

The use of Mean Residual Life in testing departures from Exponentiality

S.Rao Jammalamadaka

Department of Statistics and Applied Probability
University of California
Santa Barbara, CA 93106-3110

Emanuele Taufer

Department of Computer and Management Sciences
University of Trento
Trento, I 38100.

Summary. We utilize the important characterization that $E(X - t|X > t)$ is a constant for $t \in [0, \infty)$ if and only if X is distributed as an exponential random variable, in order to construct a new test procedure for exponentiality. We discuss asymptotic distribution theory and other properties of the proposed procedure. Simulation studies indicate that the proposed statistic has very good power in a large variety of situations.

Keywords Kolmogorov-Smirnov statistic, Mean Residual Life, Quantile process, Test for exponentiality, Wiener process.

1 Introduction

Mean residual life (MRL) is a very well-known and central concept in reliability and survival analysis; if X denotes a non-negative random variable (r.v.) with distribution function (d.f.)

F , then the MRL at time t is defined as

$$m(t) = E(X - t | X > t) = \frac{\int_t^\infty \bar{F}(x) dx}{\bar{F}(t)} \quad (1.1)$$

where $\bar{F} = 1 - F$ is the survival function. If $m(t)$ is non-increasing (or non-decreasing) in t , then X is said to have a decreasing (or increasing) MRL distribution (DMRL or IMRL respectively). MRL is related to the hazard rate $\lambda(t)$ by the expression

$$\lambda(t) = \frac{1 + m'(t)}{m(t)} \quad (1.2)$$

and it can be shown that the class of DMRL (IMRL) distributions includes the class of increasing (decreasing) failure rate distributions (IFR and DFR respectively). For a review of these concepts and other properties of MRL, we refer the reader to Patel (1983) and Guess and Proschan (1988).

The exponential distribution is a natural boundary between DMRL and IMRL distributions, having a constant mean residual life. Indeed, it has been shown by Shanbhag (1970) that the exponential distribution can be characterized by constancy of MRL, more precisely

$$m(t) = \theta, \quad \forall t > 0 \quad (1.3)$$

if and only if F is exponential with mean θ . The above characterization can be easily shown to be equivalent to either

$$\int_t^\infty \bar{F}(x) dx = \bar{F}(t) \theta, \quad \forall t > 0 \quad (1.4)$$

or

$$E[\min(X, t)] = F(t) \theta \quad \forall t > 0. \quad (1.5)$$

Characterizations based on MRL have been used to build tests for exponentiality against specialized as well as omnibus alternatives. Hollander and Proschan (1972), Bergman and Klefsjö (1989) and Bandyopadhyay and Basu (1990) provide tests against DMRL distributions. Koul (1978) and Bhattacharjee and Sen (1995) provide tests for the larger class of New Better than Used in Expectation (NBUE) alternatives (recall that a distribution is NBUE if the equal sign in (1.4) is replaced by a $<$). More recently Baringhaus and Henze (2000) and Taufer (2000) provide omnibus tests for exponentiality. All these papers exploit equations (1.4) and (1.5) - in what follows, we utilize equation (1.3) directly in order to provide new tests for exponentiality. This approach is quite different from that used in the above mentioned papers and, as we will see, leads to some powerful alternative tests for exponentiality.

Note that testing for exponentiality still attracts considerable attention and is the topic of a good amount of recent research; besides the above mentioned contributions, other authors provide test statistics for detecting departures from the hypothesis of exponentiality against specific or general alternatives. Alwasel (2001), Ahmad and Alwasel (1999) use the lack of memory property of the exponential distribution. Klar (2001) exploits the integrated distribution function while Jammalamadaka and Taufer (2003) consider a characterization based on normalized spacings. Grzegorzewski and Wieczorkowski (1999) and Ebrahimi et Al. (1992) make use of the maximum entropy principle. Other omnibus tests for exponentiality have been developed by Henze and Meintanis (2002), Henze (1993), Baringhaus and Henze

(1991, 1992) who use estimators of the Laplace transform.

Specialized tests are provided by Del Castillo and Puig (1999a, 1999b), Gatto and Jammalamadaka (2002), Klar (2003) and Klar (2000) who extends the work of Jammalamadaka and Lee (1998). For a review of earlier contributions, the interested reader is referred to Shapiro (1995), Ascher (1990), and Doksum and Yandell (1984).

