An objective comparison between various goodness-of-fit tests for exponentiality

N Smit

orcid.org/ 0000-0002-4570-033X

Dissertation submitted in partial fulfilment of the requirements for the degree *Master of Science* in Statistics at the
North-West University

Supervisor:	Prof JS Allison
Co-Supervisor:	Dr IJH Visagie
Assistant Supervisor:	Prof L Santana

Graduation May 2018
22730656

Acknowledgements

I would like to express my sincere gratitude to:

- Prof J.S. Allison, my supervisor, for his guidance, assistance, expertise and patience.
- Prof L. Santana and Dr I.J.H. Visagie, my co-supervisors, for their willingness to help and advise.
- My parents, for providing me with the opportunity to continue my studies and for their support.
- Close friends and family, for their words of encouragement and motivation when they were needed.

Abstract

Keywords: bootstrap, exponential distribution, goodness-of-fit testing, tuning parameter.

The exponential distribution is a popular model both in practice and in theoretical work. As a result, a multitude of tests based on varied characterisations have been developed for testing the hypothesis that observed data are realised from this distribution. Many of the recently developed tests contain a tuning parameter, usually appearing in a weight function. In this dissertation, we compare the powers of 20 tests for exponentiality, some containing a tuning parameter and some that do not. To ensure an objective comparison between each of the tests, we employ a data-dependent choice of the tuning parameter for those tests that contain these parameters.

The numerical comparisons presented are conducted for various sample sizes and for a large number of alternative distributions. The results of the simulation study show that the test with the best overall performance is the Baringhaus and Henze test, followed closely by the test by Henze and Meintanis; both tests contain a tuning parameter. The score test by Cox and Oakes performs the best among those tests that do not include a tuning parameter.

Opsomming

Sleutelwoorde: skoenlusmetode, eksponensiële verdeling, passingstoetse, verstelbare parameter.

Die eksponensiële verdeling is 'n gewilde model vir praktiese toepassing sowel as teoretiese werk. As gevolg daarvan is menigte toetse, gebaseer op verskeie karakteriserings, ontwikkel om die hipotese te toets dat waargenome data vanuit hierdie verdeling gerealiseer is. Heelwat van die mees onlangs ontwikkelde toetse bevat 'n verstelbare parameter, wat gewoonlik in 'n gewigsfunksie voorkom. In hierdie verhandeling vergelyk ons die onderskeidingsvermoë van 20 toetse vir eksonensialiteit, sommige van die toetse bevat 'n verstelbare parameter en ander nie. Om 'n objektiewe vergelyking tussen die toetse te verseker gebruik ons 'n data-afhanklike-keuse van die verstelbare parameter vir die toetse wat wel hierdie parameter bevat.

Die numeriese vergelykings word getref gebaseer op verskeie steekproefgroottes en vir 'n groot aantal alternatiewe verdelings. Die resultate van die simulasiestudie wys dat die Baringhaus en Henze-toets die algehele beste prestasie toon, maar die waargenome onderskeidingsvermoë van die toets van Henze en Meintanis is vergelykbaar. Beide hierdie toetse bevat verstelbare parameters. Die tellingstoets van Cox en Oakes was die beste presteerder onder die toetse wat nie verstelbare parameters bevat nie.

Contents

Chapter 1

Introduction and motivation

The exponential distribution is a popular choice of model both in practice and in theoretical work. For this reason a great deal of research has been dedicated to the large number of ways in which it can be uniquely characterised. This has ultimately led to a multitude of tests for testing the hypothesis that observed data are realised from the exponential distribution.

Several authors have written review papers on this topic, describing and comparing a number of tests, see, for example, Spurrier (1984), Ascher (1990), and Henze and Meintanis (2002). However, the most recent review paper on this topic was written more than 10 years ago by Henze and Meintanis (2005). Since then, a number of new tests have been proposed, see for example Jammalamadaka and Taufer (2006), Haywood and Khmaladze (2008), Mimoto and Zitikus (2008), Wang (2008), Volkova (2010), Grané and Fortiana (2011), Abbasnejad et al. (2012), Baratpour and Habibi Rad (2012), Volkova and Nikitin (2013), Meintanis et al. (2014), and Zardasht et al. (2015).

Furthermore, many of the tests for exponentiality contain a tuning parameter, often appearing in a weight function. The fact that the powers of these tests are functions of the tuning parameter complicates the comparisons between tests. In many papers the authors evaluate the power of the test over a grid of possible values of this parameter, but the problem with this approach is that the optimal choice of the tuning parameter is unknown in practice. In these papers the authors often provide a so-called ‘compromise’ choice; this is a choice of the tuning parameter that provides reasonably high power for the majority of the alternatives considered in their finite sample studies. Examples of papers that contain these compromise choices include Henze and Meintanis (2002), Henze and Meintanis (2005), and Meintanis et al. (2014). However, while these fixed choices of the parameter are able to produce high powers against a number of alternatives, they can also produce abysmally low powers against other alternatives. Naturally, in practice, the distribution of the realised data is unknown, meaning that the power of tests employing the compromise choice might be suspect.

A method to choose the value of the tuning parameter data-dependently is proposed in Allison and Santana (2015). This approach removes the practical problem of choosing the tuning parameter and also allows one to directly compare the powers achieved by various goodness-of-fit tests.

The aim of this dissertation is to objectively compare the powers of various tests for exponentiality. Where applicable, the methodology detailed in Allison and Santana (2015) is used in order to choose the value of the tuning parameter data-dependently; this allows a fair '*objective*' comparison between the tests containing a tuning parameter and those without one.

The remainder of this dissertation is organised as follows: In Chapter 2, numerous goodness-of-fit tests for exponentiality, based on a variety of characterisations of the exponential distribution, are discussed. The bootstrap is discussed in Chapter 3. The topics included here is the bootstrap principle and hypothesis testing using the bootstrap. Chapter 4 presents the results of an extensive Monte Carlo study of the empirical powers of the majority of the tests discussed in Chapter 2 against numerous alternatives to the exponential distribution (These distributions are classified according to the shape of the corresponding hazard rates). This chapter also discusses the data-dependent choice of the tuning parameter and includes a real-world example to which the tests considered in the simulation study are applied. The dissertation is concluded with some final remarks in Chapter 5.

One publication has already stemmed from this research, namely "An 'apples to apples' comparison of various tests for exponentiality", and was published in Computational Statistics (see Allison et al., 2017).

Chapter 2

Goodness-of-fit testing for the exponential distribution

Let $X_1, X_2, ..., X_n$ be a sequence of independent and identically distributed continuous realisations of a random variable X. Denote the exponential distribution with expectation $1/\lambda$ by $Exp\,(\lambda)$. The composite goodness-of-fit hypothesis to be tested is

$$H_0 : \text{the distribution of } X \text{ is } Exp\,(\lambda)\,,$$

for some $\lambda > 0$, against general alternatives.

The majority of the test statistics that will be considered are based on the scaled values $Y_j = X_j\hat{\lambda}$, where $\hat{\lambda} = 1/\bar{X}_n$ with $\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$. The use of scaled values is motivated from the invariance property of the exponential distribution with respect to scale transformations. Since X follows an exponential distribution if, and only if, cX is exponentially distributed for every $c > 0$, one would not expect a scale transformation to influence the conclusion drawn regarding the exponentiality of X. As a result, the test statistic depends on the data only through scaled versions of the original data, and the conclusions drawn regarding the exponentiality of $X_1, ..., X_n$ and $Y_1, ..., Y_n$ should be the same. In the remainder of the paper, denote the order statistics of X_j and Y_j by $X_{(1)} < X_{(2)} < ... < X_{(n)}$ and $Y_{(1)} < Y_{(2)} < ... < Y_{(n)}$ respectively.

In this chapter short descriptions of the tests for exponentiality is provided, that will be compared to one another in the Monte Carlo study in Chapter 4. These tests are arranged according to the characteristics of the exponential distribution that the test is based on. These tests are chosen because they provide a diverse selection of established tests (tests that have been shown to perform well in terms of power) and newly developed tests, and simultaneously considering tests that contain a tuning parameter as well as those that do not.

In addition to the tests presented in this chapter, references are provided to numerous other tests for exponentiality not included in this study.

2.1 Tests based on the empirical characteristic function

In recent years many goodness-of-fit tests have been developed which are based on the characteristic function (CF). Typically in these tests the CF of a random variable X, given by

$$\phi(t) = E\left[e^{itX}\right],$$

is estimated by the empirical characteristic function (ECF) of the data $X_1, ..., X_n$, defined as

$$\phi_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{itX_j}.$$

Standard methods for testing that employ the ECF utilise the L2-type distance

$$\int_{-\infty}^{\infty} |\phi_n(t) - \phi(t)|^2 w_\gamma(t)dt,$$

which incorporates the CF, ECF and a parametric weight function $w_\gamma(\cdot)$, which usually satisfy the conditions $\int_{-\infty}^{\infty} t^2 w_\gamma(t)dt < \infty$, $w_\gamma(t) = w_\gamma(-t)$, and $w_\gamma(t) \geq 0, \ \ \forall\ t$, and is indexed by a tuning parameter γ.

There has been considerable discussion in the literature on the choice of $w_\gamma(t)$. Popular choices are $w_\gamma(t) = e^{-\gamma|t|}$ or $w_\gamma(t) = e^{-\gamma t^2}$. Both of these correspond to kernel-based choices with $e^{-\gamma|t|}$ being a multiple of the standard Laplace density as kernel with bandwidth equal to $1/\gamma$ and $e^{-\gamma t^2}$ a multiple of the standard normal density as kernel with bandwidth equal to $1/(\gamma\sqrt{2})$.

For various tests for exponentiality that incorporate the ECF, the interested reader is referred to Henze and Meintanis (2002) and Henze and Meintanis (2005) and the references therein. However, for the purposes of this dissertation we will only focus on the 'Epps and Pulley' test proposed in Epps and Pulley (1986) and a more recent test based on the concept of the probability weighted empirical characteristic function (PWECF) proposed in Meintanis et al. (2014).

Epps and Pulley (1986) test (EP_n)

The test proposed in Epps and Pulley (1986) is based on the difference between the ECF, $\phi_n(t)$, of $X_1, X_2, \ldots, X_n$ and the CF of the exponential distribution, $\phi_0(t, \lambda) = \lambda/(\lambda - it)$. If the data are exponentially distributed with parameter λ, then $\phi_n(t)$ should be close to $\phi_0(t, \lambda)$.

Estimating λ by $\widehat{\lambda} = 1/\bar{X}_n$, the test is based on the idea that the quantity

$$\int_0^{\infty} \left(\phi_n(t) - \phi_0(t, \frac{1}{\bar{X}_n})\right) w(t)dt,$$

should be small under the null hypothesis, where

$$w(t) = \frac{1}{2\pi(1 + i\bar{X}_n t)}.$$

The normalised Epps and Pulley test statistic simplifies to

$$\begin{aligned} EP_n =& \sqrt{48n}\int_0^\infty \left(\phi_n(t) - \frac{1}{1-i\bar{X}_n t}\right)\frac{\bar{X}_n}{2\pi(1+i\bar{X}_n t)}dt \\ =& \sqrt{48n}\left[\frac{1}{n}\sum_{j=1}^n e^{-Y_j} - \frac{1}{2}\right]. \end{aligned}$$

This test rejects H_0 for large values of $|EP_n|$. The null distribution of this test statistic was shown to be standard normal in Epps and Pulley (1986). Furthermore, the test was also shown to be consistent against absolutely continuous alternative distributions with monotone hazard rates, strictly positive supports and finite expected values. In a number of studies it has been shown that this test is resonably powerful, see for example Henze and Meintanis (2005).

PWECF ($PW^1_{n,\gamma}$ and $PW^2_{n,\gamma}$)

There has been a lot of discussion regarding the form of the weight function when using goodness-of-fit tests based on the ECF and CF. Fortunately, Meintanis et al. (2014) provides a statistically meaningful way to choose the weight function. This choice reduces the problem to only choosing a tuning parameter γ, typically still contained in the weight function. The probability weighted characteristic function (PWCF) is defined as

$$\chi(t;\gamma) = E\left[W(X;\gamma t)e^{itX}\right] = \int_{-\infty}^{\infty} W(x;\gamma t)e^{itx}dF_\lambda(x),$$

where the probability weight function is given by

$$W(x,\beta) = \left[F_\lambda(x)(1-F_\lambda(x))\right]^{|\beta|}, \qquad \beta\in\mathbb{R},\ x\in\mathbb{R}, \tag{2.1}$$

and where $F_\lambda(\cdot\,)$ denotes the exponential distribution function with parameter λ. Note that the weight function in (2.1) places more weight at the centre of the distribution than in the tails. The probability weighted empirical characteristic function (PWECF) is then defined as

$$\chi_n(t;\gamma) = \frac{1}{n}\sum_{j=1}^n \widehat{W}(X_j;\gamma t)e^{itX_j}, \qquad t\in\mathbb{R}, \tag{2.2}$$

where the estimated probability weight is given by

$$\widehat{W}(X_j;\beta) = \left[F_{\widehat{\lambda}}(x)(1-F_{\widehat{\lambda}}(x))\right]^{|\beta|}, \qquad \beta\in\mathbb{R},\ x\in\mathbb{R},$$

and where $F_{\widehat{\lambda}}(\cdot\,)$ denotes the exponential distribution function with estimated parameter $\widehat{\lambda}$.

Meintanis et al. (2014) employs these expressions and develops a test for exponentiality based on the $L2$-norm between $\chi_n(t;\gamma)$ and $\chi(t;\gamma)$. The resulting test statistic is given by

$$PW^1_{n,\gamma} = n\int_{-\infty}^{\infty}|\chi_n(t;\gamma)-\chi(t;\gamma)|^2dt. \tag{2.3}$$

Note that the weight function that plagues other tests based on the ECF no longer appears in the test statistic, since the weight function has been incorporated within the PWECF and PWCF functions themselves. In Meintanis et al. (2014), the limiting null distribution of the test statistic is derived and it is shown that this test is consistent for a very large class of alternative distributions. In a finite sample simulation study, the test was also found be quite powerful against a variety of alternative distributions.

The test statistic in (2.3) can be simplified to

$$\begin{aligned} PW_{n,\gamma}^{1} = &-\frac{2}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\gamma\ln\left[(1-Z_j)\,Z_j\,(1-Z_k)\,Z_k\right]}{(X_j-X_k)^2+\gamma^2\ln^2\left[(1-Z_j)\,Z_j\,(1-Z_k)\,Z_k\right]} \\ &+\frac{2}{n}\sum_{j=1}^{n}\int_0^1\frac{\gamma\ln\left[(1-Z_j)\,Z_j\,(1-u)\,u\right]}{\left[X_j+\ln(1-u)\right]^2+\gamma^2\ln^2\left[(1-Z_j)\,Z_j\,(1-u)\,u\right]}du, \end{aligned}$$

where $Z_j = \exp(-Y_j)$. In the Monte Carlo simulation study preseted in Meintanis et al. (2014) the power of this test was evaluated over a grid of possible choices of the tuning parameter γ. However, for practical applications the authors suggest using $\gamma = 1$, because this choice fared well for the majority of the alternatives considered in their paper. We will henceforth refer to this type of recommended choice of the parameter as the *compromise choice.*

In Meintanis et al. (2014), the weight function is chosen to give more weight to the centre of the distribution. In this dissertation we also consider a weight function that places greater weight on the tails. This alternative choice for the weight function appearing in (2.2) is given by

$$\widetilde{W}(X_j;\beta) = \left[\frac{1}{4} - F_{\widehat{\lambda}}(x)(1-F_{\widehat{\lambda}}(x))\right]^{|\beta|}, \qquad \beta\in\mathbb{R},\ x\in\mathbb{R},$$

and the test statistic resulting from (2.3) when employing this weight function is denoted by $PW_{n,\gamma}^{2}$. Based on some preliminary Monte Carlo studies, we recommend using $\gamma = 0.1$ as the compromise choice.

Both $PW_{n,\gamma}^{1}$ and $PW_{n,\gamma}^{2}$ reject for large values.

Henze and Meintanis (2005) test ($S_{n,\gamma}$)

The following characterization of the exponential distribution was proven in Meintanis and Iliopoulos (2003): *A random variable X is exponentially distributed if, and only if, its CF satisfies the equation*

$$\begin{aligned} |\phi(t)|^2 &= C^2(t) + S^2(t) \\ &= C(t), \end{aligned}$$

where $|\phi(t)|^2$ is the squared modulus of $\phi(t)$, with $C(t) = \mathrm{E}[\cos(tX)]$ denoting the real part and $S(t) = \mathrm{E}[\sin(tX)]$ denoting the imaginary part of $\phi(t)$.

That is, the squared modulus and the real part of the CF are equal for the exponential distribution. Henze and Meintanis (2005) introduced a test for exponentiality based on the above characterization. The test statistic is given by

$$S_{n,\gamma} = n\int_0^{\infty}\left[|\phi_n(t)|^2 - C_n(t)\right]^2\exp(-\gamma t^2)dt, \tag{2.4}$$

where $\gamma > 0$ is a tuning parameter and $C_n(t)$ is the real part of the ECF, $\phi_n(t)$. The test statistic in (2.4) can be simplified to the closed-form expression

$$\begin{aligned}
S_{n,\gamma} &= \frac{1}{4n}\sqrt{\frac{\pi}{\gamma}}\sum_{j=1}^{n}\sum_{k=1}^{n}\left[\exp\left(-\frac{Y_{jk-}^2}{4\gamma}\right)+\exp\left(-\frac{Y_{jk+}^2}{4\gamma}\right)\right] \\
&- \frac{1}{2n^2}\sqrt{\frac{\pi}{\gamma}}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\left[\exp\left(-\frac{[Y_{jk-}-Y_l]^2}{4\gamma}\right)+\exp\left(-\frac{[Y_{jk-}+Y_l]^2}{4\gamma}\right)\right] \\
&+ \frac{1}{4n^3}\sqrt{\frac{\pi}{\gamma}}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n}\left[\exp\left(-\frac{[Y_{jk-}-Y_{lm-}]^2}{4\gamma}\right)+\exp\left(-\frac{[Y_{jk-}+Y_{lm-}]^2}{4\gamma}\right)\right],
\end{aligned}$$

where $Y_{jk-} = Y_j - Y_k$ and $Y_{jk+} = Y_j + Y_k$ (see Appendix B.1 for the derivation). This test rejects the null hypothesis for large values of $S_{n,\gamma}$. Henze and Meintanis (2005) explains that $S_{n,\gamma}$ is related to the first component, $\hat{U}_{n2}$, of the smooth test for exponentiality (see Baringhaus et al., 2000), given by

$$\hat{U}_{n2} = \frac{\sqrt{n}}{2}\left(\frac{1}{n}\sum_{j=1}^{n}Y_j^2 - 2\right).$$

The mentioned relation is that $\lim_{\gamma\to\infty}\gamma^{5/2}S_{n,\gamma} = \frac{3}{4}\sqrt{\pi}\hat{U}_{n2}^2$. $S_{n,\gamma}$ is not included in the simulation study due to the excessive computer time required in order to calculate critical values. Henze and Meintanis (2005) found that $S_{n,\gamma}$ performs poorly, compared to other tests such as $BH_{n,\gamma}$ (see Baringhaus and Henze, 1991), against alternatives with decreasing hazard rates.

2.2 Tests based on the empirical Laplace transform

In general, the Laplace transform (LT) of a random variable X is defined as $E\left[e^{-tX}\right]$. For a standard exponential random variable, Y, the Laplace transform is given by

$$\psi(t) = E\left[e^{-tY}\right] = \frac{1}{1+t}.$$

Employing the scaled data $Y_1, \ldots, Y_n$, $\psi(t)$ can be estimated by the empirical Laplace transform (ELT),

$$\psi_n(t) = \frac{1}{n}\sum_{j=1}^{n}e^{-tY_j}.$$

We consider two test statistics based on the ELT, namely the 'Baringhaus and Henze (1991)' test and the 'Henze and Meintanis (2002)' test.

Baringhaus and Henze (1991) test ($BH_{n,\gamma}$)

Baringhaus and Henze (1991) developed a test based on the following differential equation that characterises the exponential distribution: $(1+t)\psi'(t) + \psi(t) = 0$, for all $t \in \mathbb{R}$.

Their test makes use of he following weighted $L2$-norm

$$BH_{n,\gamma} = n \int_0^\infty \left[(1+t)\psi_n'(t) + \psi_n(t)\right]^2 \exp(-\gamma t)dt, \tag{2.5}$$

where $\gamma > 0$ is a constant tuning parameter. It is easy to show (see Appendix B.2 for the derivation) that the statistic in (2.5) simplifies to

$$BH_{n,\gamma} = \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\left[\frac{(1-Y_j)(1-Y_k)}{Y_j+Y_k+\gamma} - \frac{Y_j+Y_k}{(Y_j+Y_k+\gamma)^2} + \frac{2Y_jY_k}{(Y_j+Y_k+\gamma)^2} + \frac{2Y_jY_k}{(Y_j+Y_k+\gamma)^3}\right].$$

Baringhaus and Henze (1991) showed that the test statistic has a nondegenerate limiting null distribution and also that the test is consistent against a class of alternative distributions with strictly positive, finite mean. The compromise choice for γ suggested in Baringhaus and Henze (1991) is $\gamma = 1$. This test rejects exponentiality for large values of $BH_{n,\gamma}$.

Henze and Meintanis (2002) test ($L_{n,\gamma}$)

The natural idea of creating a test for exponentiality by measuring the $L2$-distance between the ELT and the LT for the standard exponential distribution was first proposed in Henze (1993). The proposed test statistic has the following form:

$$H_{n,\gamma} = n \int_0^\infty \left(\psi_n(t) - \frac{1}{1+t}\right)^2 \exp(-\gamma t)dt. \tag{2.6}$$

This test statistic should produce a value close to zero if the null hypothesis is true. However, the equation in (2.6) does not simplify to a simple closed-form expression and requires numerical integration. To overcome this issue Henze and Meintanis (2002) proposes the following form of the test statistic:

$$\begin{aligned} L_{n,\gamma} &= n \int_0^\infty \left[(1+t)\psi_n(t) - 1\right]^2 \exp(-\gamma t)dt \\ &= n \int_0^\infty \left[\psi_n(t) - \frac{1}{1+t}\right]^2 (1+t)^2 \exp(-\gamma t)dt, \end{aligned} \tag{2.7}$$

where $\gamma > 0$ (see Appendix B.3 for the derivation). The statistic in (2.7) simplifies to the following closed-form expression:

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

Two possible compromise choices for the parameter γ are suggested for practical applications in Henze and Meintanis (2002); $\gamma = 0.75$ and $\gamma = 1$. For the purpose of this dissertation, $\gamma = 0.75$ is used as the compromise choice. This test rejects H_0 for large values of $L_{n,\gamma}$.

2.3 Tests based on the empirical distribution function

The use of distance measures based on the empirical distribution function (EDF) is one of the earliest approaches to goodness-of-fit testing. The EDF based on the scaled data $Y_1, \ldots, Y_n$ is defined as

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

where $I(\cdot)$ denotes the indicator function and $x \in \mathbb{R}$. The tests considered measure the discrepancy between the standard exponential distribution function and the EDF. The most famous of these include the Kolmogorov-Smirnov and Cramér-von Mises tests (see, for example, D'Agostino and Stephens, 1986), which are discussed below. Another test, based on the integrated EDF, can be found in Klar (2001), but is not discussed here.

Kolmogorov-Smirnov (KS_n)

The Kolmogorov-Smirnov test statistic is given by:

$$KS_n = \sup_{x \geq 0} \left| F_n(x) - \left(1 - e^{-x}\right) \right|. \tag{2.8}$$

The test statistic in (2.8) can be simplified to

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

where

$$\begin{aligned} KS_n^+ &= \max_{1 \leq j \leq n} \left[\frac{j}{n} - \left(1 - e^{-Y_{(j)}}\right)\right], \\ KS_n^- &= \max_{1 \leq j \leq n} \left[\left(1 - e^{-Y_{(j)}}\right) - \frac{j-1}{n}\right]. \end{aligned}$$

This test rejects the null hypothesis for large values of KS_n.

Cramér-von Mises (CM_n)

The Cramér-von Mises test statistic for testing exponentiality is given by

$$CM_n = n \int_0^{\infty} \left[F_n(x) - \left(1 - e^{-x}\right)\right]^2 e^{-x} dx. \tag{2.9}$$

The test statistic in (2.9) can be simplified (see Appendix B.4 for the derivation) to

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

Large values of CM_n will lead to the rejection of the null hypothesis.

