distributions considered. In the cases where censoring is present we considered various censoring distributions. Some remarks on the asymptotic properties of the proposed tests were included. Another result of independent interest was presented; a test, based on the Laplace transform, initially proposed for use with full samples was amended to allow for testing for the Weibull distribution in the presence of censoring. The techniques developed in the paper are illustrated using two practical examples. Based on the observed numerical powers, we recommend using the newly proposed test which is based on Stein's method, when testing the hypothesis under consideration in the presence of censoring.

In the first two articles we make some computational assumptions regarding the Kaplan-Meier estimator; these assumptions are common in the statistical literature. The impact of some of these assumptions was explored in Article 3.

On the effect of the Kaplan-Meier estimator's assumed tail behaviour on goodness-of-fit testing

The numerical results in the article indicate that the set-to-1 assumption required when using the Kaplan-Meier estimator of the distribution function plays an unexpectedly important role when testing goodness-of-fit hypotheses. It should also be noted that this assumption is important not only when estimating the lifetime distribution, but also when estimating the censoring distribution. Furthermore, it should be noted that the highest powers are often achieved when the set-to-1 assumption is made for one of the two distributions being estimated but not the other.

Only goodness-of-fit testing for censored samples in the independently and identically distributed case was explored in the first three articles. In the conference proceedings we introduced a goodness-of-fit test for the parametric proportional hazards model.

On an omnibus test for the parametric proportional hazards model

In the proceedings, we proposed an omnibus test of fit for the parametric Cox proportional hazards model in the presence of random right censoring. The proposed test resulted from a modification of an existing test for the uniform distribution. This test is demonstrated to be able to detect deviations from the hypothesised model when the baseline distribution or the regression component of the model is misspecified. Two modified classical tests are considered and the Monte Carlo study suggested that the Cramér-von Mises as well as the newly proposed tests are especially powerful. We also outlined the procedure required to use the newly modified test in the framework of independent and identically distributed random variables. For practical purposes we recommend using the newly proposed test.

6.3 Concluding remarks and future research

The overall aim of this thesis was to perform Monte Carlo simulation studies to compare the performances of various goodness-of-fit tests. The main comparison was between the modified versions of the classical tests, namely Cramér-von Mises and Kolmogorov-Smirnov tests, and the newly modified tests for the three different classes of distributions considered. The newly modified tests outperform the classical tests for the majority of alternatives considered for all three of the lifetime distributions. These new modifications are derived from tests that are based on either the characteristic function or the Laplace transform. In most cases of random right censoring we would recommend using one of the newly modified test statistics presented in the first three articles.

We now conclude by suggesting some future research possibilities. The tests based on the characteristic function, Laplace transform as well as Stein's method all contain a tuning parameter. In our study we used a grid of values to obtain the recommendations made. It would be interesting to see if a data dependent choice of the tuning parameter would be more suitable, such a method has been proposed for use in the full sample case. Therefore, future research will entail the development of a data dependent method to choose this type of tuning parameter in the presence of random censoring. Furthermore, it is important to explore some of the theoretical aspects of the newly proposed classes of tests, such as the asymptotic null distribution as well as the consistency of the tests. Additionally, these modified tests can be used to test the goodness-of-fit of various survival models in the presence of random right censoring, the proposed tests for exponentiality are particularly useful in this setting. One such model is the conventional Cox proportional hazards model, another is the mixture cure model. The techniques developed in this thesis may be useful when performing goodness-of-fit tests for these models.