2 The test Statistic and its properties

2.1 Construction of the test statistic

Let X_1, \dots, X_{n+1} be a random sample from a distribution F with order statistics, $X_{(1)} \leq \dots \leq X_{(n+1)}$ and suppose we wish to test the hypothesis

$$H_0 : F(x) = 1 - e^{-x/\theta}, \theta > 0 \quad \text{versus} \quad H_1 : F(x) \neq 1 - e^{-x/\theta}, \theta > 0.$$

In order to exploit the characterization $m(t) = \theta$, under exponentiality, define the “sample mean residual life after $X_{(k)}$ ” as

$$\begin{aligned} \bar{X}_{>k} &= \frac{1}{n-k+1} \sum_{i=k+1}^{n+1} (X_{(i)} - X_{(k)}) \\ &= \frac{1}{n-k+1} \sum_{i=k+1}^{n+1} (n-i+2)(X_{(i)} - X_{(i-1)}). \end{aligned} \tag{2.6}$$

For convenience, we define

$$Y_i = (n-i+2)(X_{(i)} - X_{(i-1)}), \quad i = 1, \dots, n+1 \tag{2.7}$$

as the “normalized spacings”. Observe that under the null hypothesis of exponentiality, we have that

$$E(\bar{X}_{>k}) = E(\bar{X}) = \theta, \quad k = 1, \dots, n. \quad (2.8)$$

Therefore if we plot the sequence $\bar{X}, \bar{X}_{>1}, \dots, \bar{X}_{>n}$ on a chart, under H_0 this should be approximately constant around the true (unknown) value θ . This intuitive graphical approach would suggest to use a distance measure between these “residual sample means” in order to build a test statistic for H_0 . One simple and natural way to do this is to exploit a Kolmogorov-Smirnov type distance, that is, reject H_0 when

$$T'_n = \max_{1 \leq k \leq n} \frac{|\bar{X} - \bar{X}_{>k}|}{\bar{X}} \quad (2.9)$$

is large. Note that division by the sample mean makes T'_n scale-free. However unfortunately it turns out that, as it is, T'_n does not converge to zero even under the null hypothesis; this may be immediately seen if we note that in particular, $\bar{X}_{>n} = Y_{n+1}$ which is exponentially distributed with mean θ under the null hypothesis, no matter what the sample size is. To consider the behavior of T'_n a bit more carefully, let $S(i) = \sum_{j \leq i} \xi_j$ where ξ_j are *i.i.d.* exponential r.v.'s with mean 1 and let $i = n - k + 1$ then we obtain

$$\begin{aligned} T'_n &= \max_{1 \leq i \leq n} \left| 1 - \frac{\bar{X}_{>n-i+1}}{\bar{X}} \right| \\ &\stackrel{D}{=} \max_{1 \leq i \leq n} \left| 1 - \frac{S(i)}{i} \frac{(n+1)}{S(n+1)} \right| \\ &\leq \max_{1 \leq i \leq n} \left| 1 - \frac{S(i)}{i} \right| + \max_{1 \leq i \leq n} \frac{S(i)}{i} \left| 1 - \frac{(n+1)}{S(n+1)} \right| \\ &\rightarrow \max_{1 \leq i < \infty} \left| 1 - \frac{S(i)}{i} \right| + o_p(1) \quad \text{as } n \rightarrow \infty. \end{aligned} \quad (2.10)$$

Thus even though the statistic T'_n is built using differences which are close to 0 under the hypothesis, the *largest* difference does not converge to 0 even under exponentiality. Thus we need to address the question whether T'_n can be modified to make it useful for testing exponentiality.

Before going further, we relate T'_n to other test statistics already proposed in the literature. Note that $\bar{X}_{>k}$ is the total time on test (TTT) transform after $X_{(k)}$ divided by the empirical distribution function (edf) evaluated at t , $t \in [X_{(k)}, X_{(k+1)})$. Denoting the TTT statistic by

$$D_{n+1}(t) = \sum_{i=1}^k Y_i + (n - k + 1)(t - X_{(k)}), \quad t \in [X_{(k)}, X_{(k+1)}) \quad (2.11)$$

then, after some manipulation we can rewrite

$$T'_n = \max_{1 \leq k \leq n} \frac{n+1}{n-k+1} \left| \frac{D_{n+1}(X_{(k)})}{(n+1)\bar{X}} - \frac{k}{n+1} \right|. \quad (2.12)$$

One may compare these statistics with those proposed by Koul (1978) and later by Bhattacharjee and Sen (1995), to test against NBUE alternatives in uncensored and censored cases respectively and also with Baringhaus and Henze (2000) for testing H_0 against omnibus alternatives. The key feature of T'_n is the weight $(n-k+1)^{-1}$ which comes up naturally in our approach to the problem. The question of interest here is, of course, if this approach can be more fruitful, especially since some power simulations indicated that T'_n does not have good power for certain alternatives to exponentiality. The reason, perhaps, is to be found in the ‘high’ variance associated with the last few residual means.