2.4 Tests based on mean residual life

In reliability theory and survival analysis the mean residual life (MRL) of a non-negative random variable X at time t, defined as the expected value of the amount of life time remaining after time t, is expressed as

$$m(t) = E\left[X - t | X > t\right] = \frac{\int_t^\infty S(x)dx}{S(t)},$$

where $S(t) = 1 - F(t)$ is the survival function. It was shown in Shanbhag (1970) that the exponential distribution is characterised by a constant MRL, i.e., for the exponential distribution we have that

$$m(t) = E(X) = \frac{1}{\lambda}, \quad \forall t > 0. \tag{2.10}$$

It can be shown that the characterisation in (2.10) is equivalent to

$$E\left(\min\left\{X, t\right\}\right) = \frac{F(t)}{\lambda}, \quad \forall t > 0, \tag{2.11}$$

or

$$\int_t^\infty S(x)dx = \frac{S(t)}{\lambda}, \quad \forall t > 0. \tag{2.12}$$

Tests based on the MRL (and the various forms of the characterising properties given in (2.10) to (2.12)) to test for exponentiality can be found in Baringhaus and Henze (2000), Jammalamadaka and Taufer (2006), and Taufer (2000). A generalisation of the test in Baringhaus and Henze (2000) which includes a more general weight function can be found in Baringhaus and Henze (2008). The two tests considered, namely the Jammalamadaka and Taufer test from Jammalamadaka and Taufer (2006) and the Baringhaus and Henze test from Baringhaus and Henze (2000), employ the characterisations in (2.10) and (2.11), respectively. The test proposed by Taufer in Taufer (2000), however, makes use of the characterisation in (2.12). This test is not considered in this study.

Baringhaus and Henze (2000) ($\overline{KS}_n$ and $\overline{CM}_n$)

In Baringhaus and Henze (2000), the authors introduce Kolmogorov-Smirnov and Cramér-von Mises type tests based on the MRL. The test statistic of the Kolmogorov-Smirnov version of the test is given by

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

where

$$\begin{aligned} \overline{KS}_n^+ &= \max_{j \in \{0,1,\ldots,n-1\}} \left[\frac{1}{n}\left(Y_{(1)} + \ldots + Y_{(j)}\right) + Y_{(j+1)}\left(1 - \frac{j}{n}\right) - \frac{j}{n}\right], \\ \overline{KS}_n^- &= \max_{j \in \{0,1,\ldots,n-1\}} \left[\frac{j}{n} - \frac{1}{n}\left(Y_{(1)} + \ldots + Y_{(j)}\right) - Y_{(j)}\left(1 - \frac{j}{n}\right)\right]. \end{aligned}$$

The Cramér-von Mises type test statistic is:

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

The null hypothesis is rejected for large values of $\overline{KS}_n$ and $\overline{CM}_n$. The asymptotic null distributions of $\overline{KS}_n$ and $\overline{CM}_n$ are identical to the asymptotic null distributions of KS_n and CM_n when used to test for a standard uniform distribution. Baringhaus and Henze (2000) showed that these two tests are consistent against each fixed alternative distribution with positive mean.

Jammalamadaka and Taufer (2006) ($J_{n,\gamma}$)

In Jammalamadaka and Taufer (2006), the authors develop a test based on the characterization in (2.10) by first defining what they call the 'sample MRL after $X_{(k)}$' as follows:

$$\begin{aligned}\bar{X}_{>k} =& \frac{1}{n-k+1}\sum_{j=k+1}^{n+1}\left(X_{(j)} - X_{(k)}\right)\\ =& \frac{1}{n-k+1}\sum_{j=k+1}^{n+1}(n-j+2)\left(X_{(j)} - X_{(j-1)}\right).\end{aligned}$$

Under exponentiality it follows that

$$E\left[\bar{X}_{>k}\right] = E\left[\bar{X}_n\right] = \frac{1}{\lambda}, \quad k = 1,2,\ldots,n. \tag{2.13}$$

Using (2.13), a Kolmogorov-Smirnov type statistic is proposed in Jammalamadaka and Taufer (2006) as a possible test for exponentiality:

$$J_n' = \max_{1\leq k\leq n}\frac{\left|\bar{X}_n - \bar{X}_{>k}\right|}{\bar{X}_n}.$$

Unfortunately, it was shown that this version of the test statistic does not converge to zero even under the null hypothesis of exponentiality. To overcome this problem and some other issues plaguing the statistic J_n', Jammalamadaka and Taufer (2006) constructs a trimmed test statistic whereby some of the last residual means are removed from the calculation. The resulting test statistic has the form

$$J_{n,\gamma} = \max_{1\leq k\leq n-\lfloor n^\gamma\rfloor}\frac{n^{\frac{\gamma}{2}}\left|\bar{X}_n - \bar{X}_{>k}\right|}{\bar{X}_n}, \qquad \gamma\in(0,1), \tag{2.14}$$

where $\lfloor x\rfloor = floor(x)$ and γ is the trimming parameter which indicates how many of the last residual means are discarded. This test rejects the null hypothesis for large values of $J_{n,\gamma}$.

In Jammalamadaka and Taufer (2006), the authors derive the asymptotic null distribution of $J_{n,\gamma}$ and also prove that the test is consistent for every fixed non-exponential alternative distribution with finite mean. In addition, it is shown that the powers of the test are highly sensitive to the choice of γ, but that a compromise choice of $\gamma = 0.9$ (i.e., when a large proportion of the last mean residuals are trimmed) produces the highest powers for the majority of the alternatives considered.

2.5 Tests based on entropy

For a non-negative continuous random variable X with density function $f(x)$, the entropy (sometimes referred to as the *differential entropy*) is given by

$$DE(X) = -\int_0^\infty f(x) \ln f(x) dx. \tag{2.15}$$

Initial attempts (see, for example, Grzegorzewski and Wieczorkowski, 1999 and Ebrahimi et al., 1992) to construct tests for exponentiality based on the entropy exploited the characterisation that, among all distributions with support $[0, \infty)$ and fixed mean, the quantity $DE(X)$ is maximised if X follows an exponential distribution. However, these tests are not explored further, instead we focus on two more recent tests based on the cumulative residual entropy (CRE). The CRE, introduced in Rao et al. (2004), is an alternative information measure which replaces the density function in (2.15) with the survival function, and is defined as

$$CRE(X) = -\int_0^\infty S(x) \ln S(x) dx,$$

where $S(x) = 1 - F(x)$ is the survival function.

Zardasht et al. (2015) (ZP_n)

The first test for exponentiality based on the CRE information measure considered is found in Zardasht et al. (2015). Let X and Z be non-negative random variables with distribution functions F and G, respectively. The test is based on the CRE of the so-called *comparison distribution function*, $D(u) = F(G^{-1}(u))$ (see Parzen, 1998). Calculating the CRE of a random variable with distribution function $D(u)$ and simplifying, the following expression is obtained

$$\mathcal{C}(X, Z) = -\int_0^\infty S(x) \ln S(x) dG(x). \tag{2.16}$$

If W is exponentially distributed with parameter $\lambda > 0$, then (2.16) can be expressed as

$$\mathcal{C}(W, Z) = \int_0^\infty x\lambda e^{-x\lambda} dG(x),$$

which is a measure used to compare the distribution function of Z to that of the exponential distribution. If Z is also exponentially distributed, then it easily follows that $\mathcal{C}(W, Z) = \frac{1}{4}$. The authors of Zardasht et al. (2015) based their test statistic on the difference between an estimator for $\mathcal{C}(W, Z)$ and $\frac{1}{4}$. The resulting test statistic is thus

$$ZP_n = \frac{1}{n}\sum_{j=1}^{n} Y_j e^{-Y_j} - \frac{1}{4}.$$

This test rejects exponentiality for both small and large values of ZP_n. Zardasht et al. (2015) go on to show that $\sqrt{n}ZP_n \overset{\mathcal{D}}{\to} N(0, 5/382)$, but did not formally prove the consistency of the test.

Baratpour and Habibi Rad (2012) (BR_n)

The next test considered is based on the cumulative Kullback-Leibler (CKL) divergence (and indirectly on the CRE) introduced in Baratpour and Habibi Rad (2012). If W_1 and W_2 are two non-negative continuous random variables with distribution functions H and G, respectively, then the CKL divergence between these two distributions is defined as

$$CKL(H, G) = \int_0^\infty (1 - H(x)) \ln \frac{1 - H(x)}{1 - G(x)} dx - \left[E(W_1) - E(W_2)\right].$$

Note that the CKL divergence is somewhat similar to the classical Kullback-Leibler divergence, with the density functions replaced by survival functions.

The authors make use of the fact that, if the null hypothesis is true, then $CKL(F, F_0) = 0$. Rewriting the CKL measure in terms of the CRE measure, and plugging in the necessary estimates, they arrive at the following test statistic

$$BR_n = \frac{\sum_{j=1}^{n-1} \frac{n-j}{n} \left(\ln \frac{n-j}{n}\right) \left(X_{(j+1)} - X_{(j)}\right) + \frac{\sum_{j=1}^n X_j^2}{2\sum_{j=1}^n X_j}}{\frac{\sum_{j=1}^n X_j^2}{2\sum_{j=1}^n X_j}}.$$

The asymptotic distribution under the null hypothesis is not derived in Baratpour and Habibi Rad (2012), however, it is shown that the test is consistent.

This test rejects H_0 for large values of BR_n.

2.6 Tests based on normalized spacings

It has been shown (see, for example, Jammalamadaka and Goria, 2004) that transforming the data can increase the power of tests for exponentiality against certain alternatives. A widely used transformation is to convert the data to the so-called normalized spacings, defined as

$$D_j = (n - j + 1)\left(X_{(j)} - X_{(j-1)}\right), \qquad j = 1, ..., n,$$

with $X_{(0)} = 1$. To find tests for exponentiality that use normalised spacings, the reader is referred to Jammalamadaka and Taufer (2003) and Jammalamadaka and Goria (2004), and for a test where these spacings are used to test for exponentiality in the presence of type-II censoring, see Balakrishnan et al. (2002). We consider two other tests based on spacings; one found in Gail and Gastwirth (1978) and a modification of a test in Gnedenko et al. (1969) which is found in Harris (1976).

Gini test (G_n)

A test statistic that employs normalised spacings for testing exponentiality is described in D'Agostino and Stephens (1986) and is given by:

$$DS_n = \sum_{j=1}^{n-1} U_j = 2n - \frac{2}{n} \sum_{j=1}^{n} jY_{(j)}, \tag{2.17}$$

where

$$U_k = \frac{\sum_{j=1}^{k} D_j}{\sum_{j=1}^{n} X_j}, \qquad \text{for } k = 1, \ldots, n-1,$$

and follows a standard uniform distribution under H_0. This test rejects H_0 for both small and large values of DS_n.

An additional test based on the so-called Gini index, proposed in Gail and Gastwirth (1978), makes use of the following test statistic

$$G_n = \frac{\sum_{j=1}^{n} \sum_{k=1}^{n} |Y_j - Y_k|}{2n(n-1)}. \tag{2.18}$$

It is easy to see that the following relationship holds between the test statistics in (2.17) and (2.18):

$$G_n = 1 - \frac{DS_n}{n-1}.$$

Similar to DS_n, this test rejects the null hypothesis for both small and large values.

Unfortunately, both of these tests have been shown not to be universally consistent.

Harris' modification of Gnedenko's F-test ($HM_{n,r}$)

In Gnedenko et al. (1969), Gnedenko proposed a test for exponentiality which involved ordering a sample of size n and then splitting the n elements into two groups; the r smallest elements and the second containing the remaining $n-r$ elements. The test statistic, given by

$$GD_{n,r} = \frac{\sum_{j=1}^{r} D_j / r}{\sum_{j=r+1}^{n} D_j / (n-r)}, \tag{2.19}$$

follows an F distribution with $2r$ and $2(n-r)$ degrees of freedom under H_0.

A modification of the test in (2.19) was introduced in Harris (1976). This modification can be used to accommodate testing for exponentiality in the presence of hypercensoring and is referred to as *Harris' modification of Gnedenko's F-test.* For this test, the sample spacings are split into three groups: The first group contains the first r spacings, the last group contains the last r last spacings, and the remaining $n-2r$ spacings form the second group. The test is based on the elements in the second group and the test statistic is given by

$$HM_{n,r} = \frac{\left(\sum_{j=1}^{r} D_j + \sum_{j=n-r+1}^{r} D_j\right) / 2r}{\left(\sum_{j=r+1}^{n-r} D_j\right) / (n-2r)}.$$

In Harris (1976), it is recommended that r is chosen to be equal to $n/4$, and this is also the value of r used in the simulation study presented Chapter in 4.

The null hypothesis is rejected for small and large values of both $GD_{n,r}$ and $HM_{n,r}$.

2.7 Tests based on a score function

The score function, defined as the gradient of the log likelihood function, is a powerful tool that can be used to test statistical hypotheses. We consider one test, developed in Cox and Oakes (1984), that employs this score function to test for exponentiality.

Cox and Oakes (1984) (CO_n)

A score test is introduced in Cox and Oakes (1984) that, when applied to censored data, has the following form

$$CO_n = d + \sum_{j=1}^{n} \ln(X_j) - d\frac{\sum_{j=1}^{n} X_j \ln(X_j)}{\sum_{j=1}^{n} X_j},$$

where $d \leq n$ is the number of uncensored data points. However, when $d = n$ (i.e., in the uncensored case) and one uses the scaled data $Y_1, \ldots, Y_n$, the statistic becomes

$$CO_n = n + \sum_{j=1}^{n} (1 - Y_j)\ln(Y_j).$$

The test rejects H_0 for both large and small values of CO_n and it is shown using finite sample simulation studies in both Ascher (1990) and Henze and Meintanis (2005) that the test is quite powerful against a wide variety of non-exponential alternatives.

It follows that $\sqrt{6/n}(CO_n/\pi)$ has a standard normal asymptotic null distribution and is consistent against alternative distributions with $\mathrm{E}(X) < \infty$ and $\mathrm{E}(X \ln X - \ln X) \neq 1$, as discussed in, for example, Henze and Meintanis (2002).

2.8 Tests based on order statistics

Hegazy-Green tests (HG_n^1 and HG_n^2)

In Hegazy and Green (1975), two goodness-of-fit tests for uniformity are presented. These tests can, however, easily be adjusted to test for exponentiality by suitably transforming the data.

If the random variable X has distribution $F(x)$, then the random variable $U = F(X)$ is uniformly distributed over $[0, 1]$. Define

$$U_j = F(X_j),$$

then two test statistics are proposed (respectively based on an L1-norm and an L2-norm)

$$HG_n^1 = \frac{1}{n}\sum_{j=1}^{n} \left|U_{(j)} - \mathrm{E}(U_{(j)})\right|,$$

$$HG_n^2 = \frac{1}{n}\sum_{j=1}^{n} \left[U_{(j)} - \mathrm{E}(U_{(j)})\right]^2,$$

where $U_{(1)} < U_{(2)} < \ldots < U_{(n)}$ denote the order statistics of U_j. When employing the transformation discussed above to test whether the original data are exponentially distributed, the test statistics simplify to

$$HG_n^1 = \frac{1}{n}\sum_{j=1}^{n} \left|U_{(j)} + \ln\left(1 - \frac{j}{n+1}\right)\right|,$$

and

$$HG_n^2 = \frac{1}{n}\sum_{j=1}^{n}\left[U_{(j)} + \ln\left(1 - \frac{j}{n+1}\right)\right]^2.$$

Hegazy and Green (1975) reported that HG_n^1 and HG_n^2 produce very similar powers. The tests reject the null hypothesis for large values of HG_n^1 and HG_n^2. Finite sample results as well as asymptotic results are provided in Hegazy and Green (1975). This test is not included in the simulation study presented in Chapter 4.

2.9 Tests based on other characterizations and properties

Over the years, a multitude of tests for exponentiality have been developed by utilising a number of interesting and varied characterisations and properties of the exponential distribution, but it would not be possible to address all of them in a single study. These tests utilise characterisations such as the Arnold-Villasenor characterisation (see Jovanović et al., 2015), the Rossberg characterisation in Volkova (2010), and various other characterisations (see, for example, Abbasnejad et al., 2012 and Noughabi and Arghami, 2011a). Other tests for exponentiality, not included in this study, include more tests for exponentiality based on order statistics (see Bartholomew, 1957, Hahn and Shapiro, 1967, Jackson, 1967, Shapiro and Wilk, 1972, and Wong and Wong, 1979), tests based on transformations to uniformity (see Seshadri et al., 1969), and tests based on maximum correlations (see Grané and Fortiana, 2011), to name but a few. However, for the purposes of the simulation study conducted in this dissertation, we will consider the following five tests: the Ahsanullah test (Volkova and Nikitin, 2013), a test based on likelihood ratios (Noughabi, 2015), a test based on transformed data (Noughabi and Arghami, 2011b), the Atkinson test (Mimoto and Zitikus, 2008), and a test based on the lack-of-memory property (Ahmad and Alwasel, 1999). The Ahsanullah test is chosen because no finite sample results for this test are available in Volkova and Nikitin (2013), whereas the remaining four are chosen because of their good power performance in finite sample studies found in the literature.

Tests based on Ahsanullah's characterisation (AH_n^1 and AH_n^2)

Assume that the distribution F belongs to a class of distributions $\mathcal{F}$ that are all strictly monotone and whose hazard rate function, $f(x)/S(x)$, is either increasing or decreasing monotonically. Ahsanullah proved the following characterisation of the exponential distribution in Ahsanullah (1978): *Let* $X_1, X_2, ..., X_n$ *be non-negative iid random variables with distribution function* F*. A necessary and sufficient condition for* F *to be exponential is that for some* j *and* k*, the statistics* $(n-j)(X_{(j+1)} - X_{(j)})$ *and* $(n-k)(X_{(k+1)} - X_{(k)})$ *are identically distributed for* $1 \leq j < k < n$*.*

In Volkova and Nikitin (2013), Volkova and Nikitin consider the following specific settings of this characterization: $n = 2$, $j = 0$ and $k = 1$. Under these settings, the characterization takes the following form: *Let* X *and* Y *be non-negative iid random variables from the class* $\mathcal{F}$*.* X *is then exponentially distributed if* $|X - Y|$ *and* $2\min\{X, Y\}$ *are identically distributed.*

The test statistic suggested in Volkova and Nikitin (2013), derived from this characterization, is

$$AH_n^1 = \int_0^\infty \left[H_n(t) - G_n(t)\right] dF_n(t),$$

where

$$H_n(t) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} I\left(|X_j - X_k| < t\right), \qquad t > 0,$$

$$G_n(t) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} I\left(2\min\{X_j, X_k\} < t\right), \qquad t > 0.$$

If the null hypothesis is true, then H_n and G_n should be close to one another. The test therefore rejects H_0 for small or large values of AH_n^1. The authors showed that

$$\sqrt{n}AH_n^1 \overset{\mathcal{D}}{\to} N\left(0, \frac{647}{42525}\right),$$

and calculated local Bahadur efficiencies under common parametric alternatives. However, the finite sample performance of their test statistic was not investigated. In addition, we also consider the more common Cramer-von Mises type distance where the squared difference between H_n and G_n is used; the corresponding statistic is denoted by

$$AH_n^2 = \int_0^\infty \left[H_n(t) - G_n(t)\right]^2 dF_n(t).$$

This new form of the test will reject H_0 for large values of the test statistic.

A test based on likelihood ratios (ZA_n)

Consider the following two generic statistics,

$$Z = \int_{-\infty}^{\infty} Z(t) dw(t) \tag{2.20}$$

and

$$Z_{\max} = \sup_{t\in(-\infty,\infty)} \{Z(t)w(t)\}, \tag{2.21}$$

where $Z(t)$, $dw(t)$ and $w(t)$ are appropriately chosen functions. It is easy to show (see, for example, Zhang, 2002) that if one chooses $Z(t) = X^2(t)$, where

$$X^2(t) = \frac{n[F_n(t) - F_0(t)]^2}{F_0(t)[1 - F_0(t)]}$$

is the Pearson chi-squared statistic, then the statistics in equations (2.20) and (2.21) become the traditional Anderson-Darling, Cramer-von Mises, and Kolmogorov-Smirnov test statistics for specific choices of $dw(t)$ and $w(t)$.

However, Zhang (2002) suggests using the likelihood ratio statistic $G^2(t)$ instead of the $X^2(t)$ statistic, where $G^2(t)$ is defined as

$$G^2(t) = 2n\left\{F_n(t)\log\left(\frac{F_n(t)}{F_0(t)}\right) + [1-F_n(t)]\log\left(\frac{1-F_n(t)}{1-F_0(t)}\right)\right\}.$$

Choosing $Z(t) = G^2(t)$, the authors obtain the following easy-to-calculate versions of the tests statistics for certain choices of $dw(t)$ and $w(t)$:

- Setting $dw(t) = F_n(t)^{-1}\{1-F_n(t)\}^{-1}dF_n(t)$ in (2.20), the following statistic is obtained:

$$ZA_n = -\sum_{j=1}^{n}\left(\frac{\log(1-\exp(-Y_{(j)}))}{n-j+0.5} - \frac{Y_{(j)}}{j-0.5}\right).$$

- Setting $dw(t) = F_0(t)^{-1}\{1-F_0(t)\}^{-1}dF_0(t)$ in (2.20), the following approximate statistic is obtained:

$$ZC_n = \sum_{j=1}^{n}\left(\log\left\{\frac{(1-\exp(-Y_{(j)}))^{-1}-1}{(n-0.5)/(j-0.75)-1}\right\}\right)^2.$$

- Setting $w(t) = 1$ in (2.21), the following statistic is obtained:

$$ZK_n = \max_{1\le j\le n}\left((j-0.5)\log\left\{\frac{j-0.5}{n(1-\exp(-Y_{(j)}))}\right\} + (n-j+0.5)\log\left\{\frac{n-j+0.5}{n(\exp(-Y_{(j)}))}\right\}\right).$$

All of these tests reject H_0 for large values of the test statistics.

The finite sample performance of these three new tests for testing the hypothesis of *normality* are investigated in Zhang (2002), where it is found that the ZA_n and ZC_n versions of these statistics perform well, even when compared to traditionally powerful tests for normality, such as the Shapiro-Wilk test. In Noughabi (2015) the finite sample performance of these tests is investigated when testing for exponentiality. The authors conclude that, among these three tests, ZA_n performs best. As a result only ZA_n is included in the Monte Carlo study. Note that while the finite sample performance of these tests were extensively studied in Noughabi (2015), the derivation of the asymptotic null distribution and consistency of these tests were not discussed.

A test using transformed data (NA_n)

The test proposed in Noughabi and Arghami (2011b) employs the rather simple idea that, for a uniform distribution, the quantity $xf_U(x)$ will be equal to $F_U(x)$, where $x \in [0,1]$, $f_U(\cdot)$ is the uniform density function and $F_U(\cdot)$ is the uniform distribution function. Therefore, given data $V_1, V_2, \ldots, V_n$, a test statistic proposed to test for uniformity is

$$T_n = \frac{1}{n}\sum_{j=1}^{n}\left|V_j\,\widehat{f}(V_j) - F_U(V_j)\right|, \tag{2.22}$$

where $\widehat{f}(\cdot\,)$ is the kernel density estimator defined as

$$\widehat{f}(x) = \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{x - V_j}{h}\right),$$

with $K(\cdot\,)$ the standard normal density function and h the bandwidth chosen using Silverman's normal rule of thumb, $h = 1.06sn^{-1/5}$ (see Silverman, 1986), where s is the sample standard deviation.

The test for exponentiality proceeds by exploiting the following characterisation of exponentiality (see Alzaid and Al-Osh, 1992): *For two independent random observations W_1 and W_2 from a distribution G, the random variable $W_1/(W_1 + W_2)$ is uniformly distributed if, and only if, G is the exponential distribution.*

Subsequently, given the order statistics $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$, construct the transformed data set

$$Z_{ij} = \frac{X_{(i)}}{X_{(i)} + X_{(j)}}, \qquad i \neq j,\ i, j = 1, 2, \ldots, n.$$

Under the hypothesis of exponentiality, these newly transformed values will have a uniform distribution. The test statistic given in (2.22) can consequently be used to test deviations from exponentiality for these transformed data:

$$NA_n = \frac{1}{n(n-1)}\sum\sum_{i \neq j} \left| Z_{ij}\widehat{f}(Z_{ij}) - F_U(Z_{ij}) \right|.$$

The test rejects the null hypothesis for large values of NA_n.

In Noughabi and Arghami (2011b) the authors investigate the finite sample performance of their newly proposed test, but do not derive any asymptotic results.

Another test using transformed data can be found in Dhumal and Shirke (2014), but this test will not be discussed in this dissertation.

The Atkinson test ($AT_{n,\gamma}$)

In Lee et al. (1980) the authors propose tests for exponentiality based on the ratio

$$Q_F(\gamma) = \frac{E[X^\gamma]}{(E[X])^\gamma},$$

for $\gamma > 0$, which is equal to $\Gamma(1+\gamma)$ if X is exponentially distributed.