This observation as well as the desire to overcome the problem in the tail noted for T'_n , motivates us to construct *trimmed* test statistics, whereby some of the last residual means are discarded from T'_n . This has to be done in such a way as to be able to estimate, asymptotically, $m(t)$ over the whole real line. With that in mind, we define

$$T_n = \max_{1 \leq k \leq n - [n^\gamma]} \frac{|\bar{X} - \bar{X}_{>k}|}{\bar{X}}, \quad \gamma \in (0, 1). \quad (2.13)$$

We see that the comparison of the sequence of the residual means goes up to the term with index $n - [n^\gamma]$ where γ is a parameter which determines the number of 'later' residual means to be discarded and $[n^\gamma]$ denotes the greatest integer in n^γ . We will now investigate the properties of this trimmed statistic T_n .

Remark 1. *A nice bonus of our approach is that the statistic T_n can be straightforwardly adapted to the more general case of unknown location*

$$H_0 : F = 1 - e^{-(x-\theta_1)/\theta_2}; \quad \text{vs.} \quad H_1 : F \neq 1 - e^{-(x-\theta_1)/\theta_2}, \quad \theta_1 \in R, \theta_2 > 0, x \geq \theta_1.$$

by simply replacing \bar{X} with $\bar{X}_{>1}$. This does not change the asymptotic distribution of the test statistics.

2.2 Asymptotic properties

In terms of the asymptotic properties of the statistic T_n , we first note that from (2.6), there are at least $[n^\gamma] + 1$ terms in $\bar{X}_{>k}$ for $1 \leq k \leq n - [n^\gamma]$. Under H_0 , the normalized spacings are exponentially distributed with mean θ hence we have

$$\bar{X}_{>k} \xrightarrow{a.s.} \theta \quad 1 \leq k \leq n - [n^\gamma],$$

by the strong law of large numbers, the speed of convergence depending on the choice of the parameter γ . From this it follows that, under exponentiality of the observations

$$T_n \xrightarrow{a.s.} 0.$$

The convergence properties of T_n can be studied under more general conditions, which we do in the following theorem.

Theorem 1. *Let $m(t) < \infty$ and F be a continuous d.f. with mean θ . Then, as $n \rightarrow \infty$,*

$$\max_{1 \leq k \leq n - [n^\gamma]} |\bar{X} - \bar{X}_{>k}| \xrightarrow{a.s.} \sup_{0 \leq t < \infty} |\theta - m(t)| \quad (2.14)$$

The proof of this theorem relies on results obtained earlier by Koul (1978) and is postponed to the end of the paper. The asymptotic distribution of T_n is given by the following result.

Theorem 2. *Let $\gamma \in (0, 1)$, then under the null hypothesis of exponentiality*

$$n^{\gamma/2} T_n \xrightarrow{D} \sup_{0 \leq t \leq 1} |W(t)| \quad (2.15)$$

where $W(t)$ is a Wiener process.

Observe that these two theorems do not hold without the “trimming.” The consistency of our proposed test procedures is a consequence of Theorems 1 and 2, which we state below.

Corollary 1. *The test rejecting the hypothesis of exponentiality for large values of T_n is consistent against each fixed non-exponential alternative distribution with finite mean.*

Theorem 2 shows us that the appropriate normalizing constants depend on trimming. The proof of this theorem relies on asymptotic results for functionals of the uniform quantile process in a weighted metric, its proof being postponed to the last section. From the proof of Theorem 2 we will see that result (2.15) still holds if we substitute n^γ by any $na(n)$ such that $a(n) \rightarrow 0$ and $na(n) \rightarrow \infty$, as $n \rightarrow \infty$.

Simulations show that the asymptotic approximation works better when the trimming is not too large; however, T_n tends to be too conservative even for large sample sizes.

For practical implementation of the test for small to moderate sample sizes, it becomes necessary to evaluate the critical points empirically. Table I below provides such critical values for tests of size $\alpha = 0.05$, obtained by Monte Carlo simulations with 10,000 replications, giving them high accuracy. For other sample sizes and values of γ , a computer program written in “Gauss” is available from the authors upon request.

Table I. Empirical critical points of $n^{\gamma/2}T_n$. Tests of size $\alpha = 0.05$.