However, an approach whereby the quantity $Q_F(\gamma)$ is raised to the power $1/\gamma$ to create the following ratio

$$R_F(\gamma) = \frac{E[X^\gamma]^{1/\gamma}}{E[X]},$$

is adopted in Mimoto and Zitikus (2008). Naturally, if X is exponentially distributed, then $R_F(\gamma)$ equals $\Gamma(1+\gamma)^{1/\gamma}$ for $\gamma \neq 0$, and equals $\exp(-\epsilon)$ when $\gamma \to 0$, where $\epsilon = 0.577215...$ is the Euler constant. The test statistic proposed in Mimoto and Zitikus (2008), called the *Atkinson statistic*, is based on the difference between an empirical estimator of $R_F(\gamma)$ and $\Gamma(1+\gamma)^{1/\gamma}$, for γ values between -1 and 1, but $\gamma \neq 0$. The test statistic is given by

$$AT_{n,\gamma} = \sqrt{n}\left| R_n(\gamma) - \Gamma(1+\gamma)^{1/\gamma} \right|, \tag{2.23}$$

where

$$R_n(\gamma) = \frac{1}{\bar{X}_n}\left[\frac{1}{n}\sum_{j=1}^{n} X_j^{\gamma}\right]^{1/\gamma}.$$

In the limit where $\gamma \to 0$ the quantity $R_F(\gamma)$ has the form

$$R_F(0) = \frac{\exp\left(E[\log(X)]\right)}{E[X]},$$

the numerator of which is consistently estimated by the geometric mean $G_n = \prod_{j=1}^{n} X_j^{1/n}$. Therefore, when $\gamma = 0$, the resulting test statistic, called the *Moran statistic for exponentiality*, has the form

$$AT_{n,0} = \sqrt{n}\left|\frac{G_n}{\bar{X}_n} - \exp(-\epsilon)\right|,$$

(see Moran, 1951). For all choices of γ, the test rejects the null hypothesis for large values.

Extensive Monte Carlo power studies are presented in Mimoto and Zitikus (2008) where it is found that values of γ close to 0 and close to 0.99 produce the highest power for most alternatives considered. For the purposes of this dissertation, a compromise choice of $\gamma = 0.01$ is selected. In addition, the authors of Mimoto and Zitikus (2008) establish the asymptotic null distribution and consistency of the test statistic $AT_{n,\gamma}$.

A test based on the lack-of-memory property (AA_n)

It is well-known that the exponential distribution is characterized by the "memoryless property". Let X be a positive random variable with distribution function F. Let t be a positive number such that $F(t) < 1$. The distribution F is said to lack memory at t if

$$P(X > s + t | X > t) = P(X > s).$$

Refer to Shimizu (1979) for more information on the lack-of-memory property.

In Ahmad and Alwasel (1999) the authors utilise the memoryless property to develop a goodness-of-fit test for exponentiality. Let S denote the survival function of X, thus $S(x) = 1 - F(x)$. X is exponentially distributed (with parameter $\lambda > 0$) if, and only if,

$$S(2x) = S^2(x), \qquad \forall\, x \geq 0.$$

The L2-norm based on the difference between $S(2x)$ and $S^2(x)$ is

$$\Delta_2(F) = \int_0^{\infty} \left[S(2x) - S^2(x)\right]^2 dF(x). \tag{2.24}$$

When the EDF is used to estimate F, equation (2.24) is simplified to obtain the test statistic proposed in Ahmad and Alwasel (1999);

$$AA_n = \frac{1}{n}\sum_{j=1}^{n}\left[S_n(2X_{(j)}) - \left(\frac{n-j}{n}\right)^2\right]^2,$$

where $S_n(x) = 1 - F_n(x)$. This test rejects H_0 for large values of AA_n. Ahmad and Alwasel (1999) states that the null limiting distribution of AA_n is not normal. The authors then provide an estimate for AA_n which is asymptotically normal under both the null and alternative hypothesis. The test procedure to obtain power estimates, as well as some finite sample results are also provided in Ahmad and Alwasel (1999). This test is not included in the simulation study in this dissertation due to the large amount of computer time required for the calculation of the critical values when making use of the proposed test procedure.

Chapter 3

The bootstrap

The bootstrap is an automated, computerised resampling technique that, since its introduction by Bradley Efron in the 1970s, has proved to be successful in many problems of inference too complex to address adequately by means of traditional analytical methods.

In fact, apart from being straightforward to implement, in many situations the bootstrap has been shown to produce results that are superior to results obtained by traditional methods in many situations, especially when the sample size is small or when underlying model assumptions cannot be verified (Hall, 1992).

Efron and Tibshirani (1993) states that *"The bootstrap [...] enjoys the advantage of being completely automatic. [It] requires no theoretical calculations, and is available no matter how mathematically complicated the estimator may be."*

A standard introductory text on the bootstrap is Efron and Tibshirani (1993), a more advanced, but still practical text, is Davidson and Hinkley (1997). More formal discussions of the bootstrap, along with important theoretical results and proofs, are given in standard texts such as Hall (1992) and Shao and Tu (1995).

3.1 The bootstrap principle

Let $\boldsymbol{X}_n = \{X_1, X_2, ..., X_n\}$ denote a random sample from an unknown distribution function F. In this chapter we shall give an overview of how the bootstrap can be used to draw inferences from $\boldsymbol{X}_n$ about a parameter vector $\theta = \theta(F)$, for some known functional $\theta(\cdot)$. The *plug-in estimator* for θ is given by $\widehat{\theta}_n = \theta(\widehat{F})$, with $\widehat{F}$ an appropriate estimator for F.

Typically the statistician would draw inferences about θ based on the distributional properties of $\widehat{\theta}_n$, which depend on F. Having only one sample at their disposal it might be a daunting task to uncover these properties safe in instances where assumptions can be made about F or where specific details (such as asymptotics) of $\widehat{\theta}_n$ are known or can be derived.

The main idea behind the bootstrap is to mimic the mechanism F that generated the original sample by resampling from $\widehat{F}$ to obtain what is termed a *bootstrap (re)sample*, which we denote by $\boldsymbol{X}_n^* = \{X_1^*, X_2^*, ..., X_n^*\}$. The process of sampling from F is then imitated by instead sampling from $\widehat{F}$. The idea is that the distribution of $\widehat{\theta}_n^*$ (conditional on $\widehat{F}$) will be "close" to the distribution of $\widehat{\theta}_n$ (conditional on F).

The statistician may draw many such samples to obtain many realisations of $\widehat{\theta}_n^* = \theta(\widehat{F}^*)$, where $\widehat{F}^*$ is a bootstrap approximation for $\widehat{F}$, which can be used to obtain an approximate distribution of $\widehat{\theta}_n^*$. A popular choice, and the only choice we consider, for $\widehat{F}$ is

$$F_n(x) = \frac{1}{n}\sum_{j=1}^{n} I(X_j \leq x),$$

the EDF of $\boldsymbol{X}_n$. Independently sampling from this choice of $\widehat{F}$ leads to the *nonparametric bootstrap*. Sampling from F_n is equivalent to sampling with replacement from $\boldsymbol{X}_n$. In most cases of interest to us we will simply take $\widehat{F} = F_n$, but the results in this chapter are generally applicable to other choices.

The following subsections are devoted to illustrating how the bootstrap is employed in some commonly occurring applications.

3.1.1 Estimation of sampling distributions

Let $R_n(\boldsymbol{X}_n; F)$ denote a random variable of interest, which may depend on both the sample $\boldsymbol{X}_n$ and the unknown distribution function F. The sampling distribution of the random variable $R_n(\boldsymbol{X}_n; F)$ is given by

$$H_n(x) := P\left(R_n(\boldsymbol{X}_n; F) \leq x\right) \qquad \forall x \in \mathbb{R}, \tag{3.1}$$

where P is the probability measure characterised by F. Replacing the distribution function F by an appropriate estimator $\widehat{F}$, we obtain the traditional bootstrap estimator for $H_n(x)$:

$$\widehat{H}_n(x) := P\left(R_n(\boldsymbol{X}_n^*; \widehat{F}) \leq x \mid \boldsymbol{X}_n\right) = P^*\left(R_n(\boldsymbol{X}_n^*; \widehat{F}) \leq x\right) \qquad \forall x \in \mathbb{R}, \tag{3.2}$$

where P^* refers to the conditional probability law of $\boldsymbol{X}_n^*$ given $\boldsymbol{X}_n$. The notation P^* will be used throughout the text to denote this conditional probability measure.

$\widehat{H}_n(x)$ may be approximated by the Monte Carlo algorithm given below.

Approximating $\widehat{H}_n(x)$

1. Generate a sample $X_1^*, X_2^*, ..., X_n^*$ from the EDF, F_n, by sampling with replacement from $X_1, X_2, ..., X_n$.
2. Calculate the statistic $\widehat{\theta}_n^* = \widehat{\theta}(X_1^*, X_2^*, ..., X_n^*)$ for the sample generated in step (1).
3. Independently repeat steps (1) and (2) B times. The statistic calculated in step (2) in the b^{th} iteration will be denoted by $\widehat{\theta}_{n,b}^*$. The result is that we obtain the following bootstrap replications: $\widehat{\theta}_{n,1}^*, \widehat{\theta}_{n,2}^*, ..., \widehat{\theta}_{n,B}^*$.
4. Approximate $\widehat{H}_n(x)$ by
$$\widehat{H}_B(x) := \frac{1}{B}\sum_{b=1}^{B} I\left(\widehat{\theta}_{n,b}^* \leq x\right).$$

3.1.2 Estimation of standard error

A standard error is defined as the standard deviation of the sampling distribution of some statistic. It describes the accuracy with which a population parameter is estimated by an estimator.

We consider the problem of estimating the standard error of the estimator $\widehat{\theta}_n$, denoted by

$$\mathrm{se}\left(\widehat{\theta}_n\right) := \sqrt{\mathrm{Var}\left(\widehat{\theta}_n\right)}. \tag{3.3}$$

The *ideal bootstrap estimate* of $\mathrm{se}\left(\widehat{\theta}_n\right)$ is given by

$$\mathrm{se}^*\left(\widehat{\theta}_n^*\right) = \sqrt{\mathrm{Var}^*\left(\widehat{\theta}_n^*\right)} = \sqrt{\mathrm{E}^*\left(\widehat{\theta}_n^* - \mathrm{E}^*(\widehat{\theta}_n^*)\right)^2}, \tag{3.4}$$

where E^* and Var^* respectively denote the expected value and variance taken with respect to F_n. In most cases (3.4) cannot be calculated explicitly from the sample data. However, in many cases, the ideal bootstrap estimate of the standard error can be effectively approximated by the following algorithm given in Efron and Tibshirani (1986).

Approximating $\mathrm{se}^*\left(\widehat{\theta}_n^*\right)$

1. Generate a sample $X_1^*, X_2^*, ..., X_n^*$ from the EDF, F_n, by sampling with replacement from $X_1, X_2, ..., X_n$.
2. Calculate the statistic $\widehat{\theta}_n^* = \widehat{\theta}(X_1^*, X_2^*, ..., X_n^*)$ for the sample generated in step (1).
3. Independently repeat steps (1) and (2) B times. The statistic calculated in step (2) in the b^{th} iteration will be denoted by $\widehat{\theta}_{n,b}^*$. The result is that we obtain the following bootstrap replications: $\widehat{\theta}_{n,1}^*, \widehat{\theta}_{n,2}^*, ..., \widehat{\theta}_{n,B}^*$.
4. Approximate the ideal bootstrap standard error $\mathrm{se}^*(\widehat{\theta}_n^*)$ by

$$\widehat{\mathrm{se}} := \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\widehat{\theta}_{n,b}^* - \widehat{\theta}_{n,\bullet}^*\right)^2},$$

where

$$\widehat{\theta}_{n,\bullet}^* = \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_{n,b}^*.$$

By the strong law of large numbers it can be shown that $\widehat{\mathrm{se}} \overset{a.s.}{\to} \mathrm{se}^*(\widehat{\theta}_n^*)$ as $B \to \infty$.

In very few cases there exist explicit formulae for the ideal bootstrap estimate of $\mathrm{se}(\widehat{\theta}_n)$. If we choose, for example, the parameter of interest as the population mean, our estimator is the sample mean $\widehat{\theta}_n = \theta(F_n) = \bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$. The bootstrap equivalent is then $\widehat{\theta}_n^* = \theta(\widehat{F}_n) = \bar{X}_n^* = \frac{1}{n}\sum_{j=1}^{n} X_j^*$. In this case (3.3) becomes

$$\mathrm{se}\left(\bar{X}_n\right) := \sqrt{\mathrm{Var}\left(\bar{X}_n\right)} = \frac{\sigma}{\sqrt{n}},$$

where $\sigma = \sqrt{\mathrm{Var}(X)}$. Since we can show that

$$\begin{aligned}
\mathrm{E}^*(\bar{X}_n^*) =& \mathrm{E}^*\left(\frac{1}{n}\sum_{j=1}^{n} X_j^*\right) \\
=& \frac{1}{n}\sum_{j=1}^{n} \mathrm{E}^*(X_j^*) \\
=& \frac{1}{n}\sum_{j=1}^{n} \mathrm{E}^*(X_1^*) \qquad\qquad X_1^*, X_2^*, ..., X_n^* \text{ are i.i.d.} \\
=& \mathrm{E}^*(X_1^*) \\
=& \sum_{j=1}^{n} X_i \frac{1}{n} \\
=& \bar{X}_n,
\end{aligned}$$

and so

$$\widehat{\sigma}_n^2 = \mathrm{Var}^*(X_j^*) = \int (x - \bar{X}_n)^2 dF_n = \frac{1}{n}\sum_{j=1}^{n} (X_j - \bar{X}_n)^2,$$

and the bootstrap estimate in (3.4) becomes

$$\mathrm{se}^*(\bar{X}_n^*) = \sqrt{\mathrm{Var}^*\left(\bar{X}_n^*\right)} = \frac{1}{n}\sqrt{\sum_{j=1}^{n} \mathrm{Var}^*(X_j^*)} = \frac{\widehat{\sigma}_n}{\sqrt{n}}.$$

3.2 Hypothesis testing

In this section some methods that apply the bootstrap to perform hypothesis testing will be discussed.

Two guidelines for performing bootstrap-based hypothesis testing are provided by Hall and Wilson (1991): (1) resample in a way that reflects the null hypothesis, (2) employ methods that are already recognised as having good features in the related problem of confidence interval construction (such as using asymptotically pivotal statistics). For further reading on using the bootstrap for hypothesis testing refer to Chernick (1999), and Good (2000). A method for conducting bootstrap hypothesis testing will now be discussed.

This method involves transforming the sample data $\boldsymbol{X}_n$ so as to mimic the null hypothesis. Let $\boldsymbol{X}_n = \{X_1, X_2, ..., X_n\}$ be a random sample from an unknown distribution F. Further, let the parameter $\theta = \theta(F)$ be some functional of F. Consider the right-sided hypothesis (a left-sided or two-sided test follows the same general reasoning)

$$H_0 : \theta(F) = \theta_0 \qquad H_A : \theta(F) > \theta_0,$$

where θ_0 is a constant.

Let $T_n(\boldsymbol{X}_n)$ be an appropriate test statistic, $C_n(\alpha)$ be the critical value and α be the significance level of the test. Then, this test rejects the null hypothesis if, and only if,

$$T_n(\boldsymbol{X}_n) \geq C_n(\alpha),$$

where $P_{H_0}\left(T_n(\boldsymbol{X}_n) \geq C_n(\alpha)\right) \approx \alpha$.

Since F is unknown, the critical value $C_n(\alpha)$ is also unknown. We can estimate $C_n(\alpha)$ by the bootstrap estimator $C_n(\alpha, \boldsymbol{X}_n)$. When applying the bootstrap one finds the bootstrap sample $\boldsymbol{X}_n^* = \{X_1^*, X_2^*, ..., X_n^*\}$ by sampling with replacement from F_n (the EDF of $\boldsymbol{X}_n$). However, to mimic H_0 we need $\theta(F_n) = \theta_0$, but this is hardly ever the case. Hence, the original data, $\boldsymbol{X}_n$, need to be transformed.

Denote the transformed variables by

$$V_j^0 = V_j(\boldsymbol{X}_n; \theta_0), \qquad j = 1, 2, ..., n.$$

Now, the bootstrap sample is given by $\boldsymbol{V}_n^{0*} = \{V_1^{0*}, V_2^{0*}, ..., V_n^{0*}\}$, which is obtained by sampling with replacement from G_n, the EDF of $\boldsymbol{V}_n^0 = \{V_1^0, V_2^0, ..., V_n^0\}$. Choose $V_j(\boldsymbol{X}_n, \theta_0)$, $j = 1, 2, ..., n$, so that $\theta(G_n) = \theta_0$. Finally, the bootstrap estimator $C_n(\alpha, \boldsymbol{X}_n)$ is chosen in such a way that it satisfies the following expression:

$$P^*\left(T_n(\boldsymbol{V}_n^{0*}) \geq C_n(\alpha, \boldsymbol{X}_n)\right) \approx \alpha.$$

The following algorithm can be used to find $\widehat{C}_n(\alpha, \boldsymbol{X}_n)$, which is the Monte Carlo approximation of $C_n(\alpha, \boldsymbol{X}_n)$.

Approximating $C_n(\alpha, \boldsymbol{X}_n)$

1. Given sample data $\boldsymbol{X}_n = \{X_1, X_2, ..., X_n\}$ from F.
2. Find the transformation $V_j^0 = V_j(\boldsymbol{X}_n, \theta_0)$ for $j = 1, 2, ..., n$, such that the EDF of this new data G_n has the property $\theta(G_n) = \theta_0$.
3. Obtain the bootstrap sample $\boldsymbol{V}_n^{0*} = \{V_1^{0*}, V_2^{0*}, ..., V_n^{0*}\}$ by sampling with replacement from $\boldsymbol{V}_n^0 = \{V_1^0, V_2^0, ..., V_n^0\}$. Calculate $T_n(\boldsymbol{V}_n^{0*})$ and denote the result by T_1^*.
4. Independently repeat step (3) B times to obtain the bootstrap replications $T_1^*, T_2^*, ..., T_B^*$.
5. Obtain the order statistics $T_{(1)}^* \leq T_{(2)}^* \leq ... \leq T_{(B)}^*$.
6. Approximate the critical value $C_n(\alpha, \boldsymbol{X}_n)$ by

   $$\widehat{C}_n(\alpha, \boldsymbol{X}_n) = T^*_{(\lfloor B(1-\alpha) \rfloor)},$$

   where $\lfloor x \rfloor$ is the largest integer smaller than x.

The bootstrap p-value is given by

$$p_{boot} = P^*\left(T_n(\boldsymbol{V}_n^{0*}) \geq T_n(\boldsymbol{X}_n)\right),$$

which can be approximated by the following algorithm:

Approximating p_{boot}

1. Given sample data $\boldsymbol{X}_n = \{X_1, X_2, ..., X_n\}$ from F.
2. Find the transformation $V_j^0 = V_j(\boldsymbol{X}_n, \theta_0)$ for $j = 1, 2, ..., n$, such that the EDF of this new data, G_n, has the property $\theta(G_n) = \theta_0$.
3. Obtain the bootstrap sample $\boldsymbol{V}_n^{0*} = \{V_1^{0*}, V_2^{0*}, ..., V_n^{0*}\}$ by sampling with replacement from $\boldsymbol{V}_n^0 = \{V_1^0, V_2^0, ..., V_n^0\}$. Calculate $T_n(\boldsymbol{V}_n^{0*})$ and denote the result by T_1^*.
4. Independently repeat step (3) B times to obtain the bootstrap replications $T_1^*, T_2^*, ..., T_B^*$.
5. Approximate the p-value with

$$\widehat{p}_{boot} = \frac{1}{B} \sum_{b=1}^{B} I\,(T_b^* \geq T_n(\boldsymbol{X}_n)).$$

The bootstrap power of the test at a given alternative θ_A, where $\theta_A = \{\theta : \theta > \theta_0\}$, is given by

$$P_{boot} = P^* \left(T_n(\boldsymbol{V}_n^{A*}) \geq C_n(\alpha, \boldsymbol{X}_n)\right),$$

which can be approximated by the following algorithm:

Approximating P_{boot}

1. Given sample data $\boldsymbol{X}_n = \{X_1, X_2, ..., X_n\}$ from F.
2. Approximate the critical value $C_n(\alpha, \boldsymbol{X}_n)$ by $\widehat{C}_n(\alpha, \boldsymbol{X}_n)$ as discussed in the algorithm for approximating $C_n(\alpha, \boldsymbol{X}_n)$.
3. Find the transformation $V_j^A = V_j(\boldsymbol{X}_n, \theta_A)$ for $j = 1, 2, ..., n$, such that the EDF of this new data, H_n, has the property $\theta(H_n) = \theta_A$.
4. Obtain the bootstrap sample $\boldsymbol{V}_n^{A*} = \{V_1^{A*}, V_2^{A*}, ..., V_n^{A*}\}$ by sampling with replacement from $\boldsymbol{V}_n^A = \{V_1^A, V_2^A, ..., V_n^A\}$. Calculate $T_n(\boldsymbol{V}_n^{A*})$ and denote the result by T_1^{A*}.
5. Independently repeat step (3) B times to obtain the bootstrap replications $T_1^{A*}, T_2^{A*}, ..., T_B^{A*}$.
6. Approximate the power of the test at an alternative θ_A with

$$\widehat{P}_{boot} = \frac{1}{B} \sum_{b=1}^{B} I\left(T_b^{A*} \geq \widehat{C}_n(\alpha, \boldsymbol{X}_n)\right).$$

For some examples on different hypothesis tests that incorporate the transformation method, see Efron and Tibshirani (1993), Westfall and Young (1993), Davidson and Hinkley (1997), Martin (2007), Fisher and Hall (1990), and Boos and Brownie (1989).

In this chapter we have concentrated specifically on bootstrap hypothesis testing because a main focus of this dissertation is to consider the perfomance of tests that employ a data dependent tuning parameter chosen by maximising the bootstrap power. This procedure will be explained in more detail in Chapter 4.

Chapter 4

Monte Carlo simulations

In this chapter Monte Carlo simulations are used to evaluate the power of various tests discussed in Chapter 2.

Note that the following tests are excluded from the Monte Carlo study: $S_{n,\gamma}$, HG_n^1, HG_n^2 and AA_n, since they either have been shown not to be very powerful or they have strenuous computational requirements. The other 20 tests that are included in the simulation study are based on a wide variety of characterisations for the exponential distribution. These tests were chosen since they provide a diversity in terms of comparing established tests to newly developed tests, and also in terms of comparing tests that contain a tuning parameter to those that do not contain tuning parameters. Most of these tests were also reported in the literature to have good power performances in finite sample studies. We will, however, first discuss the methodology used to choose the value of the tuning parameter data-dependently (for those tests containing a tuning parameter).

4.1 A data-dependent choice of the tuning parameter

Many of the tests discussed in Chapter 2 contain a tuning parameter γ, typically appearing in a weight function (see for example the test statistics in Allison and Santana (2015), Alzaid and Al-Osh (1992), and Baringhaus and Henze (2000)). As stated earlier, authors typically approach the selection of this parameter by evaluating the power performance of their tests across a grid of values of the tuning parameter and then suggesting a compromise choice for the parameter by selecting a value that fares well for the majority of the alternatives considered. However, there is general agreement that a data-dependent choice of the tuning parameter is required for practical implementation.

Consider a generic test statistic which contains a tuning parameter γ denoted $T_{n,\gamma}$, whose critical values, denoted by $\widetilde{C}_{n,\gamma}(\alpha)$, can be obtained through Monte Carlo simulation. A possible data-dependent choice of the parameter γ proposed by Allison and Santana (2015) can be obtained by maximising the bootstrap power of the test as follows:

$$\widehat{\gamma} = \widehat{\gamma}(\boldsymbol{X}_n) = \arg\sup_{\gamma\in\mathbb{R}} P^* \left(T_{n,\gamma}(\boldsymbol{Y}_n^*) \geq \widetilde{C}_{n,\gamma}(\alpha)\right),$$

where $\boldsymbol{Y}_n^* = \{Y_1^*, Y_2^*, ..., Y_n^*\}$ denotes a bootstrap sample taken with replacement from $\boldsymbol{Y}_n$, and P^* is the law of $\boldsymbol{Y}_n^*$ given $\boldsymbol{Y}_n$. In Allison and Santana (2015) the following algorithm used to approximate the ideal bootstrap estimator $\widehat{\gamma}$ is provided.

Approximating $\widehat{\gamma}$

1. Fix a grid of γ values: $\gamma \in \{\gamma_1, \gamma_2, ..., \gamma_k\}$.
2. Generate a bootstrap sample $\boldsymbol{Y}_n^* = \{Y_1^*, Y_2^*, ..., Y_n^*\}$ by sampling with replacement from $\boldsymbol{Y}_n = \{Y_1, Y_2, ..., Y_n\}$.
3. Calculate the test statistic $T_{n,\gamma}(\boldsymbol{Y}_n^*)$ for the sample generated in step (2), $j = 1, 2, ..., k$.
4. Repeat steps (2) and (3) a large number of times (say B times). Denote the resulting test statistics by $T^*_{n,\gamma_j,1}, T^*_{n,\gamma_j,2}, ..., T^*_{n,\gamma_j,B}$, $j = 1, 2, ..., k$.
5. Calculate
$$\widehat{P}_{boot,\gamma_j} = \frac{1}{B}\sum_{b=1}^{B} I\left(T^*_{n,\gamma_j,b} \geq \widetilde{C}_{n,\gamma_j}(\alpha)\right) \qquad j = 1, 2, ..., k.$$
6. Calculate
$$\widehat{\gamma}_B = \widehat{\gamma}_B(\boldsymbol{X}_n) = \arg \max_{\gamma \in \{\gamma_1, \gamma_2, ..., \gamma_k\}} \widehat{P}_{boot,\gamma}.$$

The numerical results reported in Tables A.1 – A.6 in Chapter 4 relating to test statistics containing a tuning parameter are obtained using the estimated tuning parameter as described above. The estimated powers obtained using the compromise choice of γ are reported in parentheses in these tables. The details related to the choice of the grid used for each test are discussed in the next section.

4.2 Simulation setting

Throughout the simulation study a significance level of 5% is used and the critical values of all tests are calculated based on 10 000 independent Monte Carlo replications. All calculations are done in R (R Core Team, 2013).

Power estimates are calculated for sample sizes $n \in \{10, 20, 30, 50, 75, 100\}$ using 5 000 independent Monte Carlo replications for various alternative distributions. These alternative distributions, given in Table 4.1, are chosen since they are commonly employed alternatives to the exponential distribution, which has a constant hazard rate (CHR). The distributions considered include those with increasing hazard rates (IHR), decreasing hazard rates (DHR), as well as non-monotone hazard rates (NMHR).

In order to determine the power of the six tests containing a tuning parameter ($BH_{n,\gamma}$, $L_{n,\gamma}$, $PW^1_{n,\gamma}$, $PW^2_{n,\gamma}$, $J_{n,\gamma}$, $AT_{n,\gamma}$) when using the data-dependent choice of the parameter (discussed in Section 4.1), we first need to approximate the empirical powers of these tests for each value of γ in a sequence of γ values. The empirical power based on the data-dependent choice is then calculated as described in Allison and Santana (2015). In each case $B = 250$ bootstrap replications are used to evaluate the bootstrap power of the tests. The following grids of values of the parameter are used for the respective tests:

- For $BH_{n,\gamma}$, $L_{n,\gamma}$, $PW^1_{n,\gamma}$, and $PW^2_{n,\gamma}$ the grid of γ values is given by
$$\gamma \in \{0.1, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 5\}.$$

- For $J_{n,\gamma}$, the grid of γ values is

$$\gamma \in \{0.1, 0.3, 0.5, 0.7, 0.9\}.$$

- The grid of γ values used for $AT_{n,\gamma}$ is

$$\gamma \in \{-0.99, -0.75, -0.5, -0.25, -0.01, 0.01, 0.25, 0.5, 0.75, 0.99\}.$$

4.3 Simulation results

Tables A.1 – A.6 show the estimated powers of the various tests discussed in Chapter 2 for sample sizes $n \in \{10, 20, 30, 50, 75, 100\}$ against each of the alternative distributions given in Table 4.1. The entries in these tables are the percentage of 5 000 independent Monte Carlo samples that resulted in the rejection of H_0 rounded to the nearest integer. Note that, for the tests containing a tuning parameter, the primary entry is the approximate power for the test based on the *data-dependent choice* of the parameter, $\widehat{\gamma}$, while the approximate power of the test based on the *compromise choice* appears in parentheses along-side it. To ease comparisons between the results, the highest power for each alternative distribution is highlighted.

The primary aim of this dissertation is to compare the power of these tests against a wide range of alternative distributions. Some general conclusions relating to the reported estimated powers of the various tests are presented below. For the second part of the analysis of the results, only the tests containing tuning parameters are considered. The powers achieved by tests employing the data-dependent choice proposed in Allison and Santana (2015) are compared here with those associated with the compromise choice of the parameter.

The performance of the tests are greatly affected by the shape of the hazard rate of the alternative distribution considered. Consequently, the overall results, as well as the results categorised according to the shape of the hazard rate classified as increasing, decreasing, or non-monotone are discussed.

4.4 Power comparisons

For the purposes of the comparison between the power of the various tests the data-dependent choice (and not the compromise choice) of the tuning parameter for the tests containing such a parameter is used.

Consider the performance of the tests in general against all alternatives. The powers of HM_n do not compare favourably to those of the other tests; this test reveals lower powers against the majority of the alternatives. For small samples, AH_n^2, BR_n and NA_n also exhibit lower powers against the majority of the alternatives. The tests that generally perform well are CO_n, ZA_n, $AT_{n,\widehat{\gamma}}$, $BH_{n,\widehat{\gamma}}$ and $L_{n,\widehat{\gamma}}$. The CM_n and $\overline{CM}_n$ also perform relatively well against the majority of the alternatives, especially for large samples.

Now consider the results pertaining to the alternatives with increasing hazard rates. Against these alternatives HM_n, KS_n, AH_n^1, $J_{n,\widehat{\gamma}}$, $PW_{n,\widehat{\gamma}}^1$ and $PW_{n,\widehat{\gamma}}^2$ exhibit lower powers for all sample sizes considered. BR_n has higher power in the case of small sample sizes,

Alternative	$f(x)$	Notation
Gamma	$\frac{1}{\Gamma(\theta)}x^{\theta-1}e^{-x}$	$\Gamma(\theta)$
Weibull	$\theta x^{\theta-1}\exp(-x^{\theta})$	$W(\theta)$
Power	$\frac{1}{\theta}x^{(1-\theta)/\theta}$, $0 < x < 1$	$PW(\theta)$
Lognormal	$\exp\left\{-\frac{1}{2}\left(\log(x)/\theta\right)^2\right\} / \left\{\theta x\sqrt{2\pi}\right\}$	$LN(\theta)$
Dhillon	$\frac{\theta+1}{x+1}\exp\left\{-\left(\log(x+1)\right)^{\theta+1}\right\}\left(\log(x+1)\right)^{\theta}$	$DH(\theta)$
Chen	$2\theta x^{\theta-1}\exp\left\{x^{\theta}+2\left(1-\exp\left(x^{\theta}\right)\right)\right\}$	$CH(\theta)$
Linear failure rate	$(1+\theta x)\exp(-x-\theta x^2/2)$	$LF(\theta)$
Extreme value	$\frac{1}{\theta}\exp\left(x+\frac{1-e^{x}}{\theta}\right)$	$EV(\theta)$
Half normal	$\left(\frac{2}{\pi}\right)^2\exp\left(\frac{-x^2}{2}\right)$	HN
Beta	$\frac{\Gamma(\theta_1+\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)}x^{\theta_1-1}(1-x)^{\theta_2-1}$	$B(\theta_1,\theta_2)$
Exponential power	$\exp\left\{1-\exp\left(x^{\theta}\right)\right\}\exp\left\{x^{\theta}\right\}\theta x^{\theta-1}$	$EP(\theta)$
Exponential logarithmic	$\frac{1}{-\ln\theta}\frac{(1-\theta)e^{-x}}{1-(1-\theta)e^{-x}}$	$EL(\theta)$
Exponential Nadarajah Haghighi (1)*	$\frac{\theta(1+x)^{-0.5}e^{1-(1+x)^{0.5}}}{2\left[1-e^{1-(1+x)^{0.5}}\right]^{1-\theta}}$	$ENH1(\theta)$
Exponential Nadarajah Haghighi (2) *	$\frac{2\theta(1+x)e^{-x^2-2x}}{\left[1-e^{-x^2-2x}\right]^{1-\theta}}$	$ENH2(\theta)$
Beta exponential	$\theta e^{-x}(1-e^{-x})^{\theta-1}$	$BEX(\theta)$
Exponential geometric	$\frac{(1-\theta)e^{-x}}{(1-\theta e^{-x})^2}$	$EG(\theta)$

Table 4.1: Various choices of the alternative distributions. * see Lemonte (2013).

but its power relative to the other tests decreases with sample size. The opposite is true for $L_{n,\widehat{\gamma}}$, which reveals a relative increase in power with sample size. The two tests based on mean residual life, $\overline{KS}_n$ and $\overline{CM}_n$, perform relatively well for all sample sizes. The Cramér-von Mises type statistic for Ahsanullah's test, AH_n^2, and NA_n also perform well, especially for small sample sizes. The following tests exhibit high powers in the case of large sample sizes: G_n, EP_n, ZA_n and $BH_{n,\widehat{\gamma}}$.

Now turn your attention to the alternatives with decreasing hazard rates. HM_n, AH_n^2, BR_n and NA_n perform poorly for all sample sizes. In turn, the tests for which large powers are observed are CO_n, $BH_{n,\widehat{\gamma}}$ and $L_{n,\widehat{\gamma}}$. Furthermore, $\overline{CM}_n$, G_n, EP_n and $AT_{n,\widehat{\gamma}}$ perform well, especially for large samples, while $PW_{n,\widehat{\gamma}}^2$ provides higher relative powers in the case of small samples.

The results pertaining to the alternatives with non-monotone hazard rates are as follows. The tests generally demonstrating the lowest powers are HM_n, BR_n and NA_n. For small sample sizes AH_n^2 performs poorly, while G_n, and EP_n exhibit relatively low powers in the case of large samples. However, ZA_n, $AT_{n,\widehat{\gamma}}$, $BH_{n,\widehat{\gamma}}$ and $L_{n,\widehat{\gamma}}$ generally perform well for all sample sizes. The original probability weighted characteristic function test, $PW_{n,\widehat{\gamma}}^1$, where the weights emphasise the centre of the distribution, does well in the case of larger samples. On the other hand, the alternative formulation of this test with the weight function allocating the majority of the weight to the tails of the distribution, $PW_{n,\widehat{\gamma}}^2$, exhibits relatively high power, especially for small samples. The same is true for CO_n.

In summary, the powers achieved by HM_n are generally substantially lower than those of the remaining tests. Other tests that do not generally achieve good results are AH_n^2, BR_n, and ZA_n. The tests that perform well are $BH_{n,\widehat{\gamma}}$, $L_{n,\widehat{\gamma}}$, $AT_{n,\widehat{\gamma}}$ and CO_n. The test that performs the best overall is $BH_{n,\widehat{\gamma}}$, closely followed by $L_{n,\widehat{\gamma}}$. Note that only one of the tests reported to perform relatively poorly contain a tuning parameter, while only one of the tests reported to achieve high powers do not contain such a parameter; CO_n performs the best among those tests that do not include a tuning parameter.

4.5 Comparisons based on the choice of the tuning parameter

Six of the goodness-of-fit test statistics considered contain tuning parameters. Below, the powers achieved by these tests using two different values of the tuning parameter is compared. The first value is chosen data-dependently using the method detailed in Allison and Santana (2015), while the second is the compromise choice recommended in the relevant literature. As was the case above, the discussion below does not only refer to the overall performance of the tests; the performance of the tests against alternatives with increasing, decreasing and non-monotone hazard rates are also discussed separately.

Consider the overall results first. For smaller sample sizes there is little to choose between the powers obtained using $AT_{n,\gamma}$ based on the choices of the tuning parameter. However, as the sample size increases, use of the data-dependent choice generally results in a slight increase in relative power. On the other hand, when using $J_{n,\gamma}$ the choice between the tuning parameters is unimportant for large samples, but for smaller samples the data-dependent choice leads to slightly higher powers. For both $BH_{n,\gamma}$ and $L_{n,\gamma}$ the data-dependent choice leads to higher powers than the compromise choice. Interestingly, the compromise choice outperforms the data-dependent choice in the case of the original PWECF test, $PW^1_{n,\gamma}$, by a small margin, while the data-dependent choice leads to vast improvements in the powers associated with $PW^2_{n,\gamma}$ (giving more weight towards the tails of the distribution), especially for larger samples.

Next, consider alternative distributions with increasing hazard rates. In this case the use of either method for the choice of the tuning parameter leads to little difference in powers obtained using the $AT_{n,\gamma}$, $BH_{n,\gamma}$, $PW^1_{n,\gamma}$ and $L_{n,\gamma}$ tests. The performance of $J_{n,\gamma}$ is slightly improved by using the compromise choice, while the performance of $PW^2_{n,\gamma}$ is greatly improved when using the data-dependent choice of the tuning parameter.

Turning your attention to the alternative distributions with decreasing hazard rates, we see that the observed powers are not substantially affected by the choice of tuning parameter in the case of the following tests: $AT_{n,\gamma}$, $BH_{n,\gamma}$ and $L_{n,\gamma}$. For both $PW^1_{n,\gamma}$ and $PW^2_{n,\gamma}$ the compromise choice of the tuning parameter outperforms the data-dependent choice. The power of $J_{n,\gamma}$ is substantially improved when using the data-dependent choice, especially for small samples.

Finally, consider the performance of the tests against alternatives with non-monotone hazard rates. When using $PW^1_{n,\gamma}$ the powers can be increased by using the compromise choice, especially for small samples. However, substantial improvements in the power of $PW^2_{n,\gamma}$ are realised when the data-dependent choice is used, especially in the case of larger samples. The powers of $BH_{n,\gamma}$ and $L_{n,\gamma}$ are higher when the data-dependent choice is used than is the case for the compromise choice. The performance of $AT_{n,\gamma}$ is not substantially affected by the choice of the tuning parameter for small samples, but using the data-dependent choice leads to improved power in the case of larger samples. When using $J_{n,\gamma}$ the data-dependent choice outperforms the compromise choice for small samples.

It is interesting to note that in the cases where the compromise choice of the tuning parameter outperforms the data-dependent choice the difference in realised power is usually small. However, there are cases where the power associated with the data-dependent choice vastly outperforms the compromise choice. As an example, consider the power of $PW^2_{n,\gamma}$ against samples of size 75 generated from a lognormal distribution with parameter 0.8. The power using the compromise choice is estimated to be 0%, while the estimated power associated with the data-dependent choice is estimated to be 96%. Various other instances of this phenomenon can be observed in the reported powers.

To conclude this section, a short illustration of how the choice of the tuning parameter affects the power of two of the tests considered in the study is provided. For this purpose, consider the tests $L_{n,\gamma}$ and $J_{n,\gamma}$ for sample size $n = 20$. In order to more easily visualise the behaviour of the powers across the γ values Figures 4.1 and 4.2 present the powers obtained for tests $L_{n,\gamma}$ and $J_{n,\gamma}$, respectively, for each choice of γ in the grid of selected γ values. The powers are calculated for five different alternative distributions. For each test, the compromise choice of the tuning parameter is indicated by a vertical dashed line in the relevant figure.

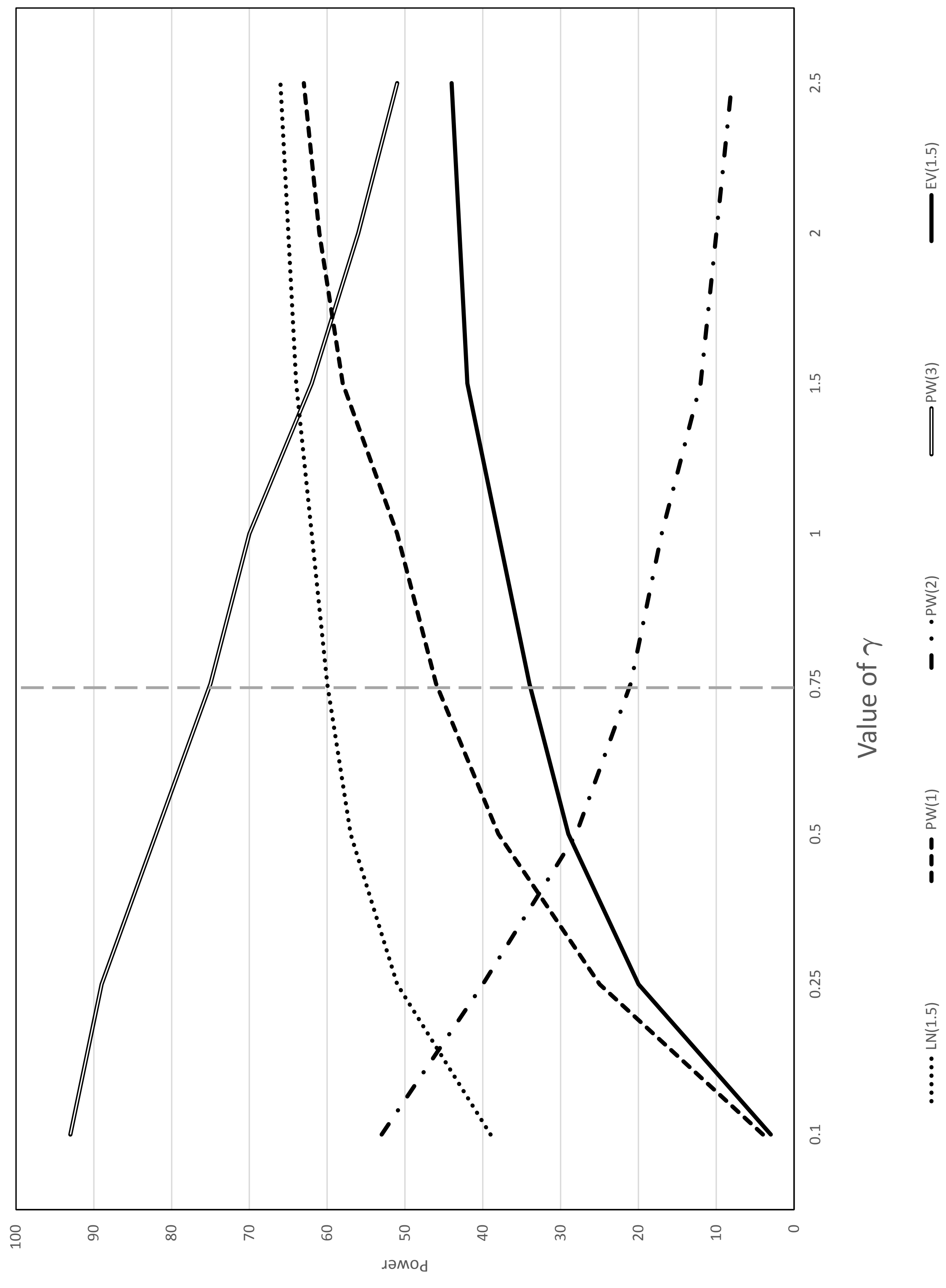

Figure 4.1: Powers for $L_{n,\gamma}$ for $n = 20$ for various alternatives.

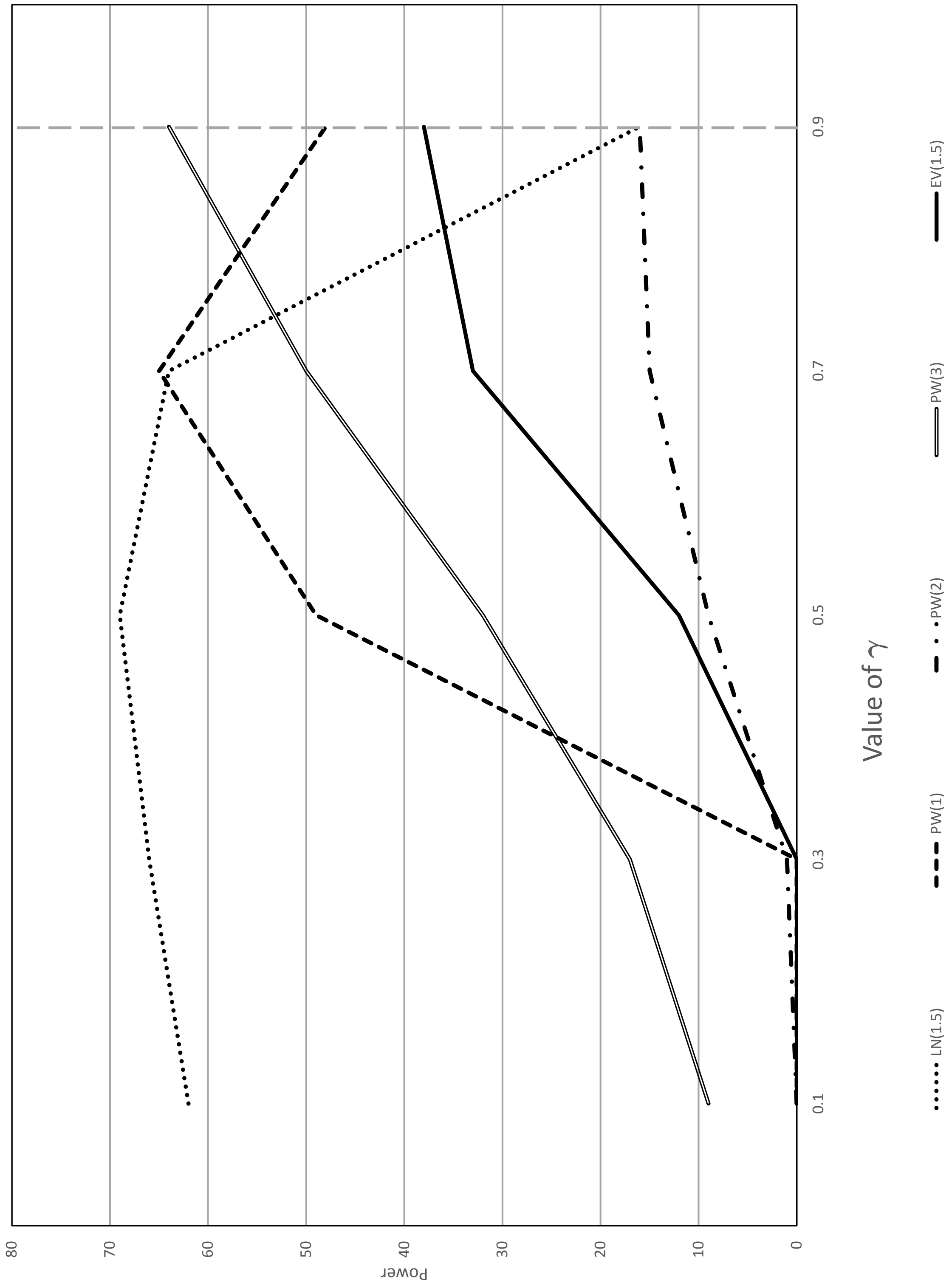

Figure 4.2: Powers for $J_{n,\gamma}$ for $n = 20$ for various alternatives.

It is clear from the figures that the power of the tests is highly dependent on the choice of γ. The compromise choice performs moderately well in many of the alternatives, but in some cases it produces low powers relative to other choices of γ (see, e.g., $L_{n,\gamma}$ for alternative PW(2) and $J_{n,\gamma}$ for alternatives LN(1.5) and PW(1)). Furthermore, the main entries in Tables 4.2 and 4.3 correspond to the powers presented in the figures, whereas the values stated in parentheses in these tables denote the percentage of times (out of 5 000 independent Monte Carlo simulations) that the data-dependent procedure selected the γ value that corresponds to the γ value given in the column heading. These tables are provided to show that the procedure for obtaining the data-dependent choice of the tuning parameter *most frequently* selects the value of γ that produces the *highest power* for a given alternative. Consider, for example, $L_{n,\gamma}$ for the alternative PW(2), where the maximum power of 53% is obtained at $\gamma = 0.1$. The percentage of times that the procedure chose $\gamma = 0.1$ is 68%, and the power of the test based on the data-dependent choice is 43%. In contrast, the power associated with the compromise choice is only 21%.

γ	**0.1**	**0.25**	**0.5**	**0.75**	**1**	**1.5**	**2**	**2.5**	**5**	$\widehat{\gamma}$
$LN(1.5)$	39(4)	51(5)	57(7)	60(6)	62(9)	64(9)	65(12)	66(15)	67(33)	63
$PW(1)$	4(3)	25(5)	38(3)	46(3)	51(2)	58(5)	61(10)	63(15)	66(54)	59
$PW(2)$	53(68)	40(10)	28(2)	21(3)	17(1)	12(2)	10(3)	8(3)	5(8)	43
$PW(3)$	93(83)	89(11)	82(3)	75(2)	70(1)	62(0)	56(0)	51(0)	37(0)	91
$EV(1.5)$	3(5)	20(6)	29(5)	34(4)	38(6)	42(10)	43(13)	44(17)	45(34)	39

Table 4.2: Percentage of 5 000 samples that resulted in the rejection of H_0 (main entries) and the percentage of times that the procedure selected the specific value of γ (in parentheses) based on test $L_{n,\gamma}$ for $n = 20$.

γ	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	$\widehat{\gamma}$
$LN(1.5)$	62(18)	66(17)	69(33)	64(29)	16(3)	64
$PW(1)$	0(0)	0(0)	49(23)	65(50)	48(27)	57
$PW(2)$	0(0)	1(0)	9(24)	15(48)	16(28)	14
$PW(3)$	9(0)	17(1)	32(10)	50(39)	64(50)	55
$EV(1.5)$	0(0)	0(0)	12(8)	33(43)	38(49)	34

Table 4.3: Percentage of 5 000 samples that resulted in the rejection of H_0 (main entries) and the percentage of times that the procedure selected the specific value of γ (in parentheses) based on test $J_{n,\gamma}$ for $n = 20$.

4.6 Equal distance comparisons

In this section we depart from the choices of the parameters of the alternative distributions used for the alternatives in Table 4.1, in order to compare the powers of the tests. Kullback and Leibler (1951) defines a measure of divergence between two distributions, known as the symmetrical Kullback-Leibler distance.