γ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
n= 10	1.289	1.439	1.606	1.358	1.226	1.368	1.251	1.139	0.777
n= 20	1.536	1.780	1.612	1.536	1.570	1.621	1.480	1.281	0.937
n= 30	1.708	2.021	1.842	1.812	1.688	1.664	1.547	1.405	1.095
n= 40	1.699	1.627	1.647	1.725	1.706	1.656	1.654	1.472	1.106
n= 50	1.790	1.738	1.763	1.890	1.766	1.773	1.646	1.504	1.142
n= 60	1.790	1.756	1.843	1.795	1.886	1.818	1.721	1.522	1.176
n= 80	1.806	1.796	1.923	1.898	1.908	1.851	1.756	1.606	1.233
n= 100	1.905	1.926	2.045	1.900	1.957	1.893	1.785	1.611	1.260
n= 120	1.863	1.900	1.827	1.951	1.969	1.906	1.815	1.663	1.306
n= 140	1.924	1.960	1.892	1.885	1.970	1.910	1.886	1.713	1.342
n= 160	1.943	2.018	1.9416	1.959	1.977	1.960	1.869	1.710	1.357
n= 180	1.908	1.965	1.931	2.005	1.9451	1.955	1.855	1.707	1.365
n= 200	1.978	2.076	2.026	1.9596	1.934	1.973	1.891	1.703	1.360

Remark 2. *As we have seen, trimming is necessary in order to avoid the last few residual means which happen to have high variability and this, in turn, allows, as it has been shown in Theorems 1 and 2, to obtain consistent test statistics. This strongly suggests that trimming, i.e. the value $[n^\gamma]$, has to increase markedly along with sample size in order to have powerful procedures. This is further confirmed by the simulation study of the next section.*

Remark 3. *The null distribution of T_n depends on that of an ordered uniform random sample. Hence, as noted by Gupta and Richards (1997), the distribution of T_n shares the same invariance property with several other tests for exponentiality, and remain the same for all random vectors X_1, \dots, X_{n+1} having a multivariate Liouville distribution.*

3 Power Estimates and Examples

In this section we consider some power estimates obtained by the method of Monte Carlo. For selected values of the parameter θ , we generated 10,000 samples of size 20, 50 and 80 for the following alternative distributions (all for $\theta > 0$):

- Weibull, $f(x) = \theta x^{\theta-1} \exp\{-x^\theta\} \mathbf{1}_{(x \geq 0)}$;
- Power, $f(x) = \theta^{-1} x^{(1-\theta)/\theta} \mathbf{1}_{(x \in [0,1])}$;
- Lomax, $f(x) = (1 + \theta x)^{-\frac{(\theta+1)}{\theta}} \mathbf{1}_{(x \geq 0)}$;
- Dhillon, $f(x) = \theta x^{\theta-1} \exp\{x^\theta + 1 - e^{x^\theta}\} \mathbf{1}_{(x \geq 0)}$;
- Log-logistic, $f(x) = \frac{\theta x^{\theta-1}}{(1+x^\theta)^2} \mathbf{1}_{(x \geq 0)}$;

- Compound Rayleigh, $f(x) = \frac{2\theta x}{(1+x^2)^{\theta+1}} \mathbf{1}_{(x \geq 0)}$.

These distributions are commonly considered in power studies of tests for exponentiality, in addition they cover a variety of situations which differ from the point of view of the hazard rate (and hence MRL). We have distribution with monotone failure rate as well as mixed cases with either increasing and then decreasing failure rate (IDFR) or decreasing and then increasing failure rate (DIFR). A complete classification is summarized in Table II.

Table II: FR and MRL classification of the distributions used in simulations, $\theta > 0$.

Distribution	IFR (DMRL)	DFR (IMRL)	DIFR (IDMRL)	IDFR (DIMRL)
Weibull(θ)	$\theta > 1$	$\theta < 1$	-	-
Power (θ)	$\theta \leq 1$	-	$\theta > 1$	-
Lomax (θ)	-	$\theta > 0$	-	-
Dhillon(θ)	$\theta \geq 1$	-	$\theta < 1$	-
Log-logistic(θ)	-	$\theta \leq 1$	-	$\theta > 1$
Compound Rayleigh(θ)	-	-	-	$\theta > 0$

As a yardstick for T_n we consider the classical Kolmogorov-Smirnov statistic (KS_n) with estimated mean and the statistic L_n developed by Baringhaus and Henze (2000). The former is well known and is typically used in virtually all empirical power studies, the latter is based on a Kolmogorov-Smirnov type distance and exploits the equivalent characterization (1.5) exploiting in practice a different weighting scheme for the distances between the general mean and the mean residual lives, for details see the above mentioned paper but also, refer to formula (2.12). Hence, this comparison allows us to address the question raised in Section 2 whether the weights $(n - k + 1)^{-1}$ may bring some advantage over the other ways to use MRL for testing procedures.

Figures 1 to 4 below show the empirical power obtained for tests of size 0.05 for various values of the trimming parameter γ going from 0.1 to 0.9, by steps of size 0.1; in order to assess the limit to which trimming can be stretched, we also consider the values $\gamma = 0.93$ and $\gamma = 0.95$. These Figures show that T_n performs quite well.