4.6.1 Kullback-Leibler (KL) distance

Let f and g be the probability density functions (pdf) of some distributions, F and G respectively. The KL distance between F and G is given by

$$K(f,g) = \mathcal{K}(f,g) + \mathcal{K}(g,f),$$

where $\mathcal{K}(f,g)$ is defined as the mean information for discrimination between F and G, given by

$$\mathcal{K}(f,g) = \int \ln\left(\frac{f(x)}{g(x)}\right) f(x)dx.$$

In general one would have that $\mathcal{K}(f,g) \neq \mathcal{K}(g,f)$ and that the KL distance can be written as follows:

$$\begin{aligned}
K(f,g) &= \mathcal{K}(f,g) + \mathcal{K}(g,f) \\
&= \int \ln\left(\frac{f(x)}{g(x)}\right) f(x)dx + \int \ln\left(\frac{g(x)}{f(x)}\right) g(x)dx \\
&= \int \ln\left(\frac{f(x)}{g(x)}\right) f(x)dx + \int -\ln\left(\frac{f(x)}{g(x)}\right) g(x)dx \\
&= \int [f(x) - g(x)] \ln\left(\frac{f(x)}{g(x)}\right) dx.
\end{aligned}$$

4.6.2 Results

The KL distances between the alternatives used in the simulation study and the standard exponential distribution varies. In order to eliminate this variation one would choose values for the parameters of the alternatives which would result in equal KL distances from the exponential distribution. Mimoto and Zitikus (2008) provides values of θ for comparing various alternatives to the standard exponential distribution, with KL distances $K(f,g) = 0.2$ and $K(f,g) = 0.5$.

We made use of sixteen of the alternatives identified in Mimoto and Zitikus (2008), eight for which $K(f,g) = 0.2$ and eight for which $K(f,g) = 0.5$, for comparing the Monte Carlo power estimates against these equally distanced alternatives for some of the tests discussed in this dissertation. The specific alternatives were chosen to include alternatives with increasing, decreasing and non-monotone hazard rates.

The tests used in this simulation study were CM_n, CO_n, $\overline{CM_n}$, ZP_n, EP_n, $AT_{n,\gamma}$, $BH_{n,\gamma}$ and $L_{n,\gamma}$. Many of these tests were found to be the most powerful tests in the Monte Carlo study conducted in Sections 4.2 – 4.5. These tests also provide diversity in terms of how recently the test has been developed, the characteristics of the exponential distribution they are based on, and whether they contain a tuning parameter or not.

Tables 4.4 – 4.7 display the power estimates of the chosen tests, for sample sizes of $n = 20$ and $n = 50$, when the KL distances are set at $K(f,g) = 0.2$ and $K(f,g) = 0.5$ respectively. For the tests that contain a tuning parameter, the primary entry is the approximate power of the test based on the *data-dependent choice* of the parameter, $\hat{\gamma}$, while the approximate power of the test based on the *compromise choice* appears in parentheses along-side it. Again, the highest power for each alternative distribution is highlighted.

Alternative	$\mathbf{CM_n}$	$\mathbf{CO_n}$	$\overline{\mathbf{CM_n}}$	$\mathbf{ZP_n}$	$\mathbf{EP_n}$	$\mathbf{AT_{n,\hat{\gamma}}}$	$\mathbf{BH_{n,\hat{\gamma}}}$	$\mathbf{L_{n,\hat{\gamma}}}$
IHR								
$\Gamma(1.76)$	33	37	31	36	33	**48**(**48**)	35(37)	34(37)
$W(1.41)$	36	38	36	37	38	44(**46**)	36(38)	35(36)
$CH(1.05)$	19	18	20	17	20	21(**22**)	18(19)	16(17)
DHR								
$\Gamma(0.58)$	37	**56**	36	45	38	43(49)	52(43)	54(48)
$W(0.71)$	40	**54**	42	45	44	43(44)	50(47)	50(49)
NMHR								
$LN(1.1)$	19	13	**21**	9	19	14(7)	18(17)	16(11)
$CH(0.59)$	32	51	30	41	32	41(48)	51(40)	**54**(45)
$DH(1.05)$	27	28	23	28	24	**44**(38)	28(28)	28(31)

Table 4.4: Monte Carlo power estimates for $n = 20$ when $K(f, g) = 0.2$

Alternative	$\mathbf{CM_n}$	$\mathbf{CO_n}$	$\overline{\mathbf{CM_n}}$	$\mathbf{ZP_n}$	$\mathbf{EP_n}$	$\mathbf{AT_{n,\hat{\gamma}}}$	$\mathbf{BH_{n,\hat{\gamma}}}$	$\mathbf{L_{n,\hat{\gamma}}}$
IHR								
$\Gamma(1.76)$	74	82	74	79	75	84(**88**)	80(81)	77(83)
$W(1.41)$	76	84	79	79	81	78(**85**)	80(83)	77(82)
$CH(1.05)$	46	44	51	40	**52**	43(41)	50(49)	48(42)
DHR								
$\Gamma(0.58)$	72	**87**	71	78	72	84(**87**)	85(80)	86(84)
$W(0.71)$	77	**87**	79	79	80	83(85)	84(84)	84(85)
NMHR								
$LN(1.1)$	39	22	**41**	10	35	30(9)	34(30)	29(17)
$CH(0.59)$	68	**85**	65	76	65	81(**85**)	84(75)	**85**(81)
$DH(1.05)$	60	64	53	67	49	**83**(77)	71(63)	71(72)

Table 4.5: Monte Carlo power estimates for $n = 50$ when $K(f, g) = 0.2$

Alternative	$\mathbf{CM_n}$	$\mathbf{CO_n}$	$\overline{\mathbf{CM_n}}$	$\mathbf{ZP_n}$	$\mathbf{EP_n}$	$\mathbf{AT_{n,\hat{\gamma}}}$	$\mathbf{BH_{n,\hat{\gamma}}}$	$\mathbf{L_{n,\hat{\gamma}}}$
IHR								
$\Gamma(2.41)$	69	76	67	74	70	84(85)	72(75)	71(75)
$W(1.686)$	71	76	71	74	75	76(81)	71(75)	69(73)
$CH(1.18)$	38	37	39	35	41	38(41)	35(37)	34(31)
DHR								
$\Gamma(0.44)$	67	86	65	75	67	79(83)	83(74)	85(79)
$W(0.587)$	73	85	74	77	75	80(82)	83(80)	83(82)
NMHR								
$LN(1.5)$	62	60	65	55	66	61(44)	64(65)	63(60)
$CH(0.478)$	68	87	67	76	69	80(84)	84(76)	85(81)
$DH(1.54)$	69	75	66	74	68	86(85)	72(74)	71(76)

Table 4.6: Monte Carlo power estimates for $n = 20$ when $K(f, g) = 0.5$

Alternative	$\mathbf{CM_n}$	$\mathbf{CO_n}$	$\overline{\mathbf{CM_n}}$	$\mathbf{ZP_n}$	$\mathbf{EP_n}$	$\mathbf{AT_{n,\hat{\gamma}}}$	$\mathbf{BH_{n,\hat{\gamma}}}$	$\mathbf{L_{n,\hat{\gamma}}}$
IHR								
$\Gamma(2.41)$	99	100	99	99	99	99(100)	100(100)	98(100)
$W(1.686)$	99	100	99	99	100	95(100)	99(100)	95(99)
$CH(1.18)$	78	81	84	75	86	74(76)	83(83)	80(76)
DHR								
$\Gamma(0.44)$	97	100	96	98	96	99(100)	99(98)	99(99)
$W(0.587)$	98	100	98	98	98	99(99)	99(99)	99(99)
NMHR								
$LN(1.5)$	93	92	94	87	94	93(82)	95(95)	94(92)
$CH(0.478)$	97	100	97	99	97	99(99)	99(99)	100(99)
$DH(1.54)$	99	99	98	99	98	99(100)	100(99)	99(100)

Table 4.7: Monte Carlo power estimates for $n = 50$ when $K(f, g) = 0.5$

Tables 4.8 – 4.9 summarise the results in this section. Table 4.8 shows the number of the eight alternatives against which the specified test was the most powerful (the counts include ties, meaning that if more than one test achieved the highest power, all those tests' counts will increase by 1). The Atkinson test ($AT_{n,\gamma}$) using the compromise choice of the tuning parameter and the Cox and Oakes test (CO_n) performed the best according to these counts.

In Table 4.9 the averaged power estimates for each test over the eight alternatives are given. The average power estimates can be compared in this section, since the alternatives are of equal distance from the exponential distribution. For a sample size $n = 20$ the Atkinson test ($AT_{n,\gamma}$), using the compromise choice of the tuning parameter, achieved the highest average power. The Baringhaus and Henze test ($BH_{n,\gamma}$), using the data-dependent choice of the tuning parameter, was the superior test in terms of average power for a sample size of $n = 50$. The test which attained the best average power amongst those that do not contain a tuning parameter, was the Cox and Oakes test (CO_n). For the tests containing a tuning parameter, one can conclude that using the data-dependent choice of the tuning parameter, instead of the compromise choice, results in a slight improvement in the average powers in the majority of the cases.

$\mathbf{n}$	$\mathbf{K(f,g)}$	$\mathbf{CM_n}$	$\mathbf{CO_n}$	$\overline{\mathbf{CM_n}}$	$\mathbf{ZP_n}$	$\mathbf{EP_n}$	$\mathbf{AT_{n,\hat{\gamma}}}$	$\mathbf{BH_{n,\hat{\gamma}}}$	$\mathbf{L_{n,\hat{\gamma}}}$
20	0.2	0	2	1	0	0	2(3)	0(0)	1(0)
50	0.2	0	3	1	0	1	1(4)	0(0)	1(0)
20	0.5	0	3	0	0	2	1(3)	0(0)	0(0)
50	0.5	0	5	0	0	2	0(4)	3(3)	1(2)

Table 4.8: Number of the eight alternatives against which the specified tests have performed the best (including ties).

$\mathbf{n}$	$\mathbf{K(f,g)}$	$\mathbf{CM_n}$	$\mathbf{CO_n}$	$\overline{\mathbf{CM_n}}$	$\mathbf{ZP_n}$	$\mathbf{EP_n}$	$\mathbf{AT_{n,\hat{\gamma}}}$	$\mathbf{BH_{n,\hat{\gamma}}}$	$\mathbf{L_{n,\hat{\gamma}}}$
20	0.2	30.4	36.9	29.9	32.3	31	37.3(37.8)	36(33.6)	35.9(34.3)
50	0.2	64	69.4	64.1	63.5	63.6	70.8(69.6)	71(68.1)	69.6(68.3)
20	0.5	64.6	72.8	64.3	67.5	66.4	73(73.1)	70.5(69.5)	70.1(69.6)
50	0.5	95	96.5	95.6	94.3	96	94.6(94.5)	96.8(96.6)	95.5(95.5)

Table 4.9: Average Monte Carlo power estimates for the specified tests over the eight alternatives.

4.7 Practical application

In this section, the tests considered in Chapter 2 are applied to a real-world data set: the 'Leukemia' data set given below in Table 4.10 (see Kotz and Johnson (1983) for a discussion of the original data set). These data display the survival times (in days) of 43 patients diagnosed with a certain type of Leukemia.

Table 4.11 lists the names of the 20 different tests used in this dissertation along with the value of the test statistic calculated from these data, the p-value for testing the hypothesis of exponentiality, as well as the time (in seconds) taken to compute the p-value and critical value

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 4.10: Survival times in days after diagnosis.

for each test (based on $MC = 10\,000$ replications). Where applicable, the data-dependent choice of γ used is also displayed in the table. The number of bootstrap replications in the calculation of the data-dependent choice of the tuning parameter is set to $B = 1\,000$. The final column in the table indicates whether the test is available in the software package R (R Core Team, 2013); these tests are primarily available in the package `exptest` (Novikov et al., 2013).

Test	$\widehat{\gamma}$ **value**	**Test statistic value**	p**-value**	**Time**	**Available in R?**
EP_n		1.714	0.081	0.64	Y
$PW^1_{n,\widehat{\gamma}}$	0.10	-1.872	0.309	1377.19	N
$PW^2_{n,\widehat{\gamma}}$	0.25	-1.167	0.069	1363.20	N
$BH_{n,\widehat{\gamma}}$	2.50	0.079	0.086	24.94	N
$L_{n,\widehat{\gamma}}$	5.00	0.003	0.081	16.96	N
KS_n		0.162	0.054	1.72	Y
CM_n		0.150	0.148	1.67	Y
$\overline{KS}_n$		1.274	0.069	3.79	Y
$\overline{CM}_n$		0.331	0.108	107.28	Y
$J_{n,\widehat{\gamma}}$	0.90	1.198	0.040	92.35	N
ZP_n		0.160	0.079	1.47	N
BR_n		0.098	0.038	1.92	N
G_n		0.426	0.951	0.98	Y
HM_n		1.336	0.181	1.06	Y
CO_n		12.171	0.076	0.61	Y
AH^1_n		0.113	0.081	2301.30	Y
AH^2_n		0.016	0.112	2051.38	N
ZA_n		0.160	0.085	2.11	N
NA_n		0.090	0.069	4235.27	N
$AT_{n,\widehat{\gamma}}$	0.99	0.009	0.086	7.23	N

Table 4.11: Summary of results for the Leukaemia data set.

All of the tests except $J_{n,0.9}$ and BR_n do not reject the null hypothesis of exponentiality at a significance level of $\alpha = 0.05$.

As shown in Table 4.11, none of the tests containing a tuning parameter appear in R. These tests are rather powerful and therefore it might be a worthwhile avenue for future work to create an R package that includes these tests along with the procedure to obtain the tuning parameter data-dependently.

Chapter 5

Conclusions

In this dissertation a large number of tests for exponentiality based on a wide variety of characteristics of this distribution are considered. These characteristics as well as the tests associated with them are briefly mentioned below.

The tests based on the characteristic function are the Epps and Pulley test (EP_n) as well as tests based on the probability weighted empirical characteristic function. We consider two forms of this test; the first uses the original test statistic proposed in Meintanis et al. (2014) ($PW^1_{n,\gamma}$). The weight function used in this test statistic assigns the majority of the weight to the centre of the distribution. The second formulation of the test statistic considered ($PW^2_{n,\gamma}$) gives more weight towards the tails of the distributions.

The tests based on the empirical Laplace transform are those of Baringhaus and Henze ($BH_{n,\gamma}$) as well as Henze and Meintanis ($L_{n,\gamma}$).

Another characteristic of the exponential distribution that some of the tests are based on is the distribution function. The tests associated with this characteristic are the Kolmogorov-Smirnov (KS_n) and Cramér-von Mises (CM_n) tests.

Next, consider the tests based on the mean residual life of the data. The tests considered include those of Baringhaus and Henze. Two test statistics based on mean residual life introduced in Baringhaus and Henze (2000) are considered; a Kolmogorov-Smirnov type test ($\overline{KS}_n$) and a Cramér-von Mises type test ($\overline{CM}_n$). The test of Jammalamadaka and Taufer ($J_{n,\gamma}$) is also based on this characteristic.

Another characteristic used to test for exponentiality is entropy. Two tests based on entropy are discussed; the test of Zardasht *et al.* (ZP_n) and that of Baratpour and Habibi Rad (BR_n).

Furthermore, consider two tests based on the normalised spacings of the observed data. The first of these is the Gini test (G_n) and the second is Harris' modification of Gnedenko's F-test (HM_n).

The Cox and Oakes test (CO_n) is also included in the study. This test is based on a score function.

Various other characteristics are also used. Two tests based on Ahsanullah's characterisation are considered. The first (AH^1_n) uses the original test statistic proposed in Volkova and Nikitin (2013). The second test (AH^2_n) utilizes a Cramér-von Mises type test statistic. Zhang's test (ZA_n), based on likelihood ratios, is included in the study as well as the Noughabi and Arghami test (NA_n) which uses transformed data. Finally, the Atkinson test ($AT_{n,\gamma}$), based on the Atkinson statistic, is considered.

Based on the results of the Monte Carlo study conducted in this dissertation, some brief conclusions regarding the powers of the tests considered can be made. Generally, HM_n achieves powers substantially lower than the remaining tests. In addition, the AH_n^2, BR_n, and ZA_n tests are also relatively poor performers in terms of power. However, tests that do perform well are $BH_{n,\widehat{\gamma}}$, $L_{n,\widehat{\gamma}}$, $AT_{n,\widehat{\gamma}}$ and CO_n. $BH_{n,\widehat{\gamma}}$ has the best overall performance, closely followed by $L_{n,\widehat{\gamma}}$. Note that only one of the tests reported to perform relatively poorly contain a tuning parameter, while only one of the tests reported to achieve high powers do not contain such a parameter; CO_n performs the best among those tests that do not include a tuning parameter.

In light of the results discussed above, we would advise using the data-dependent choice of the tuning parameter; this choice generally outperforms the compromise choice. It is important to note that power associated with the data-dependent choice of the tuning parameter can conceivably be increased further by evaluating the powers over finer grids of tuning parameters than the grids used in this dissertation. Because of the large number of Monte Carlo replications required for the numerical results shown in this dissertation, finer grids would substantially increase the computational burden. However, in the case where the hypothesis of exponentiality is to be tested on a single dataset the computational time required is substantially less.

A compact summary of the contents of this dissertation has been published in Computational Statistics under the name “An ‘apples to apples’ comparison of various tests for exponentiality” (see Allison et al., 2017).

Bibliography

Abbasnejad, M., Arghami, N. R. and Tavakoli, M. (2012). A goodness of fit test for exponentiality based on Lin-Won information, *Journal of the Iranian Statistical Society* **11**(2): 191–202.

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

Ahsanullah, M. (1978). On a characterization of the exponential distributon by spacings, *Annals of the Institute of Statistical Mathematics Part A* **30**: 163–166.

Allison, J. S. and Santana, L. (2015). On a data-dependent choice of the tuning parameter appearing in certain goodness-of-fit tests, *Journal of Statistical Computation and Simulation* **85**(16): 3276–3288.

Allison, J. S., Santana, L., Smit, N. and Visagie, I. J. H. (2017). An 'apples to apples' comparison of various tests for exponentiality, *Computational Statistics* **DOI 10.1007/s00180-017-0733-3**.

Alzaid, A. A. and Al-Osh, M. A. (1992). Characterisation of probability distributions based on the relation, *Sankya Ser. B* **53**: 188–190.

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

Balakrishnan, N., Ng, H. K. T. and Kannan, N. (2002). *Goodness-of-Fit Tests and Model Validity*, Birkhäuser, Boston, chapter 8: A test of exponentiality based on spacings for progressively Type-II censored data, pp. 89–111.

Baratpour, S. and Habibi Rad, A. (2012). Testing goodness-of-fit for exponential distribution based on cumulative residual entropy, *Communications in Statistics - Theory and Methods* **41**(8): 1387–1396.

Baringhaus, L., Grtler, N. and Henze, N. (2000). Weighted integral test statistics and components of smooth tests of fit, *Australian & New Zealand Journal of Statistics* **42**: 179–192.

Baringhaus, L. and 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.

Baringhaus, L. and Henze, N. (2000). Tests of fit for exponentiality based on a characterization via the mean residual life function, *Statistical Papers* **41**: 225–236.

Baringhaus, L. and Henze, N. (2008). A new weighted integral goodness-of-fit statistic for exponentiality, *Statistics and Probability Letters* **78**: 1006–1016.

Bartholomew, D. J. (1957). Testing for departure from the exponential distribution, *Biometrika* **44**(1): 253–257.

Boos, D. D. and Brownie, C. (1989). Bootstrap methods for testing homogeneity of variances, *Technometrics* **31**(1): 69–82.

Chernick, M. (1999). *Bootstrap Methods: A Guide for Practitioners and Researchers*, Wiley, Hoboken.

Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*, Chapman and Hall, London.

D'Agostino, R. and Stephens, M. (1986). *Goodness-of-fit Techniques*, Marcel Dekker, New York.

Davidson, A. and Hinkley, D. (1997). *Bootstrap Methods and their Applications*, Cambridge University Press, Cambrigde.

Dhumal, B. and Shirke, D. (2014). A modified test for testing exponentiality using transformed data, *Journal of Statistical Planning and Simulation* **84**: 397–403.

Ebrahimi, N., Habibullah, M. and Soofi, E. S. (1992). Testing exponentiality based on kullback leibler information, *Journal of the Royal Statistical Society Series B* **54**: 739–748.

Efron, B. and Tibshirani, R. J. (1986). Bootstrap methods for standard error, confidence intervals, and other measures of statistical accuracy, *Statistical Science* **1**: 54–75.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, Chapman and Hall, New York.

Epps, T. W. and 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* **48**(2): 206–213.

Fisher, N. I. and Hall, P. (1990). On bootstrap hypothesis testing, *The Australian Journal of Statistics* **32**(2): 117–190.

Gail, M. H. and Gastwirth, J. L. (1978). A scale-free goodness-of-fit test for exponentiality based on the gini statistic, *Journal of the Royal Statistical Society B* **40**: 350–357.

Gnedenko, B. V., Belyayev, Y. K. and Solovyev, A. D. (1969). *Mathematical Models of Reliability Theory*, Academic Press.

Good, P. (2000). *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*, Springer, New York.

Grané, A. and Fortiana, J. (2011). A directional test of exponentiality based on maximum correlations, *Metrika* **73**: 255–274.

Grzegorzewski, P. and Wieczorkowski, R. (1999). Entropy-based goodness-of-fit test for exponentiality, *Communications in Statistics - Theory and Methods* **28**: 1183–1202.

Hahn, G. J. and Shapiro, S. S. (1967). *Statistical Models in Engineering*, Wiley, New York.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*, Springer, New York.

Hall, P. and Wilson, S. R. (1991). Two guidelines for bootstrap hypothesis testing, *Biometrics* **47**: 757–762.

Harris, C. M. (1976). A note on testing for exponentiality, *Naval Research Logistics Quarterly* **23**: 169–175.

Haywood, J. and Khmaladze, E. (2008). On distribution-free goodness-of-fit testing of exponentiality, *Journal of Econometrics* **143**: 5–18.

Hegazy, Y. A. S. and Green, J. R. (1975). Some new goodness-of-fit tests using order statistics, *Journal of Applied Statistics* **24**(3): 299–308.

Henze, N. (1993). A new flexible class of omnibus tests for exponentiality, *Communications in Statistics - Theory and Methods* **22**: 115–133.

Henze, N. and Meintanis, S. G. (2002). Tests of fit for exponentiality based on the empirical laplace transform, *Statistics* **36**(2): 147–161.

Henze, N. and Meintanis, S. G. (2005). Recent and classical tests for exponentiality: a partial review with comparisons, *Metrika* **36**: 29–45.

Jackson, O. A. Y. (1967). An analysis of departure from the exponential distribution, *Journal of the Royal Statistical Society Series B* **29**(3): 540–549.

Jammalamadaka, S. R. and Goria, M. N. (2004). A test of goodness-of-fit based on gini's index of spacings, *Statistics and Probability Letters* **68**(2): 177–187.

Jammalamadaka, S. R. and Taufer, E. (2003). Testing exponentiality by comparing the empirical distribution function of the normalized spacings with that of the original data, *Journal of Nonparametric Statistics* **15**(6): 719–729.

Jammalamadaka, S. R. and Taufer, E. (2006). Use of mean residual life in testing departures from exponentiality, *Journal of Nonparametric Statistics* **18**(3): 277–292.

Jovanović, M., Milošević, B., Nikitin, Y. Y., Obradović, M. and Volkova, K. Y. (2015). Tests of exponentiality based on Arnold–Villasenor characterization and their efficiencies, *Computational Statistics and Data Analysis* **90**: 100–113.

Klar, B. (2001). Goodness-of-fit tests for the exponential and the normal distribution based on the integrated distribution function, *Annals of the Institute of Statistical Mathematics* **53**(2): 338–353.

Kotz, S. and Johnson, N. L. (1983). *Encyclopedia of Statistical Sciences*, Vol. 3, Wiley, New York.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency, *The Annals of Mathematical Statistics* **22**(1): 79–86.

Lee, S. C. S., Locke, C. and Spurrier, J. D. (1980). On a class of tests of exponentiality, *Technometrics* **22**: 547–554.

Lemonte, A. J. (2013). A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function, *Computational Statistics and Data Analysis* **62**: 149–170.

Martin, M. A. (2007). Bootstrap hypothesis testing for some common statistical problems: A critical evaluation of size and power properties, *Computational Statistics and Data Analysis* **51**: 6321–6342.

Meintanis, S. G. and Iliopoulos, G. (2003). Characterizations of the exponential distribution based on certain properties of its characteristic function, *Kybernetika* **39**(3): 295–298.

Meintanis, S. G., Swanepoel, J. W. H. and Allison, J. S. (2014). The probability weighted characteristic function and goodness-of-fit testing, *Journal of Statistical Planning and Inference* **146**: 122–132.

Mimoto, N. and Zitikus, R. (2008). The atikinson index, the moran statistic, and testing exponentiality, *Journal of the Japan Statistical Society* **38**(2): 187–205.

Moran, P. A. P. (1951). The random division of an interval - part II, *Journal of the Royal Statistical Society Series B* **13**(1): 147–150.

Noughabi, H. A. (2015). Testing exponentiality based on the likelihood ratio and power comparison, *Annals of Data Science* **2**(2): 195–204.

Noughabi, H. A. and Arghami, N. R. (2011a). Testing exponentiality based on characterizations of the exponential distribution, *Journal of Statistical Computation and Simulation* **81**(11): 1641–1651.