In Figure 1 we have two distributions with IFR (DMRL) i.e., Weibull with $\theta = 1.2$ and Power with $\theta = 0.5$. Here we see that the power of T_n is generally increasing with γ and it seems that it can be safely stretched to its extremes even for small sample sizes. In nearly all cases a value $\gamma = 0.9$ obtains the highest or close to highest power. For Weibull(1.2) the power of T_n is higher (especially for small samples) or at least comparable to that of both its competitors. KS_n and L_n have very high performance for the Power(0.5) but T_n obtains the same result for a large span of the truncation parameter γ . For these distributions we ran more extensive simulations with other parameter values obtaining similar conclusions.

In Figure 2 we consider distributions with DFR (IMRL), i.e. Lomax with $\theta = 0.5$ and Weibull with $\theta = 0.8$. Contrary to the situation of Figure 1, here we see that too large trimming may not be appropriate, especially for small sample sizes, note however that in such situations the power of T_n can be much higher than that of its competitors; a value of $\gamma = 0.8$ maintains comparable power with L_n and KS_n in most cases.

In Figure 3 we have Power(2) and Dhillon(0.5), i.e. two distributions with DIFR (IDMRL). Again, extreme trimming is not appropriate for small sample sizes, for which, however a value of $\gamma = 0.8$ obtains the highest power among all tests. For larger sample sizes, larger trimming

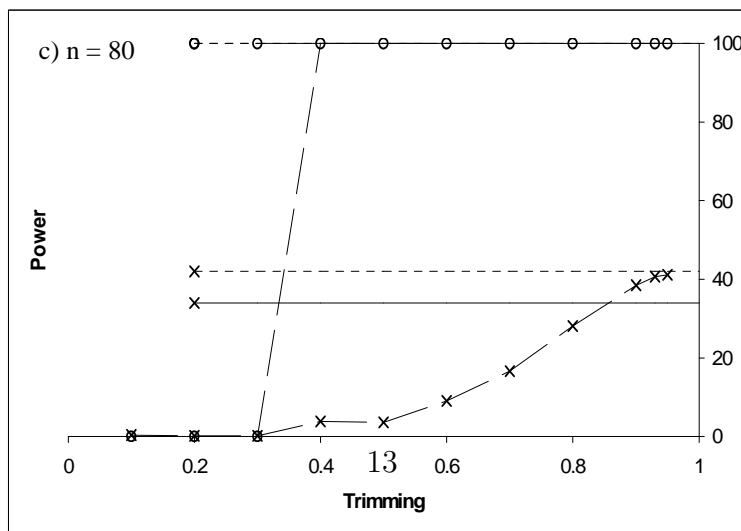
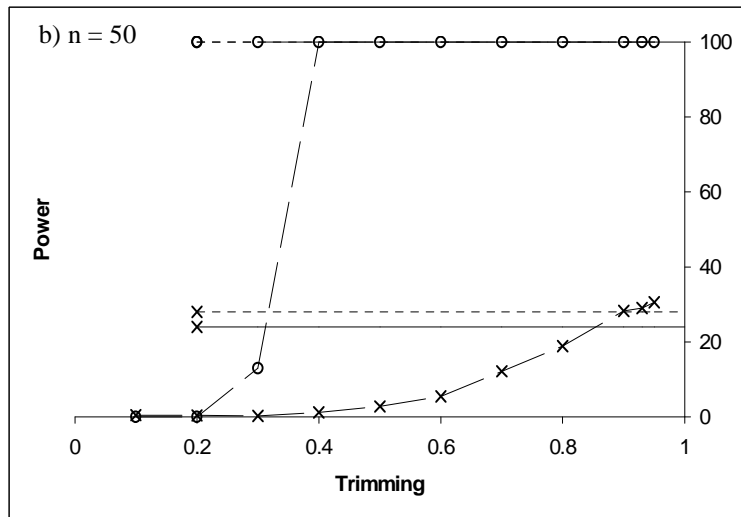
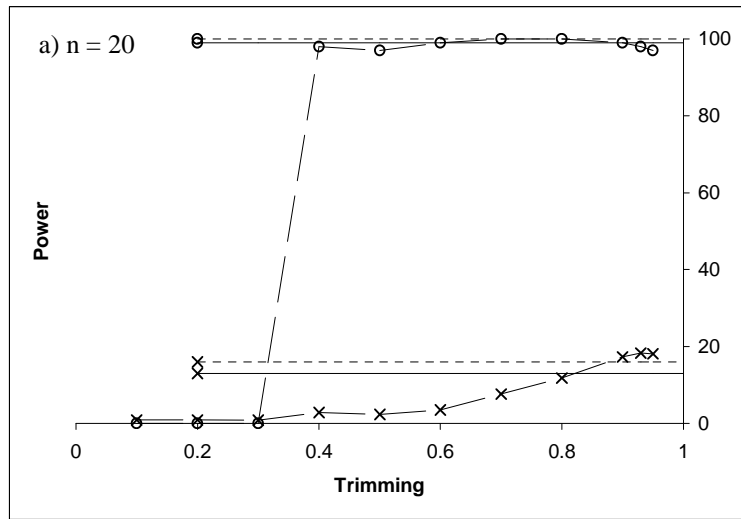