Noughabi, H. A. and Arghami, N. R. (2011b). Testing exponentiality using transformed data, *Journal of Statistical Computation and Simulation* **81**(4): 511–516.

Novikov, A., Pusev, R. and Yakovlev, M. (2013). exptest: Tests for Exponentiality. R package version 1.2.
URL: *http://CRAN.R-project.org/package=exptest*

Parzen, E. (1998). Statistical methods mining, two sample data analysis, comparison distributions, and quantile limit theorems, *Asymptotic Methods in Probability and Statistics* **1**: 611–617.

R Core Team (2013). R: A Language and Environment for Statistical Computing.
URL: *http://www.R-project.org/*

Rao, M., Chen, Y., Vemuri, B. C., IEEE and Wang, F. (2004). Cumulative residual entropy: a new measure of information, *IEEE Transactions on Information Theory* **50**(6): 1220–1228.

Seshadri, V., Csorgo, M. and Stephens, M. (1969). Tests for the exponential distribution using Kolmogorov-type statistics., *Journal of the Royal Statistical Society B* **31**: 499–509.

Shanbhag, D. N. (1970). The characterizations for exponential and geometric distributions, *Journal of the American Statistical Association* **65**(331): 1256–1259.

Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap*, Springer, New York.

Shapiro, S. S. and Wilk, M. B. (1972). An analysis of variance test for the exponential distribution (complete samples), *Technometrics* **14**(2): 355–370.

Shimizu, R. (1979). On a lack of memory property of the exponential distribution, *Annals of the Institute of Statistical Mathematics Part A* **31**: 309–313.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Spurrier, J. D. (1984). An overview of tests for exponentiality, *Communications in Statistics - Theory and Methods* **13**: 1635–1654.

Taufer, E. (2000). A new test for exponentiality against omnibus alternatives, *Stochastic Modelling and Applications* **3**: 23–36.

Volkova, K. Y. (2010). On asymptotic efficiency of exponentiality tests based on Rossberg's characterization, *Journal of Mathematical Sciences* **167**(4): 486–494.

Volkova, K. Y. and Nikitin, Y. Y. (2013). Exponentiality tests based on Ahsanullah's characterization and their efficiency, *Journal of Mathematical Sciences* **204**(1): 42–54.

Wang, B. (2008). Goodness-of-fit test for the exponential distribution based on progressively type-II censored sample, *Journal of Statistical Computation and Simulation* **78**(2): 125–132.

Westfall, P. H. and Young, S. S. (1993). *Resampling-based Multiple Testing: Examples and Methods for p-value Adjustments*, Wiley, New York.

Wong, P. G. and Wong, S. P. (1979). An extremal quotient test for exponential distributions, *Metrika* **26**(1): 1–4.

Zardasht, V., Parsi, S. and Mousazadeh, M. (2015). On empirical cumulative residual entropy and a goodness-of-fit test for exponentiality, *Statistical Papers* **56**(3): 677–688.

Zhang, J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio, *Journal of the Royal Statistical Society, Series B* **64**(2): 281–294.

Appendix A

Monte Carlo power estimates

	KS_n	CM_n	HM_n	CO_n	$\overline{KS}_n$	$\overline{CM}_n$	G_n	AH_n^1	AH_n^2	BR_n	ZP_n	EP_n	ZA_n	NA_n	$AT_{n,\widehat{\gamma}}$	$J_{n,\widehat{\gamma}}$	$BH_{n,\widehat{\gamma}}$	$PW_{n,\widehat{\gamma}}^1$	$PW_{n,\widehat{\gamma}}^2$	$L_{n,\widehat{\gamma}}$
CHR																				
$\Gamma(1)$	5	5	5	5	5	5	4	5	5	5	5	5	5	5	5(5)	5(5)	5(5)	6(5)	5(4)	5(5)
IHR																				
$\Gamma(1.5)$	11	12	5	11	14	11	10	9	**17**	13	13	11	14	**17**	16(15)	12(**17**)	11(12)	11(10)	11(1)	11(12)
$\Gamma(2)$	23	25	7	24	27	23	22	19	30	24	28	23	29	33	**35**(33)	26(32)	27(29)	21(24)	25(0)	26(27)
$W(1.2)$	9	9	4	8	11	8	7	8	**13**	11	9	8	10	**13**	12(11)	9(12)	9(10)	9(8)	8(2)	8(6)
$W(1.5)$	21	24	6	22	28	23	22	20	30	27	26	23	26	**33**	32(32)	24(30)	25(27)	21(22)	21(1)	24(24)
$PW(1)$	30	36	8	27	43	36	35	27	37	**57**	31	32	32	38	31(32)	30(28)	32(32)	37(28)	12(6)	30(25)
$CH(1)$	8	9	5	8	12	9	8	7	12	**13**	9	8	9	**13**	11(11)	9(11)	8(9)	10(8)	6(2)	8(8)
$CH(1.5)$	36	44	8	42	47	44	43	37	49	**53**	46	44	44	52	49(50)	37(43)	44(46)	38(40)	30(1)	43(40)
$LF(2)$	13	14	5	12	18	14	13	12	19	**20**	15	13	14	**20**	18(17)	14(19)	14(15)	15(13)	9(2)	13(13)
$LF(4)$	18	20	6	18	24	19	18	17	25	26	20	19	20	**27**	26(**27**)	19(24)	20(21)	20(18)	13(1)	19(18)
$EV(0.5)$	8	8	5	7	11	8	7	7	**12**	**12**	9	7	9	**12**	10(11)	9(**12**)	8(9)	10(8)	6(3)	7(8)
$EV(1.5)$	19	21	6	17	27	21	19	17	26	**30**	20	19	19	27	24(24)	20(24)	20(22)	21(18)	11(2)	19(18)
$B(2,1)$	84	94	28	91	93	95	95	91	95	**99**	93	94	93	95	89(92)	84(84)	85(93)	91(91)	57(26)	82(88)
$B(1,2)$	11	13	6	11	17	13	12	10	16	**21**	13	12	13	17	15(15)	12(15)	12(13)	13(11)	7(3)	11(11)
$EP(1.5)$	48	57	10	54	60	59	58	49	61	**69**	58	58	55	64	59(61)	48(54)	55(58)	50(52)	34(2)	53(51)
$BEX(2)$	19	21	6	20	24	20	18	18	29	22	24	20	25	**31**	**31**(28)	23(28)	21(22)	18(19)	22(0)	21(21)
DHR																				
$\Gamma(0.4)$	41	44	14	67	28	44	46	50	0	6	52	46	63	0	56(62)	44(0)	69(57)	53(64)	60(63)	**71**(62)
$\Gamma(0.7)$	8	9	7	16	5	10	11	13	1	2	12	10	12	1	11(13)	12(1)	17(14)	9(15)	15(**19**)	18(15)
$W(0.8)$	10	11	8	16	7	12	13	12	1	3	12	13	12	1	11(11)	12(1)	**18**(16)	9(16)	14(**18**)	**18**(16)
$EG(0.2)$	5	5	5	5	5	5	5	5	4	3	5	5	5	4	5(5)	6(4)	**7**(**7**)	5(6)	6(**7**)	**7**(**7**)
$EG(0.5)$	8	8	6	9	5	8	9	7	2	2	8	9	6	2	6(5)	9(2)	10(9)	5(8)	8(**11**)	10(9)
$EG(0.8)$	19	21	11	23	13	24	**25**	12	1	5	20	**25**	16	1	18(12)	22(1)	**25**(**25**)	12(20)	20(24)	24(23)
$EL(0.2)$	9	9	7	12	5	11	12	10	2	2	10	12	8	2	10(7)	11(1)	14(13)	7(12)	12(**15**)	13(13)
$EL(0.5)$	5	5	5	6	4	6	6	6	3	3	5	6	5	3	5(4)	6(3)	8(7)	5(7)	6(**9**)	8(7)
$EL(0.8)$	5	5	5	5	5	4	5	5	4	4	5	5	4	4	5(5)	5(4)	**6**(**6**)	5(5)	5(5)	**6**(5)
$ENH1(0.5)$	48	52	19	**69**	37	54	57	47	0	13	57	56	61	0	58(60)	51(0)	**69**(62)	53(63)	62(65)	68(66)
$BEX(0.4)$	39	42	14	66	26	42	44	49	0	6	50	43	61	0	54(60)	45(0)	67(53)	52(61)	58(61)	**69**(59)
$BEX(0.7)$	9	9	8	16	5	10	12	14	1	2	12	11	12	1	10(13)	12(2)	17(14)	9(16)	15(**19**)	18(15)
NMHR																				
$PW(2)$	10	11	21	19	7	8	6	13	3	6	12	6	23	3	13(20)	11(2)	24(12)	17(25)	24(**29**)	26(16)
$PW(3)$	33	36	24	63	19	31	27	41	0	2	44	29	63	0	53(62)	35(0)	66(42)	51(63)	59(62)	**69**(52)
$LN(0.8)$	16	17	9	16	17	15	14	15	27	14	19	15	23	27	**31**(25)	21(26)	18(19)	16(18)	26(0)	18(19)
$LN(1.5)$	31	34	19	33	25	37	38	11	1	13	31	38	25	1	31(19)	35(1)	**39**(**39**)	23(32)	30(37)	37(37)
$DH(1)$	13	13	7	11	15	11	10	12	20	12	13	11	16	**21**	**21**(19)	15(20)	13(14)	12(13)	16(0)	13(13)
$DH(1.5)$	29	33	10	34	35	31	30	28	42	32	38	32	40	45	**49**(45)	36(43)	36(38)	29(33)	37(0)	35(36)
$CH(0.5)$	32	34	11	54	20	34	36	39	0	4	41	36	47	0	42(46)	34(0)	55(44)	38(49)	46(51)	**56**(49)
HN	11	12	4	10	15	11	10	10	16	16	12	10	12	**17**	14(14)	11(15)	11(12)	13(11)	8(2)	10(10)
$B(0.5,1)$	10	10	21	19	7	7	6	13	3	6	12	6	23	3	13(20)	11(2)	23(11)	17(24)	24(**29**)	26(16)
$EP(0.5)$	19	20	11	38	11	20	22	28	1	2	26	21	34	1	26(34)	21(1)	40(28)	25(36)	33(38)	**42**(33)
$ENH1(2)$	7	7	7	5	6	**8**	7	4	7	4	5	**8**	6	6	7(5)	7(6)	**8**(**8**)	6(6)	7(3)	**8**(6)
$ENH2(0.4)$	32	34	13	59	18	32	33	41	0	2	42	33	55	0	45(52)	33(0)	57(41)	43(52)	48(51)	**61**(48)
$ENH2(0.7)$	6	6	8	9	5	5	5	9	3	3	8	5	9	3	6(9)	6(3)	11(7)	7(11)	9(**12**)	**12**(9)

Table A.1: Monte Carlo power estimates for $n = 10$.

	KS_n	CM_n	HM_n	CO_n	$\overline{KS}_n$	$\overline{CM}_n$	G_n	AH_n^1	AH_n^2	BR_n	ZP_n	EP_n	ZA_n	NA_n	$AT_{n,\widehat{\gamma}}$	$J_{n,\widehat{\gamma}}$	$BH_{n,\widehat{\gamma}}$	$PW_{n,\widehat{\gamma}}^1$	$PW_{n,\widehat{\gamma}}^2$	$L_{n,\widehat{\gamma}}$
CHR																				
$\Gamma(1)$	5	6	5	5	5	5	5	5	5	5	5	6	5	5	5(5)	5(5)	4(4)	5(4)	5(5)	4(4)
IHR																				
$\Gamma(1.5)$	17	21	6	24	22	20	19	18	26	19	23	21	27	28	**29**(28)	17(23)	19(20)	15(17)	19(1)	19(20)
$\Gamma(2)$	41	48	10	57	46	47	45	46	57	38	56	48	59	61	**64**(**64**)	44(51)	49(52)	40(46)	48(1)	49(53)
$W(1.2)$	12	14	5	15	16	13	13	12	18	14	15	14	16	**20**	18(18)	11(16)	12(14)	11(11)	11(1)	12(13)
$W(1.5)$	40	48	8	55	49	49	48	45	55	46	53	51	52	**59**	56(**59**)	41(50)	49(52)	38(44)	37(1)	48(49)
$PW(1)$	54	68	9	54	74	71	71	56	66	**92**	55	68	69	66	55(52)	57(48)	62(59)	63(53)	29(32)	59(46)
$CH(1)$	13	15	5	15	18	15	15	12	18	**20**	14	15	15	**20**	14(15)	12(15)	12(14)	12(10)	7(3)	11(11)
$CH(1.5)$	68	80	12	83	79	83	84	76	82	86	81	85	80	84	76(**88**)	64(71)	78(81)	68(73)	50(8)	76(76)
$LF(2)$	24	28	5	28	32	29	28	23	32	34	27	30	27	**35**	28(30)	21(27)	25(28)	22(22)	14(2)	24(24)
$LF(4)$	34	42	6	42	45	44	44	36	44	**48**	42	45	40	**48**	39(42)	33(41)	38(41)	31(33)	21(3)	36(35)
$EV(0.5)$	13	15	5	14	18	15	14	12	17	**21**	14	15	15	20	14(15)	12(15)	13(14)	12(11)	7(3)	12(12)
$EV(1.5)$	37	46	6	41	50	49	48	36	45	**60**	42	48	42	48	40(41)	34(38)	41(42)	35(34)	16(5)	39(34)
$B(2,1)$	99	**100**	46	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	97(**100**)	99(99)	99(**100**)	**100**(**100**)	93(94)	88(**100**)
$B(1,2)$	20	25	6	24	30	28	27	19	25	**43**	23	27	25	29	23(23)	19(22)	22(23)	20(17)	9(6)	20(18)
$EP(1.5)$	81	92	16	91	91	94	94	88	92	**97**	90	94	91	93	84(90)	78(82)	89(92)	83(86)	58(19)	85(87)
$BEX(2)$	34	41	9	48	39	38	37	39	51	29	48	40	53	54	**59**(57)	38(44)	43(45)	35(40)	43(0)	43(46)
DHR																				
$\Gamma(0.4)$	70	75	13	**91**	62	74	75	84	0	25	80	75	88	0	87(90)	65(64)	90(83)	84(86)	85(86)	**91**(87)
$\Gamma(0.7)$	15	18	8	28	11	18	20	24	1	5	21	19	20	1	20(23)	16(9)	27(22)	15(23)	23(28)	**29**(24)
$W(0.8)$	18	20	7	**28**	14	22	24	22	1	7	21	23	18	1	21(22)	19(7)	27(25)	15(22)	22(27)	**28**(25)
$EG(0.2)$	6	6	5	6	5	6	6	6	3	4	5	6	5	3	5(5)	6(4)	6(6)	5(6)	6(**7**)	6(6)
$EG(0.5)$	11	12	6	14	9	14	**15**	10	2	6	11	**15**	8	2	10(8)	14(3)	14(14)	7(11)	10(14)	14(13)
$EG(0.8)$	34	39	13	41	32	43	45	25	0	20	34	45	28	0	39(28)	43(7)	44(**46**)	26(34)	32(39)	42(42)
$EL(0.2)$	16	18	7	20	12	19	20	15	1	6	18	20	11	1	16(13)	18(4)	20(20)	10(16)	17(**22**)	19(19)
$EL(0.5)$	7	7	6	8	6	8	8	6	2	5	7	8	5	2	6(5)	8(3)	8(8)	5(7)	8(**10**)	8(8)
$EL(0.8)$	5	5	5	**6**	5	5	5	**6**	4	4	5	**6**	4	4	4(5)	5(4)	**6**(5)	5(5)	**6**(**6**)	5(5)
$ENH1(0.5)$	80	85	24	**93**	75	85	86	80	0	47	86	86	89	0	89(90)	76(62)	91(88)	83(87)	87(88)	91(90)
$BEX(0.4)$	68	73	12	**90**	58	72	74	84	0	21	80	74	86	0	85(89)	63(62)	88(80)	81(85)	83(85)	**90**(85)
$BEX(0.7)$	15	16	8	26	10	16	18	23	1	4	20	18	20	1	19(23)	15(8)	26(21)	15(23)	23(**28**)	**28**(24)
NMHR																				
$PW(2)$	15	18	34	25	13	11	6	20	2	15	18	7	39	3	26(32)	14(16)	36(14)	29(40)	46(**51**)	43(21)
$PW(3)$	58	63	35	84	43	53	42	73	0	4	70	45	86	0	84(87)	55(64)	88(65)	81(86)	85(87)	**91**(75)
$LN(0.8)$	31	35	17	36	29	28	25	38	55	18	37	26	60	52	**63**(48)	33(38)	35(33)	34(39)	50(0)	37(38)
$LN(1.5)$	57	62	28	59	54	65	**66**	23	1	43	53	**66**	49	0	60(43)	64(16)	65(65)	44(51)	49(56)	63(60)
$DH(1)$	21	25	11	28	23	22	20	24	36	16	28	22	36	35	**39**(35)	21(27)	23(24)	21(25)	32(0)	23(25)
$DH(1.5)$	57	67	15	74	62	65	63	66	77	47	73	65	76	80	**84**(82)	62(68)	69(71)	62(67)	68(1)	69(74)
$CH(0.5)$	57	63	11	**82**	48	62	63	72	0	16	70	63	74	0	73(77)	52(46)	79(70)	67(73)	73(75)	80(75)
HN	18	22	5	20	26	22	22	18	24	**27**	21	22	21	**27**	21(22)	16(21)	18(20)	16(15)	10(3)	17(17)
$B(0.5,1)$	15	19	34	26	12	11	6	20	2	15	18	6	40	2	26(32)	15(17)	35(13)	30(39)	46(**50**)	42(20)
$EP(0.5)$	34	39	11	62	25	36	37	54	0	5	47	37	56	0	53(60)	32(30)	63(47)	47(56)	55(59)	**66**(54)
$ENH1(2)$	10	11	10	8	9	**12**	10	4	9	9	6	11	11	6	9(6)	**12**(5)	10(10)	8(9)	9(3)	9(7)
$ENH2(0.4)$	56	60	15	83	44	56	54	74	0	8	68	55	79	0	78(82)	48(52)	83(67)	73(78)	76(78)	**86**(75)
$ENH2(0.7)$	7	7	10	12	5	6	5	12	2	3	8	5	11	2	8(12)	7(6)	13(8)	8(13)	13(**17**)	16(10)

Table A.2: Monte Carlo power estimates for $n = 20$.

	KS_n	CM_n	HM_n	CO_n	$\overline{KS}_n$	$\overline{CM}_n$	G_n	AH_n^1	AH_n^2	BR_n	ZP_n	EP_n	ZA_n	NA_n	$AT_{n,\widehat{\gamma}}$	$J_{n,\widehat{\gamma}}$	$BH_{n,\widehat{\gamma}}$	$PW_{n,\widehat{\gamma}}^1$	$PW_{n,\widehat{\gamma}}^2$	$L_{n,\widehat{\gamma}}$
CHR																				
$\Gamma(1)$	6	5	5	5	5	5	5	5	5	5	5	5	4	5	5(5)	5(5)	5(5)	5(5)	5(5)	5(5)
IHR																				
$\Gamma(1.5)$	25	28	6	33	28	28	28	26	38	21	32	28	36	39	**40**(39)	23(30)	29(30)	23(27)	24(0)	29(32)
$\Gamma(2)$	60	69	13	78	64	69	69	66	77	51	75	70	78	79	81(**82**)	59(66)	72(74)	64(69)	65(1)	71(77)
$W(1.2)$	16	18	5	20	20	19	19	16	24	18	20	20	21	**26**	22(22)	15(21)	17(18)	14(16)	13(1)	17(19)
$W(1.5)$	58	68	10	75	65	70	72	63	74	61	72	72	71	**77**	71(76)	58(66)	68(72)	57(63)	51(2)	67(71)
$PW(1)$	72	86	9	71	90	90	90	73	82	**99**	72	86	92	81	71(63)	78(66)	83(79)	84(76)	48(56)	81(65)
$CH(1)$	18	20	6	18	24	22	22	15	22	**28**	18	21	19	25	18(19)	17(20)	18(20)	17(15)	7(3)	18(17)
$CH(1.5)$	86	94	14	95	93	96	**97**	92	95	**97**	94	**97**	94	95	88(94)	83(88)	93(96)	88(92)	65(15)	87(93)
$LF(2)$	35	43	6	41	46	47	**48**	33	44	**48**	41	47	39	47	37(39)	33(39)	40(42)	33(33)	18(3)	39(36)
$LF(4)$	51	60	7	58	62	64	**65**	50	61	**65**	59	**65**	55	64	52(55)	48(55)	58(59)	47(49)	27(4)	57(52)
$EV(0.5)$	19	22	5	20	26	24	24	14	22	**30**	19	24	21	25	20(21)	16(21)	18(19)	17(15)	8(4)	17(17)
$EV(1.5)$	51	63	7	55	66	69	69	50	62	**79**	56	68	59	64	55(52)	51(53)	61(60)	52(50)	21(9)	59(50)
$B(2,1)$	**100**	**100**	62	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	99(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)	99(**100**)	96(**100**)
$B(1,2)$	31	39	7	32	45	44	45	25	35	**65**	31	43	38	38	33(32)	30(31)	36(36)	31(27)	11(9)	36(28)
$EP(1.5)$	95	99	21	98	98	99	99	97	99	**100**	98	99	99	98	91(97)	93(94)	97(99)	96(98)	72(37)	90(97)
$BEX(2)$	51	58	12	67	54	57	57	57	70	39	66	58	70	72	**77**(75)	51(57)	62(63)	54(60)	57(0)	61(67)
DHR																				
$\Gamma(0.4)$	87	90	14	**98**	81	89	89	95	27	42	93	89	96	13	96(97)	83(88)	97(94)	93(95)	95(95)	97(96)
$\Gamma(0.7)$	22	26	8	38	18	26	28	33	1	8	29	28	27	0	29(**43**)	21(19)	36(30)	23(31)	31(37)	38(33)
$W(0.8)$	24	28	7	**38**	21	31	32	30	1	12	29	32	23	0	29(31)	26(19)	36(34)	21(30)	29(35)	36(35)
$EG(0.2)$	6	6	5	6	5	6	6	6	3	4	6	6	5	3	5(5)	7(4)	6(6)	5(6)	6(**8**)	7(6)
$EG(0.5)$	15	16	6	17	13	**19**	**19**	13	1	9	15	**19**	9	1	14(11)	18(8)	18(18)	10(14)	13(17)	18(17)
$EG(0.8)$	49	55	14	58	48	60	**61**	34	0	35	48	**61**	40	0	55(43)	53(27)	59(60)	38(46)	43(49)	56(56)
$EL(0.2)$	23	26	7	28	19	28	**29**	21	1	11	25	**29**	15	0	22(18)	24(13)	28(28)	16(23)	22(27)	26(28)
$EL(0.5)$	8	8	6	8	7	9	9	7	2	5	8	9	5	2	6(6)	8(4)	9(9)	6(8)	7(**10**)	9(9)
$EL(0.8)$	**6**	5	5	5	5	**6**	**6**	5	4	5	5	5	4	4	5(5)	**6**(5)	5(5)	4(5)	5(**6**)	5(5)
$ENH1(0.5)$	92	95	27	**98**	91	95	96	94	20	69	95	96	97	10	97(**98**)	89(90)	**98**(97)	94(96)	96(96)	**98**(**98**)
$BEX(0.4)$	86	90	13	**97**	80	89	88	94	25	37	92	89	95	12	96(**97**)	81(87)	**97**(93)	93(95)	94(94)	**97**(95)
$BEX(0.7)$	20	23	9	35	15	23	24	33	1	7	26	24	24	1	27(32)	20(19)	34(28)	21(30)	29(35)	**36**(32)
NMHR																				
$PW(2)$	22	27	46	31	17	16	6	24	2	27	22	7	56	2	38(41)	20(27)	49(17)	48(56)	64(**68**)	56(27)
$PW(3)$	76	81	46	93	64	71	53	87	24	8	84	58	95	13	95(96)	76(84)	96(80)	92(96)	95(96)	**98**(90)
$LN(0.8)$	45	49	21	47	39	39	33	56	74	20	52	33	82	70	**87**(64)	38(43)	60(45)	58(63)	70(0)	61(57)
$LN(1.5)$	74	78	38	76	73	**81**	**81**	34	0	64	68	**81**	65	0	78(60)	78(49)	80(80)	63(68)	65(70)	79(76)
$DH(1)$	30	33	13	37	29	29	28	34	50	18	38	28	49	48	**57**(46)	26(34)	36(33)	34(40)	45(0)	37(40)
$DH(1.5)$	79	85	22	90	81	83	82	86	92	59	90	82	93	93	**94**(**94**)	79(83)	89(88)	84(87)	83(1)	85(92)
$CH(0.5)$	74	79	11	**92**	67	78	78	88	11	27	83	78	86	3	89(**92**)	69(74)	91(85)	83(88)	87(88)	91(89)
HN	25	30	4	28	33	33	33	23	32	**38**	28	33	28	35	28(29)	24(29)	28(29)	23(23)	11(3)	27(26)
$B(0.5,1)$	22	27	48	31	17	15	6	25	3	27	23	6	56	2	37(41)	19(26)	49(16)	47(55)	64(**67**)	56(27)
$EP(0.5)$	51	55	14	76	40	52	51	69	4	8	64	51	70	1	71(76)	47(54)	76(61)	64(71)	69(72)	**79**(69)
$ENH1(2)$	12	13	12	8	11	15	11	4	10	14	6	13	**16**	7	13(6)	14(4)	11(11)	12(11)	9(3)	10(7)
$ENH2(0.4)$	72	77	18	93	62	73	69	88	16	13	83	70	90	6	92(**94**)	71(78)	93(83)	87(90)	89(90)	**94**(89)
$ENH2(0.7)$	7	7	11	12	5	5	5	14	1	3	10	5	12	2	11(15)	8(10)	16(8)	11(16)	17(**21**)	20(11)

Table A.3: Monte Carlo power estimates for $n = 30$.

	KS_n	CM_n	HM_n	CO_n	$\overline{KS}_n$	$\overline{CM}_n$	G_n	AH^1_n	AH^2_n	BR_n	ZP_n	EP_n	ZA_n	NA_n	$AT_{n,\widehat{\gamma}}$	$J_{n,\widehat{\gamma}}$	$BH_{n,\widehat{\gamma}}$	$PW^1_{n,\widehat{\gamma}}$	$PW^2_{n,\widehat{\gamma}}$	$L_{n,\widehat{\gamma}}$
CHR																				
$\Gamma(1)$	6	5	5	5	5	5	5	5	5	6	5	5	5	5	5(5)	5(5)	5(5)	6(5)	5(5)	5(5)
IHR																				
$\Gamma(1.5)$	38	45	10	54	42	45	46	46	57	29	52	46	55	58	58(**60**)	34(43)	49(52)	40(47)	41(0)	48(55)
$\Gamma(2)$	83	91	24	96	87	91	91	91	94	67	94	91	95	95	94(**97**)	81(86)	94(95)	89(92)	86(1)	91(96)
$W(1.2)$	24	28	6	33	29	30	31	27	36	23	31	31	32	**38**	32(34)	21(28)	30(32)	21(26)	20(1)	29(32)
$W(1.5)$	82	90	17	**95**	88	92	93	88	92	80	93	93	92	92	88(94)	77(84)	92(94)	83(89)	74(2)	87(93)
$PW(1)$	93	99	9	91	99	99	99	93	96	**100**	91	98	**100**	95	89(84)	98(85)	95(96)	98(96)	77(84)	93(88)
$CH(1)$	28	35	6	31	39	39	40	25	33	**45**	30	39	33	36	29(28)	24(28)	34(35)	25(24)	10(4)	32(28)
$CH(1.5)$	98	**100**	26	**100**	**100**	**100**	**100**	99	**100**	**100**	**100**	**100**	**100**	99	94(**100**)	96(98)	99(**100**)	98(99)	83(31)	95(**100**)
$LF(2)$	54	65	6	61	66	70	**71**	54	63	69	61	**71**	60	65	59(57)	49(56)	66(66)	51(54)	27(4)	63(57)
$LF(4)$	73	83	10	81	83	87	**88**	75	82	84	80	**88**	80	84	74(74)	68(74)	85(86)	72(77)	44(6)	81(79)
$EV(0.5)$	27	33	6	31	38	38	39	25	34	**46**	29	38	34	37	32(30)	26(29)	35(35)	25(25)	10(4)	33(28)
$EV(1.5)$	75	87	7	79	88	91	92	76	83	**95**	79	90	87	83	79(71)	73(72)	88(86)	78(77)	32(16)	85(76)
$B(2,1)$	**100**	**100**	88	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)
$B(1,2)$	48	60	8	52	67	69	69	43	53	**90**	47	67	70	54	57(46)	49(44)	66(60)	52(47)	15(13)	64(47)
$EP(1.5)$	**100**	**100**	39	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	95(**100**)	99(**100**)	**100**(**100**)	**100**(**100**)	88(68)	96(**100**)
$BEX(2)$	75	82	23	90	77	81	81	87	92	50	89	81	90	92	92(**94**)	73(77)	88(88)	83(85)	81(0)	85(90)
DHR																				
$\Gamma(0.4)$	98	99	15	**100**	97	99	98	99	89	68	99	98	**100**	87	99(**100**)	95(98)	**100**(**100**)	97(**100**)	**100**(**100**)	**100**(**100**)
$\Gamma(0.7)$	33	38	10	**54**	28	39	40	48	6	12	44	40	40	4	47(52)	30(34)	52(47)	38(48)	45(51)	**54**(51)
$W(0.8)$	37	42	8	**54**	34	44	46	45	5	21	44	46	36	3	48(50)	35(34)	52(51)	36(44)	41(48)	51(53)
$EG(0.2)$	7	7	6	8	6	8	7	6	3	5	7	8	5	2	6(5)	8(5)	8(8)	6(7)	7(**9**)	8(8)
$EG(0.5)$	21	24	7	25	20	28	28	17	1	13	22	28	13	0	22(18)	23(14)	29(**30**)	16(22)	17(23)	27(28)
$EG(0.8)$	69	76	20	78	72	80	80	52	7	56	69	81	61	4	79(66)	71(53)	81(**82**)	60(67)	61(65)	79(77)
$EL(0.2)$	33	39	7	41	32	42	43	33	3	18	37	44	23	1	38(31)	33(25)	43(**45**)	27(36)	32(38)	40(43)
$EL(0.5)$	10	11	5	**12**	9	**12**	**12**	9	2	7	10	**12**	6	2	8(7)	11(6)	**12**(**12**)	7(10)	9(**12**)	11(**12**)
$EL(0.8)$	5	6	5	6	5	6	5	6	3	5	6	6	4	3	5(4)	6(4)	6(6)	6(6)	6(**7**)	6(6)
$ENH1(0.5)$	99	**100**	38	**100**	99	**100**	**100**	99	87	91	**100**	**100**	**100**	87	**100**(**100**)	98(99)	**100**(**100**)	98(**100**)	99(**100**)	**100**(**100**)
$BEX(0.4)$	98	98	14	**100**	96	98	98	99	88	62	99	98	99	85	99(**100**)	95(98)	**100**(99)	97(99)	99(99)	**100**(**100**)
$BEX(0.7)$	30	33	10	49	25	33	34	45	6	9	40	34	36	3	44(50)	29(32)	49(42)	36(45)	43(49)	**51**(48)
NMHR																				
$PW(2)$	35	47	73	42	29	28	5	32	12	57	32	6	85	8	59(58)	32(36)	68(25)	73(78)	84(**86**)	74(39)
$PW(3)$	94	96	69	99	88	91	71	97	87	24	97	77	**100**	81	99(**100**)	93(96)	**100**(95)	96(**100**)	99(**100**)	**100**(98)
$LN(0.8)$	72	77	44	66	63	60	47	86	95	23	76	45	98	92	**99**(84)	57(63)	92(68)	86(89)	88(0)	93(81)
$LN(1.5)$	91	94	56	92	92	**95**	**95**	52	12	86	87	**95**	86	6	93(83)	91(79)	94(94)	85(87)	82(85)	94(92)
$DH(1)$	46	53	23	55	45	45	41	62	75	21	61	40	75	**79**	**79**(68)	43(50)	64(56)	61(66)	66(0)	63(65)
$DH(1.5)$	96	98	44	99	97	98	97	99	99	76	99	97	99	99	99(**100**)	94(97)	99(99)	96(99)	96(1)	98(**100**)
$CH(0.5)$	92	95	13	**99**	90	94	94	97	71	49	97	94	98	65	98(**99**)	89(94)	98(97)	95(97)	97(97)	**999**(98)
HN	40	48	6	45	52	54	56	39	48	**57**	44	55	46	52	44(41)	36(41)	50(50)	37(37)	16(4)	48(41)
$B(0.5,1)$	36	48	74	41	31	28	6	32	13	58	33	7	85	8	60(60)	31(36)	68(25)	75(78)	86(**88**)	73(39)
$EP(0.5)$	72	77	20	92	64	73	70	88	44	15	84	71	88	35	91(**93**)	68(77)	92(83)	86(89)	88(89)	**93**(89)
$ENH1(2)$	16	18	19	9	16	19	14	6	15	22	6	15	**29**	8	20(7)	23(4)	15(13)	18(19)	14(2)	13(8)
$ENH2(0.4)$	92	94	24	**99**	88	92	88	97	78	26	96	89	98	71	**99**(**99**)	89(94)	**99**(96)	96(98)	98(98)	**99**(98)
$ENH2(0.7)$	11	12	17	18	8	8	6	18	2	4	15	7	19	2	17(21)	9(13)	24(11)	16(24)	25(**31**)	28(16)

Table A.4: Monte Carlo power estimates for $n = 50$.

	KS_n	CM_n	HM_n	CO_n	$\overline{KS}_n$	$\overline{CM}_n$	G_n	AH_n^1	AH_n^2	BR_n	ZP_n	EP_n	ZA_n	NA_n	$AT_{n,\widehat{\gamma}}$	$J_{n,\widehat{\gamma}}$	$BH_{n,\widehat{\gamma}}$	$PW_{n,\widehat{\gamma}}^1$	$PW_{n,\widehat{\gamma}}^2$	$L_{n,\widehat{\gamma}}$
CHR																				
$\Gamma(1)$	5	5	5	5	5	5	5	5	5	5	5	5	6	5	5(5)	5(5)	5(5)	5(5)	5(4)	5(5)
IHR																				
$\Gamma(1.5)$	54	63	11	74	57	62	63	64	74	33	70	63	72	74	73(78)	49(57)	69(70)	59(65)	56(0)	65(72)
$\Gamma(2)$	96	98	33	100	97	98	98	99	99	81	99	98	99	99	98(100)	92(95)	99(99)	96(99)	95(1)	98(100)
$W(1.2)$	34	42	6	49	41	43	44	38	48	27	45	44	45	49	46(51)	31(39)	43(46)	30(36)	26(0)	40(44)
$W(1.5)$	95	98	22	100	97	99	99	98	99	92	99	99	99	99	93(99)	91(95)	98(99)	95(98)	88(2)	95(99)
$PW(1)$	99	100	8	99	100	100	100	99	100	100	98	100	100	99	90(94)	100(97)	99(100)	100(100)	91(96)	95(98)
$CH(1)$	39	48	5	46	53	55	56	34	44	61	41	55	50	46	44(38)	36(39)	51(48)	33(34)	11(4)	47(38)
$CH(1.5)$	100	100	35	100	100	100	100	100	100	100	100	100	100	100	96(100)	98(100)	100(100)	100(100)	92(49)	98(100)
$LF(2)$	73	83	6	80	84	87	88	72	79	84	77	87	80	81	75(71)	68(74)	85(85)	70(74)	36(3)	83(77)
$LF(4)$	90	96	11	94	95	97	98	90	94	96	93	98	94	94	86(89)	85(89)	95(96)	88(91)	58(7)	92(92)
$EV(0.5)$	39	48	5	46	52	55	55	33	42	61	40	55	50	45	44(38)	37(40)	54(50)	35(36)	11(4)	50(40)
$EV(1.5)$	91	97	7	93	98	98	99	90	95	100	92	98	98	93	87(85)	91(88)	96(96)	93(93)	42(23)	94(91)
$B(2,1)$	100	100	97	100	100	100	100	100	100	100	100	100	100	100	100(100)	100(100)	100(100)	100(100)	100(100)	100(100)
$B(1,2)$	68	81	8	73	86	88	88	61	72	99	65	86	94	69	76(59)	74(62)	85(78)	70(66)	18(18)	83(62)
$EP(1.5)$	100	100	51	100	100	100	100	100	100	100	100	100	100	100	96(100)	99(100)	100(100)	100(100)	95(87)	99(100)
$BEX(2)$	92	95	29	98	92	95	94	97	98	65	97	94	99	98	97(99)	88(91)	98(97)	94(96)	92(0)	96(98)
DHR																				
$\Gamma(0.4)$	100	100	13	100	100	100	100	100	100	86	100	100	100	99	100(100)	98(100)	100(100)	98(100)	100(100)	100(100)
$\Gamma(0.7)$	48	55	9	71	44	54	56	63	23	17	61	55	57	18	64(70)	41(49)	68(62)	55(61)	59(64)	67(66)
$W(0.8)$	51	59	7	70	51	62	63	60	19	29	61	63	51	16	64(66)	46(49)	67(68)	50(59)	54(60)	65(68)
$EG(0.2)$	7	8	5	8	7	9	9	7	3	6	8	9	5	2	6(6)	8(6)	9(9)	6(7)	7(9)	8(8)
$EG(0.5)$	30	35	7	36	30	40	41	25	4	20	31	41	18	3	32(26)	30(23)	37(38)	21(28)	20(27)	34(34)
$EG(0.8)$	86	90	25	91	88	92	93	68	30	74	84	93	79	26	91(82)	84(77)	92(93)	77(83)	74(78)	90(90)
$EL(0.2)$	49	55	7	56	49	58	59	45	11	26	54	58	35	8	53(45)	43(41)	58(60)	39(48)	42(48)	54(57)
$EL(0.5)$	12	13	5	13	11	15	15	11	2	7	12	15	6	1	11(10)	13(10)	15(16)	9(12)	10(14)	14(14)
$EL(0.8)$	5	6	5	6	5	6	6	5	3	5	6	6	4	3	5(5)	6(5)	6(6)	5(6)	5(6)	6(6)
$ENH1(0.5)$	100	100	45	100	100	100	100	100	99	98	100	100	100	99	100(100)	99(100)	100(100)	99(100)	100(100)	100(100)
$BEX(0.4)$	100	100	13	100	100	100	100	100	99	82	100	100	100	99	100(100)	98(100)	100(100)	98(100)	100(100)	100(100)
$BEX(0.7)$	43	49	10	66	39	49	49	61	21	13	56	49	51	16	60(66)	38(46)	65(58)	52(59)	56(61)	65(62)
NMHR																				
$PW(2)$	55	71	85	54	48	50	6	40	40	85	44	8	98	25	77(74)	50(44)	84(34)	90(92)	95(96)	86(51)
$PW(3)$	99	100	82	100	98	99	84	100	99	55	100	90	100	98	100(100)	98(99)	100(99)	98(100)	100(100)	100(100)
$LN(0.8)$	91	93	59	81	85	81	60	97	99	27	91	56	100	98	100(94)	72(78)	100(84)	95(98)	96(0)	99(92)
$LN(1.5)$	98	99	68	99	99	99	99	70	44	96	97	99	97	36	98(94)	97(95)	99(99)	95(96)	93(94)	98(98)
$DH(1)$	66	73	32	74	62	63	55	80	89	25	79	54	91	87	91(85)	55(62)	85(71)	82(84)	82(0)	82(80)
$DH(1.5)$	100	100	58	100	100	100	100	100	100	88	100	100	100	100	100(100)	98(100)	100(100)	99(100)	98(1)	100(100)
$CH(0.5)$	99	99	12	100	98	99	99	100	96	68	100	99	100	95	100(100)	96(99)	100(100)	97(100)	100(100)	100(100)
HN	57	68	5	64	71	74	74	53	63	73	60	73	66	64	61(55)	51(57)	70(68)	53(54)	22(4)	68(56)
$B(0.5,1)$	54	70	86	53	47	49	5	42	39	85	42	6	98	26	78(75)	50(45)	84(34)	90(92)	95(96)	86(50)
$EP(0.5)$	88	91	23	98	83	88	86	97	80	23	95	86	96	75	98(99)	85(90)	98(94)	94(97)	96(97)	98(97)
$ENH1(2)$	22	26	26	10	22	26	17	7	20	30	6	19	46	9	29(7)	30(4)	21(16)	26(27)	19(1)	17(8)
$ENH2(0.4)$	99	99	30	100	97	98	97	100	98	40	100	97	100	96	100(100)	96(99)	100(100)	97(100)	100(100)	100(100)
$ENH2(0.7)$	14	14	20	20	9	9	7	24	6	4	18	7	26	3	25(29)	11(16)	31(13)	21(31)	32(38)	36(19)

Table A.5: Monte Carlo power estimates for $n = 75$.

	KS_n	CM_n	HM_n	CO_n	$\overline{KS}_n$	$\overline{CM}_n$	G_n	AH_n^1	AH_n^2	BR_n	ZP_n	EP_n	ZA_n	NA_n	$AT_{n,\widehat{\gamma}}$	$J_{n,\widehat{\gamma}}$	$BH_{n,\widehat{\gamma}}$	$PW_{n,\widehat{\gamma}}^1$	$PW_{n,\widehat{\gamma}}^2$	$L_{n,\widehat{\gamma}}$
CHR																				
$\Gamma(1)$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5(5)	5(5)	5(5)	5(5)	5(5)	5(5)
IHR																				
$\Gamma(1.5)$	68	76	14	86	71	76	76	79	84	42	82	77	84	86	81(**88**)	60(67)	82(82)	75(79)	71(0)	80(85)
$\Gamma(2)$	99	**100**	43	**100**	99	**100**	**100**	**100**	**100**	90	**100**	**100**	**100**	**100**	99(**100**)	96(99)	**100**(**100**)	98(**100**)	98(1)	**100**(**100**)
$W(1.2)$	43	52	6	61	49	54	56	52	59	35	55	57	56	**62**	54(60)	38(47)	56(59)	41(49)	36(0)	54(58)
$W(1.5)$	99	**100**	29	**100**	99	**100**	**100**	**100**	**100**	98	**100**	**100**	**100**	**100**	95(**100**)	96(99)	**100**(**100**)	98(**100**)	93(2)	98(**100**)
$PW(1)$	**100**	**100**	9	**100**	**100**	**100**	**100**	**100**	**100**	**100**	99	**100**	**100**	**100**	90(98)	**100**(99)	**100**(**100**)	**100**(**100**)	97(99)	96(**100**)
$CH(1)$	50	61	6	56	66	69	70	45	53	**75**	50	70	64	56	60(47)	47(49)	67(63)	47(48)	15(5)	66(51)
$CH(1.5)$	**100**	**100**	45	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	97(**100**)	99(**100**)	**100**(**100**)	**100**(**100**)	96(65)	**100**(**100**)
$LF(2)$	85	93	7	90	93	95	**96**	85	89	93	88	**96**	91	89	86(82)	80(85)	93(94)	84(87)	48(5)	92(88)
$LF(4)$	96	99	13	98	99	99	**100**	97	98	99	98	**100**	98	98	90(95)	94(96)	98(99)	96(97)	70(9)	96(97)
$EV(0.5)$	51	63	6	58	66	70	70	44	53	**76**	52	71	66	57	60(47)	48(49)	67(62)	46(47)	15(5)	65(50)
$EV(1.5)$	98	**100**	7	98	**100**	**100**	**100**	97	98	**100**	97	**100**	**100**	98	91(93)	96(96)	98(99)	98(98)	53(33)	96(97)
$B(2,1)$	**100**	**100**	99	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)
$B(1,2)$	80	92	9	83	94	96	95	74	82	**100**	76	95	99	80	88(71)	88(75)	93(90)	85(83)	25(26)	92(77)
$EP(1.5)$	**100**	**100**	65	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	97(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)	98(96)	**100**(**100**)
$BEX(2)$	97	99	38	**100**	97	98	98	99	**100**	74	**100**	98	**100**	**100**	98(**100**)	95(97)	**100**(**100**)	97(**100**)	97(0)	**100**(**100**)
DHR																				
$\Gamma(0.4)$	**100**	**100**	14	**100**	**100**	**100**	**100**	**100**	**100**	95	**100**	**100**	**100**	**100**	**100**(**100**)	99(**100**)	**100**(**100**)	99(**100**)	**100**(**100**)	**100**(**100**)
$\Gamma(0.7)$	59	65	11	81	55	64	66	75	42	22	71	65	68	39	77(**82**)	52(61)	78(73)	68(73)	70(74)	78(77)
$W(0.8)$	64	72	8	**82**	64	73	76	72	37	36	73	75	64	36	76(78)	55(63)	79(79)	64(72)	67(72)	77(80)
$EG(0.2)$	9	9	6	10	8	10	10	8	3	6	9	10	5	2	7(7)	9(8)	10(10)	7(8)	7(**11**)	9(9)
$EG(0.5)$	38	44	7	47	39	50	**52**	32	9	25	39	51	25	8	43(34)	37(32)	49(50)	29(36)	26(34)	46(45)
$EG(0.8)$	93	96	30	97	95	97	**98**	81	54	86	93	**98**	91	56	96(92)	90(89)	97(97)	88(92)	85(88)	96(96)
$EL(0.2)$	60	66	7	69	60	70	71	54	22	34	65	71	44	23	67(58)	54(55)	70(**73**)	51(61)	54(59)	68(70)
$EL(0.5)$	14	16	5	17	14	18	**19**	13	3	9	15	**19**	8	2	14(12)	15(12)	18(**19**)	10(14)	12(17)	17(18)
$EL(0.8)$	6	6	5	6	6	**7**	**7**	5	3	5	6	**7**	4	3	5(5)	6(5)	6(6)	6(6)	5(**7**)	6(6)
$ENH1(0.5)$	**100**	**100**	54	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**(**100**)	99(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)	**100**(**100**)
$BEX(0.4)$	**100**	**100**	14	**100**	**100**	**100**	**100**	**100**	**100**	93	**100**	**100**	**100**	**100**	**100**(**100**)	99(**100**)	**100**(**100**)	99(**100**)	**100**(**100**)	**100**(**100**)
$BEX(0.7)$	55	62	11	**79**	51	60	62	72	39	18	69	62	64	36	74(**79**)	47(57)	75(68)	65(71)	69(73)	76(74)
NMHR																				
$PW(2)$	69	85	94	65	65	69	6	49	66	97	53	8	**100**	46	88(84)	68(51)	92(46)	95(98)	99(99)	92(63)
$PW(3)$	**100**	**100**	91	**100**	**100**	**100**	93	**100**	**100**	85	**100**	96	**100**	**100**	**100**(**100**)	98(**100**)	**100**(**100**)	99(**100**)	**100**(**100**)	**100**(**100**)
$LN(0.8)$	98	99	73	88	95	92	70	**100**	**100**	32	97	67	**100**	**100**	**100**(98)	81(86)	**100**(94)	97(**100**)	98(0)	**100**(98)
$LN(1.5)$	**100**	**100**	80	**100**	**100**	**100**	**100**	81	70	99	99	**100**	99	67	99(98)	98(99)	**100**(**100**)	99(99)	97(98)	99(**100**)
$DH(1)$	79	84	43	83	75	75	64	92	**96**	27	88	63	**96**	95	96(93)	66(71)	95(83)	92(94)	93(0)	94(91)
$DH(1.5)$	**100**	**100**	73	**100**	**100**	**100**	**100**	**100**	**100**	95	**100**	**100**	**100**	**100**	**100**(**100**)	98(**100**)	**100**(**100**)	**100**(**100**)	99(1)	**100**(**100**)
$CH(0.5)$	**100**	**100**	13	**100**	**100**	**100**	**100**	**100**	99	84	**100**	**100**	**100**	**100**	**100**(**100**)	98(**100**)	**100**(**100**)	98(**100**)	**100**(**100**)	**100**(**100**)
HN	70	80	5	76	82	85	**86**	67	74	85	72	**86**	79	76	77(67)	65(68)	85(83)	68(70)	28(5)	83(72)
$B(0.5,1)$	67	84	94	63	63	67	5	48	66	97	51	7	**100**	45	87(83)	68(51)	92(44)	95(98)	99(99)	92(63)
$EP(0.5)$	96	97	29	**100**	93	96	94	99	94	32	98	95	99	93	99(**100**)	92(96)	**100**(98)	97(99)	99(99)	**100**(99)
$ENH1(2)$	27	33	32	11	28	32	21	9	26	39	6	22	**60**	12	37(7)	36(4)	28(19)	34(37)	24(1)	24(9)
$ENH2(0.4)$	**100**	**100**	38	**100**	**100**	**100**	**100**	**100**	**100**	55	**100**	**100**	**100**	**100**	**100**(**100**)	98(**100**)	**100**(**100**)	98(**100**)	**100**(**100**)	**100**(**100**)
$ENH2(0.7)$	17	19	26	27	11	11	8	30	12	4	23	8	35	8	33(35)	13(18)	39(17)	29(39)	42(**49**)	45(26)

Table A.6: Monte Carlo power estimates for $n = 100$.

Appendix B

Closed-form expression derivations

Most articles in the literature only provide the theoretical form of test statistics and, immediately thereafter, the closed-form expressions are stated. This appendix serves to provide the intermediate steps between the theoretical forms and the computational (or closed-form) expressions.