Figure 1: Power values, in percentage, of T_n (dashed), $K S_n$ (solid), L_n (dotted). Distributions are: Weibull (1.2) - x (cross) and Power (0.5) - o (circle).

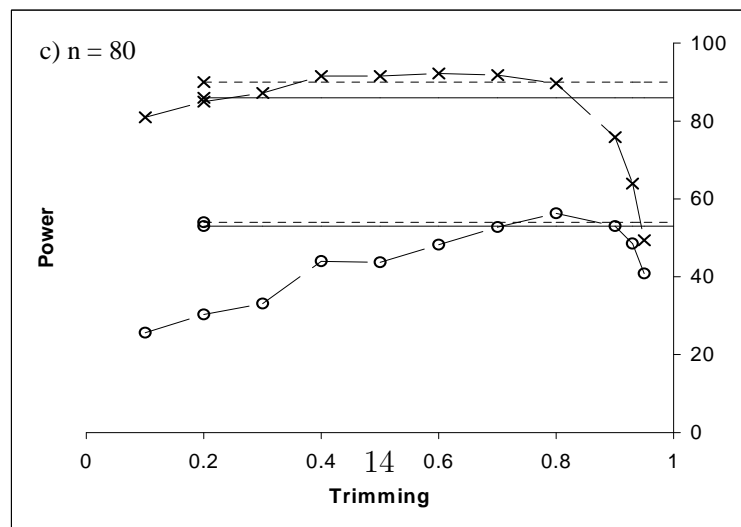
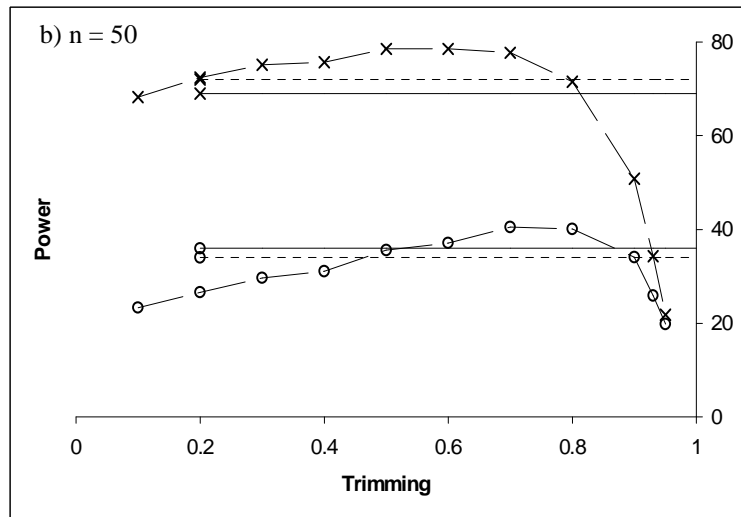
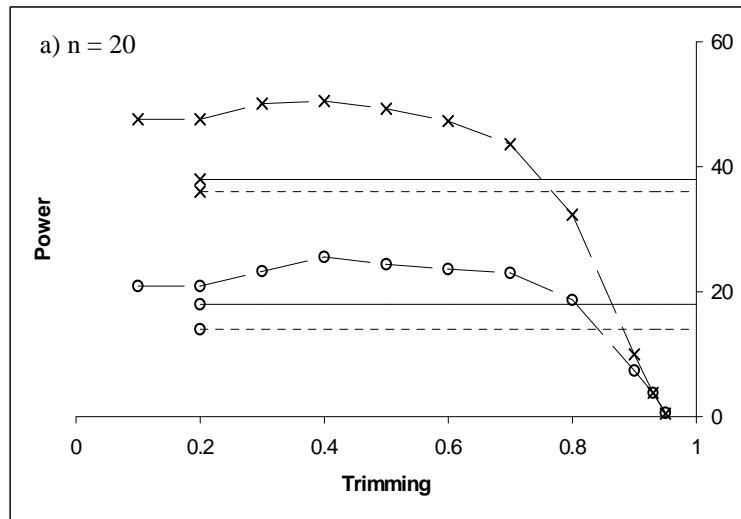


Figure 2: Power values, in percentage, of T_n (dashed), KS_n (solid), L_n (dotted). Distributions are: Lomax (0.5) - x (cross), Weibull (0.8) - o (circle).