B.1 Henze and Meintanis (2005) test ($S_{n,\gamma}$)

We start by noting that the ECF of the scaled data $Y_1, Y_2, ..., Y_n$ can be written as follows:

$$
\begin{aligned}
\phi_n(t) &= \frac{1}{n}\sum_{j=1}^{n} e^{itY_j} \\
&= \frac{1}{n}\sum_{j=1}^{n}\left[\cos\left(tY_j\right) + i\sin\left(tY_j\right)\right] \\
&= \frac{1}{n}\sum_{j=1}^{n}\cos\left(tY_j\right) + \frac{i}{n}\sum_{j=1}^{n}\sin\left(tY_j\right).
\end{aligned}
$$

The squared modulus of the ECF, $\left|\phi_n(t)\right|^2$ can be written as

$$
\begin{aligned}
\left|\phi_n(t)\right|^2 &= \left|\frac{1}{n}\sum_{j=1}^{n}\cos\left(tY_j\right) + \frac{i}{n}\sum_{j=1}^{n}\sin\left(tY_j\right)\right|^2 \\
&= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\cos\left(tY_j\right)\cos\left(tY_k\right) + \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\sin\left(tY_j\right)\sin\left(tY_k\right) \\
&= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\left[\cos\left(tY_j\right)\cos\left(tY_k\right) + \sin\left(tY_j\right)\sin\left(tY_k\right)\right] \\
&= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\cos\left(tY_j - tY_k\right) \\
&= \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\cos\left(tY_{jk-}\right), \qquad \text{where } Y_{jk-} = Y_j - Y_k.
\end{aligned}
$$

$C_n(t)$, the real part of $\phi_n(t)$, is given by

$$C_n(t) = \frac{1}{n}\sum_{j=1}^{n} \cos\Big(tY_j\Big).$$

Note that the test statistic proposed by Henze and Meintanis (2005) can be written as

$$\begin{aligned}
S_{n,\gamma} &= n\int_0^\infty \left[\Big|\phi_n(t)\Big|^2 - C_n(t)\right]^2 e^{-t^2\gamma}dt \\
&= n\int_0^\infty \left[\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\cos\Big(tY_{jk-}\Big) - \frac{1}{n}\sum_{j=1}^{n}\cos\Big(tY_j\Big)\right]^2 e^{-t^2\gamma}dt \\
&= n\int_0^\infty \left[\frac{1}{n^4}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n}\cos\Big(tY_{jk-}\Big)\cos\Big(tY_{lm-}\Big)\right. \\
&\qquad - \frac{2}{n^3}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\cos\Big(tY_{jk-}\Big)\cos\Big(tY_l\Big) \\
&\qquad\qquad \left. + \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\cos\Big(tY_j\Big)\cos\Big(tY_k\Big)\right] e^{-t^2\gamma}dt \\
&= \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\int_0^\infty \cos\Big(tY_j\Big)\cos\Big(tY_k\Big)e^{-t^2\gamma}dt \\
&\qquad - \frac{2}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\int_0^\infty \cos\Big(tY_{jk-}\Big)\cos\Big(tY_l\Big)e^{-t^2\gamma}dt \\
&\qquad\qquad + \frac{1}{n^3}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n}\int_0^\infty \cos\Big(tY_{jk-}\Big)\cos\Big(tY_{lm-}\Big)e^{-t^2\gamma}dt \\
&=: \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathcal{I}_1 - \frac{2}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\mathcal{I}_2 + \frac{1}{n^3}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n}\mathcal{I}_3.
\end{aligned}$$

The above integrals can be simplified as follows:

$$\begin{aligned}
\mathcal{I}_1 &= \int_0^\infty \cos(tY_j)\cos(tY_k)e^{-t^2\gamma}dt \\
&= \frac{1}{2}\int_0^\infty \left[\cos\Big(tY_j + tY_k\Big) + \cos\Big(tY_j - tY_k\Big)\right]e^{-t^2\gamma}dt \\
&= \frac{1}{2}\int_0^\infty \left[\cos\Big(tY_{jk+}\Big) + \cos\Big(tY_{jk-}\Big)\right]e^{-t^2\gamma}dt, \qquad \text{where } Y_{jk+} = Y_j + Y_k \\
&= \frac{1}{2}\int_0^\infty \cos\Big(tY_{jk+}\Big)e^{-t^2\gamma}dt + \frac{1}{2}\int_0^\infty \cos\Big(tY_{jk-}\Big)e^{-t^2\gamma}dt \\
&= \frac{1}{2}\left[\frac{1}{2}\sqrt{\frac{\pi}{\gamma}}\exp\Big(-\frac{Y_{jk+}^2}{4\gamma}\Big)\right] + \frac{1}{2}\left[\frac{1}{2}\sqrt{\frac{\pi}{\gamma}}\exp\Big(-\frac{Y_{jk-}^2}{4\gamma}\Big)\right] \\
&= \frac{1}{4}\sqrt{\frac{\pi}{\gamma}}\left[\exp\Big(-\frac{Y_{jk+}^2}{4\gamma}\Big) + \exp\Big(-\frac{Y_{jk-}^2}{4\gamma}\Big)\right].
\end{aligned}$$

$$\begin{aligned}
\mathcal{I}_2 &= \int_0^\infty \cos\Big(tY_{jk-}\Big)\cos\Big(tY_l\Big)e^{-t^2\gamma}dt \\
&= \frac{1}{2}\int_0^\infty \left[\cos\Big(tY_{jk-} + tY_l\Big) + \cos\Big(tY_{jk-} - tY_l\Big)\right]e^{-t^2\gamma}dt \\
&= \frac{1}{2}\int_0^\infty \cos\Big(t\big(Y_{jk-} + Y_l\big)\Big)e^{-t^2\gamma}dt + \frac{1}{2}\int_0^\infty \cos\Big(t\big(Y_{jk-} - Y_l\big)\Big)e^{-t^2\gamma}dt \\
&= \frac{1}{2}\left[\frac{1}{2}\sqrt{\frac{\pi}{\gamma}}\exp\Big(-\frac{(Y_{jk-} + Y_l)^2}{4\gamma}\Big)\right] + \frac{1}{2}\left[\frac{1}{2}\sqrt{\frac{\pi}{\gamma}}\exp\Big(-\frac{(Y_{jk-} - Y_l)^2}{4\gamma}\Big)\right] \\
&= \frac{1}{4}\sqrt{\frac{\pi}{\gamma}}\left[\exp\Big(-\frac{(Y_{jk-} + Y_l)^2}{4\gamma}\Big) + \exp\Big(-\frac{(Y_{jk-} - Y_l)^2}{4\gamma}\Big)\right].
\end{aligned}$$

$$\begin{aligned}
\mathcal{I}_3 &= \int_0^\infty \cos\Big(tY_{jk-}\Big)\cos\Big(tY_{lm-}\Big)e^{-t^2\gamma}dt \\
&= \frac{1}{2}\int_0^\infty \left[\cos\Big(tY_{jk-} + tY_{lm-}\Big) + \cos\Big(tY_{jk-} - tY_{lm-}\Big)\right]e^{-t^2\gamma}dt \\
&= \frac{1}{2}\int_0^\infty \cos\Big(t\big(Y_{jk-} + Y_{lm-}\big)\Big)e^{-t^2\gamma}dt + \frac{1}{2}\int_0^\infty \cos\Big(t\big(Y_{jk-} - Y_{lm-}\big)\Big)e^{-t^2\gamma}dt \\
&= \frac{1}{2}\left[\frac{1}{2}\sqrt{\frac{\pi}{\gamma}}\exp\Big(-\frac{(Y_{jk-} + Y_{lm-})^2}{4\gamma}\Big)\right] + \frac{1}{2}\left[\frac{1}{2}\sqrt{\frac{\pi}{\gamma}}\exp\Big(-\frac{(Y_{jk-} - Y_{lm-})^2}{4\gamma}\Big)\right] \\
&= \frac{1}{4}\sqrt{\frac{\pi}{\gamma}}\left[\exp\Big(-\frac{(Y_{jk-} + Y_{lm-})^2}{4\gamma}\Big) + \exp\Big(-\frac{(Y_{jk-} - Y_{lm-})^2}{4\gamma}\Big)\right].
\end{aligned}$$

Substituting back into $S_{n,\gamma}$ we have that

$$\begin{aligned}S_{n,\gamma} &= \frac{1}{4n}\sqrt{\frac{\pi}{\gamma}}\sum_{j=1}^{n}\sum_{k=1}^{n}\left[\exp\Big(-\frac{Y_{jk+}^2}{4\gamma}\Big)+\exp\Big(-\frac{Y_{jk-}^2}{4\gamma}\Big)\right]\\ &\quad -\frac{1}{2n^2}\sqrt{\frac{\pi}{\gamma}}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\left[\exp\Big(-\frac{(Y_{jk-}-Y_l)^2}{4\gamma}\Big)+\exp\Big(-\frac{(Y_{jk-}+Y_l)^2}{4\gamma}\Big)\right]\\ &\quad +\frac{1}{4n^3}\sqrt{\frac{\pi}{\gamma}}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n}\left[\exp\Big(-\frac{(Y_{jk-}+Y_{lm-})^2}{4\gamma}\Big)+\exp\Big(-\frac{(Y_{jk-}-Y_{lm-})^2}{4\gamma}\Big)\right].\end{aligned}$$

B.2 Baringhaus and Henze (1991) test ($BH_{n,\gamma}$)

Note that the test statistic proposed by Baringhaus and Henze (1991) can be written as

$$\begin{aligned}BH_{n,\gamma} &= n\int_0^\infty\left[(1+t)\psi_n'(t)+\psi_n(t)\right]^2 e^{-t\gamma}dt\\ &= n\int_0^\infty\left[(1+t)\Big(-\frac{1}{n}\sum_{j=1}^{n}Y_je^{-tY_j}\Big)+\frac{1}{n}\sum_{j=1}^{n}e^{-tY_j}\right]^2 e^{-t\gamma}dt\\ &= n\int_0^\infty\left[\frac{1}{n^2}(1+t)^2\sum_{j=1}^{n}\sum_{k=1}^{n}Y_jY_ke^{-t(Y_j+Y_k+\gamma)}-\frac{1}{n^2}(1+t)\sum_{j=1}^{n}\sum_{k=1}^{n}Y_je^{-t(Y_j+Y_k+\gamma)}\right.\\ &\qquad\left.-\frac{1}{n^2}(1+t)\sum_{j=1}^{n}\sum_{k=1}^{n}Y_ke^{-t(Y_j+Y_k+\gamma)}+\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}e^{-t(Y_j+Y_k+\gamma)}\right]dt\\ &= \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\left[\int_0^\infty(1+t)^2Y_jY_ke^{-t(Y_j+Y_k+\gamma)}dt-\int_0^\infty(1+t)Y_je^{-t(Y_j+Y_k+\gamma)}dt\right.\\ &\qquad\left.-\int_0^\infty(1+t)Y_ke^{-t(Y_j+Y_k+\gamma)}dt+\int_0^\infty e^{-t(Y_j+Y_k+\gamma)}dt\right]\\ &=: \frac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\Big[\mathcal{I}_1-\mathcal{I}_2-\mathcal{I}_3+\mathcal{I}_4\Big].\end{aligned}$$

The above integrals can be simplified as follows:

$$\begin{aligned}\mathcal{I}_1 &= \int_0^\infty(1+t)^2Y_jY_ke^{-t(Y_j+Y_k+\gamma)}dt\\ &= \int_0^\infty Y_jY_ke^{-t(Y_j+Y_k+\gamma)}dt+2\int_0^\infty tY_jY_ke^{-t(Y_j+Y_k+\gamma)}dt+\int_0^\infty t^2Y_jY_ke^{-t(Y_j+Y_k+\gamma)}dt\\ &= \left[-\frac{Y_jY_ke^{-t(Y_j+Y_k+\gamma)}}{Y_j+Y_k+\gamma}\right]_0^\infty+2\left[\frac{Y_jY_k\Gamma(2)}{(Y_j+Y_k+\gamma)^2}\right]+\left[\frac{Y_jY_k\Gamma(3)}{(Y_j+Y_k+\gamma)^3}\right]\\ &= \frac{Y_jY_k}{Y_j+Y_k+\gamma}+\frac{2Y_jY_k}{(Y_j+Y_k+\gamma)^2}+\frac{2Y_jY_k}{(Y_j+Y_k+\gamma)^3}.\end{aligned}$$

$$
\begin{aligned}
\mathcal{I}_2 &= \int_0^\infty (1+t)Y_j e^{-t(Y_j+Y_k+\gamma)}dt \\
&= \int_0^\infty Y_j e^{-t(Y_j+Y_k+\gamma)}dt + \int_0^\infty tY_j e^{-t(Y_j+Y_k+\gamma)}dt \\
&= \left[-\frac{Y_j e^{-t(Y_j+Y_k+\gamma)}}{Y_j+Y_k+\gamma} \right]_0^\infty + \left[\frac{Y_j\Gamma(2)}{(Y_j+Y_k+\gamma)^2} \right] \\
&= \frac{Y_j}{Y_j+Y_k+\gamma} + \frac{Y_j}{(Y_j+Y_k+\gamma)^2}.
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{I}_3 &= \int_0^\infty (1+t)Y_k e^{-t(Y_j+Y_k+\gamma)}dt \\
&= \int_0^\infty Y_k e^{-t(Y_j+Y_k+\gamma)}dt + \int_0^\infty tY_k e^{-t(Y_j+Y_k+\gamma)}dt \\
&= \left[-\frac{Y_k e^{-t(Y_j+Y_k+\gamma)}}{Y_j+Y_k+\gamma} \right]_0^\infty + \left[\frac{Y_k\Gamma(2)}{(Y_j+Y_k+\gamma)^2} \right] \\
&= \frac{Y_k}{Y_j+Y_k+\gamma} + \frac{Y_k}{(Y_j+Y_k+\gamma)^2}.
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{I}_4 &= \int_0^\infty e^{-t(Y_j+Y_k+\gamma)}dt \\
&= \left[-\frac{e^{-t(Y_j+Y_k+\gamma)}}{Y_j+Y_k+\gamma} \right]_0^\infty \\
&= \frac{1}{Y_j+Y_k+\gamma}.
\end{aligned}
$$

Substituting back into $BH_{n,\gamma}$ we have that

$$
\begin{aligned}
BH_{n,\gamma} = \frac{1}{n}\sum_{j=1}^n\sum_{k=1}^n \Bigg[& \frac{Y_jY_k}{Y_j+Y_k+\gamma} + \frac{2Y_jY_k}{(Y_j+Y_k+\gamma)^2} + \frac{2Y_jY_k}{(Y_j+Y_k+\gamma)^3} - \frac{Y_j}{Y_j+Y_k+\gamma} \\
& - \frac{Y_j}{(Y_j+Y_k+\gamma)^2} - \frac{Y_k}{Y_j+Y_k+\gamma} - \frac{Y_k}{(Y_j+Y_k+\gamma)^2} + \frac{1}{Y_j+Y_k+\gamma} \Bigg] \\
= \frac{1}{n}\sum_{j=1}^n\sum_{k=1}^n \Bigg[& \frac{1-Y_j-Y_k+Y_jY_k}{Y_j+Y_k+\gamma} - \frac{Y_j+Y_k-2Y_jY_k}{(Y_j+Y_k+\gamma)^2} + \frac{2Y_jY_k}{(Y_j+Y_k+\gamma)^3} \Bigg] \\
= \frac{1}{n}\sum_{j=1}^n\sum_{k=1}^n \Bigg[& \frac{(1-Y_j)(1-Y_k)}{Y_j+Y_k+\gamma} - \frac{Y_j+Y_k}{(Y_j+Y_k+\gamma)^2} + \frac{2Y_jY_k}{(Y_j+Y_k+\gamma)^2} + \frac{2Y_jY_k}{(Y_j+Y_k+\gamma)^3} \Bigg].
\end{aligned}
$$

B.3 Henze and Meintanis (2002) test ($L_{n,\gamma}$)

The test statistic proposed by Henze and Meintanis (2002) can be written as

$$\begin{aligned}
L_{n,\gamma} &= n\int_0^\infty \left[\psi_n(t) - \frac{1}{1+t}\right]^2 (1+t)^2 e^{-t\gamma} dt \\
&= n\int_0^\infty \left[\frac{1}{n}\sum_{j=1}^n e^{-tY_j} - \frac{1}{1+t}\right]^2 (1+t)^2 e^{-t\gamma} dt \\
&= n\int_0^\infty \left[(1+t)^2 \frac{1}{n^2}\sum_{j=1}^n\sum_{k=1}^n e^{-t(Y_j+Y_k+\gamma)} - 2(1+t)\frac{1}{n}\sum_{j=1}^n e^{-t(Y_j+\gamma)} + e^{-t\gamma}\right] dt \\
&= \frac{1}{n}\sum_{j=1}^n\sum_{k=1}^n \int_0^\infty (1+t)^2 e^{-t(Y_j+Y_k+\gamma)} dt - 2\sum_{j=1}^n \int_0^\infty (1+t) e^{-t(Y_j+\gamma)} dt + n\int_0^\infty e^{-t\gamma} dt \\
&=: \frac{1}{n}\sum_{j=1}^n\sum_{k=1}^n \mathcal{I}_1 - 2\sum_{j=1}^n \mathcal{I}_2 + n\mathcal{I}_3.
\end{aligned}$$

The above integrals can be simplified as follows:

$$\begin{aligned}
\mathcal{I}_1 &= \int_0^\infty (1+t)^2 e^{-t(Y_j+Y_k+\gamma)} dt \\
&= \int_0^\infty e^{-t(Y_j+Y_k+\gamma)} dt + 2\int_0^\infty t e^{-t(Y_j+Y_k+\gamma)} dt + \int_0^\infty t^2 e^{-t(Y_j+Y_k+\gamma)} dt \\
&= \left[-\frac{e^{-t(Y_j+Y_k+\gamma)}}{Y_j+Y_k+\gamma}\right]_0^\infty + 2\left[\frac{\Gamma(2)}{(Y_j+Y_k+\gamma)^2}\right] + \left[\frac{\Gamma(3)}{(Y_j+Y_k+\gamma)^3}\right] \\
&= \frac{1}{Y_j+Y_k+\gamma} + \frac{2}{(Y_j+Y_k+\gamma)^2} + \frac{2}{(Y_j+Y_k+\gamma)^3} \\
&= \frac{(Y_j+Y_k+\gamma)^2 + 2(Y_j+Y_k+\gamma) + 2}{(Y_j+Y_k+\gamma)^3} \\
&= \frac{Y_j^2 + Y_k^2 + \gamma^2 + 2Y_jY_k + 2Y_j\gamma + 2Y_k\gamma + 2Y_j + 2Y_k + 2\gamma + 2}{(Y_j+Y_k+\gamma)^3} \\
&= \frac{1 + (Y_j^2 + Y_k^2 + \gamma^2 + 1 + 2Y_jY_k + 2Y_j\gamma + 2Y_k\gamma + 2Y_j + 2Y_k + 2\gamma)}{(Y_j+Y_k+\gamma)^3} \\
&= \frac{1 + (Y_j+Y_k+\gamma+1)^2}{(Y_j+Y_k+\gamma)^3}.
\end{aligned}$$

$$
\begin{aligned}
\mathcal{I}_2 &= \int_0^\infty (1+t)e^{-t(Y_j+\gamma)}dt \\
&= \int_0^\infty (e^{-t(Y_j+\gamma)}dt + \int_0^\infty te^{-t(Y_j+\gamma)}dt \\
&= \left[-\frac{e^{-t(Y_j+\gamma)}}{Y_j+\gamma}\right]_0^\infty + \left[\frac{\Gamma(2)}{(Y_j+\gamma)^2}\right] \\
&= \frac{1}{Y_j+\gamma} + \frac{1}{(Y_j+\gamma)^2} \\
&= \frac{1+Y_j+\gamma}{(Y_j+\gamma)^2}.
\end{aligned}
$$

$$
\begin{aligned}
\mathcal{I}_3 &= \int_0^\infty e^{-t\gamma}dt \\
&= \left[-\frac{e^{-t\gamma}}{\gamma}\right]_0^\infty \\
&= \frac{1}{\gamma}.
\end{aligned}
$$

Substituting back into $L_{n,\gamma}$ we have that

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

B.4 Cramér-von Mises test (CM_n)

We prove this for the uniform distribution case by noting that the test statistic can be written as follows:

$$
\begin{aligned}
CM_n &= n \int_0^1 \left[F_n(x) - x \right]^2 dx \\
&= n \sum_{j=0}^{n} \int_{X_{(j)}}^{X_{(j+1)}} \left[\frac{j}{n} - x \right]^2 dx \qquad \Big(X_{(0)} = 0, X_{(n+1)} = 1 \Big) \\
&= \frac{n}{3} \left[\sum_{j=0}^{n} \left(x - \frac{j}{n} \right)^3 \right]_{X_{(j)}}^{X_{(j+1)}} \\
&= \frac{n}{3} \left[\sum_{j=0}^{n} \left(x^3 - \frac{3jx^2}{n} + \frac{3j^2x}{n^2} - \frac{j^3}{n^3} \right) \right]_{X_{(j)}}^{X_{(j+1)}} \\
&= \frac{n}{3} \sum_{j=0}^{n} \Bigg[X_{(j+1)}^3 - 3\left(\frac{j}{n}\right) X_{(j+1)}^2 + 3\left(\frac{j}{n}\right)^2 X_{(j+1)} - \left(\frac{j}{n}\right)^3 - X_{(j)}^3 + 3\left(\frac{j}{n}\right) X_{(j)}^2 \\
&\qquad - 3\left(\frac{j}{n}\right)^2 X_{(j)} + \left(\frac{j}{n}\right)^3 \Bigg] \\
&= \frac{n}{3} \sum_{j=0}^{n} \Bigg[- X_{(j)}^3 + 3\left(\frac{j}{n}\right) X_{(j)}^2 - 3\left(\frac{j}{n}\right)^2 X_{(j)} + X_{(j+1)}^3 - 3\left(\frac{j}{n}\right) X_{(j+1)}^2 + 3\left(\frac{j}{n}\right)^2 X_{(j+1)} \Bigg] \\
&= \frac{n}{3} \Bigg[- X_{(0)}^3 + 3\left(\frac{0}{n}\right) X_{(0)}^2 - 3\left(\frac{0}{n}\right)^2 X_{(0)} + X_{(1)}^3 - 3\left(\frac{0}{n}\right) X_{(1)}^2 + 3\left(\frac{0}{n}\right)^2 X_{(1)} \\
&\qquad - X_{(1)}^3 + 3\left(\frac{1}{n}\right) X_{(1)}^2 - 3\left(\frac{1}{n}\right)^2 X_{(1)} + X_{(2)}^3 - 3\left(\frac{1}{n}\right) X_{(2)}^2 + 3\left(\frac{1}{n}\right)^2 X_{(2)} \\
&\qquad - X_{(2)}^3 + 3\left(\frac{2}{n}\right) X_{(2)}^2 - 3\left(\frac{2}{n}\right)^2 X_{(2)} + X_{(3)}^3 - 3\left(\frac{2}{n}\right) X_{(3)}^2 + 3\left(\frac{2}{n}\right)^2 X_{(3)} \cdots \\
&\qquad \cdots - X_{(n)}^3 + 3\left(\frac{n}{n}\right) X_{(n)}^2 - 3\left(\frac{n}{n}\right)^2 X_{(n)} + X_{(n+1)}^3 - 3\left(\frac{n}{n}\right) X_{(n+1)}^2 + 3\left(\frac{n}{n}\right)^2 X_{(n+1)} \Bigg] \\
&= \frac{n}{3} \sum_{j=1}^{n} \left[\frac{3}{n} X_{(j)}^2 - \frac{3(2j-1)}{n^2} X_{(j)} \right] + \frac{n}{3} \Big[1 - 3 + 3 \Big] \\
&= \sum_{j=1}^{n} \left[X_{(j)}^2 - \frac{2j-1}{n} X_{(j)} \right] + \frac{n}{3} \\
&= \sum_{j=1}^{n} \left[X_{(j)} - \frac{2j-1}{2n} \right]^2 - \sum_{j=1}^{n} \left[\frac{2j-1}{2n} \right]^2 + \frac{n}{3}.
\end{aligned}
$$

But one can show that

$$\begin{aligned}\sum_{j=1}^{n}\left[\frac{2j-1}{2n}\right]^2 &= \frac{1}{4n^2}\sum_{j=1}^{n}\left[4j^2-4j+1\right]\\ &= \frac{1}{4n^2}\left[4\sum_{j=1}^{n}j^2-4\sum_{j=1}^{n}j+\sum_{j=1}^{n}1\right]\\ &= \frac{1}{4n^2}\left[4\left(\frac{n^3}{3}+\frac{n^2}{2}+\frac{n}{6}\right)-4\left(\frac{n^2}{2}+\frac{n}{2}\right)+n\right]\\ &= \frac{1}{4n^2}\left[\frac{4n^3}{3}+\frac{n}{3}\right]\\ &= \frac{n}{3}-\frac{1}{12n}\\ &= \frac{4n^2-1}{12n}.\end{aligned}$$

Substituting back into CM_n we have that

$$\begin{aligned}CM_n &= \sum_{j=1}^{n}\left[X_{(j)}-\frac{2j-1}{2n}\right]^2-\frac{4n^2-1}{12n}+\frac{n}{3}\\ &= \sum_{j=1}^{n}\left[X_{(j)}-\frac{2j-1}{2n}\right]^2+\frac{4n^2-4n^2+1}{12n}\\ &= \frac{1}{12n}+\sum_{j=1}^{n}\left[X_{(j)}-\frac{2j-1}{2n}\right]^2.\end{aligned}$$

From this, we can conclude that for the exponential distribution we find that

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