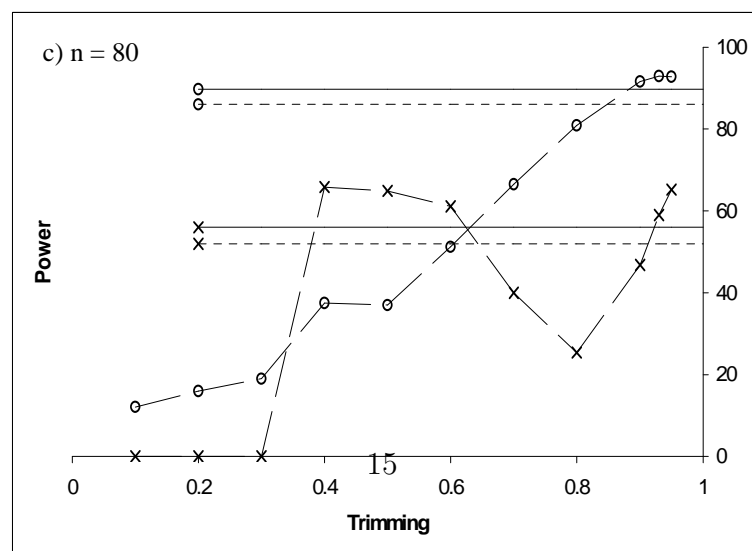
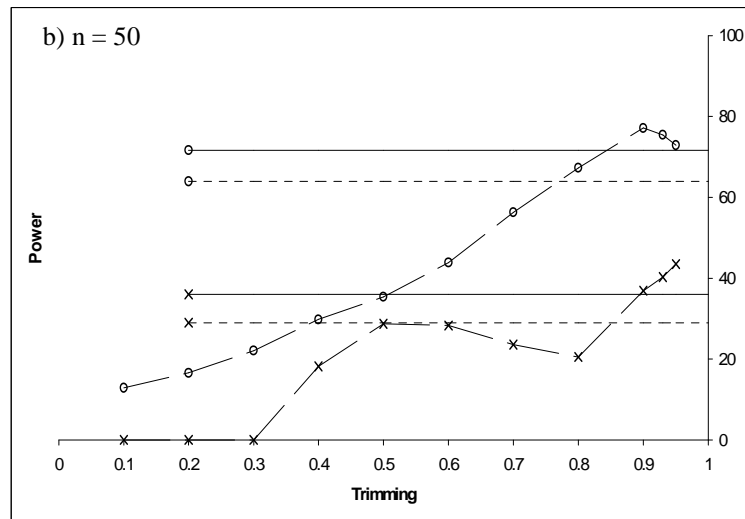
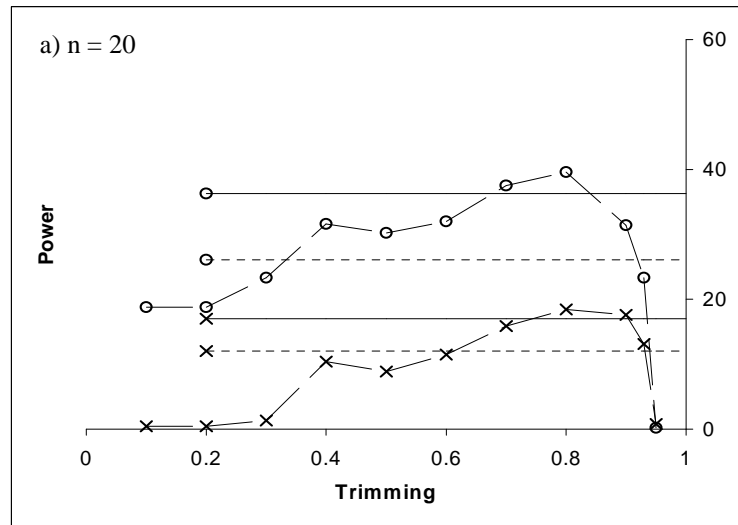


Figure 3: Power values, in percentage, of T_n (dashed), KS_n (solid), L_n (dotted). Distributions are: Power (2) - x (cross), Dhillon (0.5) - o (circle).

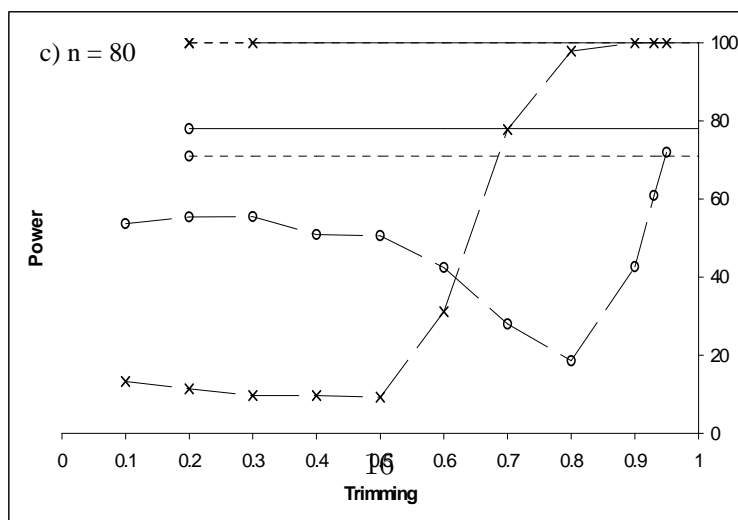
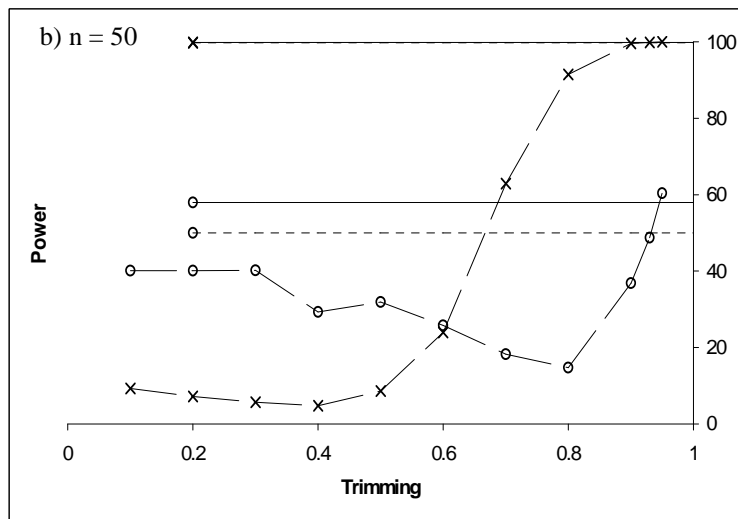
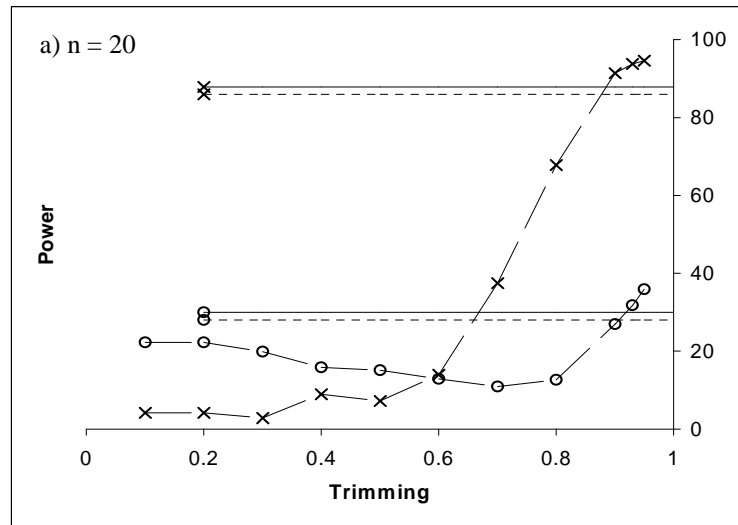


Figure 4: Power values, in percentage, of T_n (dashed), $K S_n$ (solid), L_n (dotted). Distributions are: Log-logistic (3) - x (cross), Compound Rayleigh (1) - o (circle).

becomes preferable by which T_n achieves higher power than L_n and KS_n .

Figure 4 depicts the results for Log-logistic(3) and Compound Rayleigh(1), which represent the case of IDFR (DIMRL) distributions. Again, very large trimming achieve the highest power values which are comparable to those of L_n and KS_n .

In summary we can say that trimming has quite a strong effect on the power of T_n ; indeed, for several alternatives, power is not appreciable until γ reaches a value of at least 0.4 and for several cases, but DFR, it increases nearly monotonically with γ . From the simulations, it appears that a value of $\gamma = 0.8$ or $\gamma = 0.9$ allows to match or overcome the power values of KS_n or L_n restricting considerably the values of γ to be considered in practical cases. Finally, we note that for larger sample sizes larger trimming is preferable, further substantiating the recommendations made in Remark 2.

For a specific application of the test, we consider the classical data set given in Table III (see Kotz and Johnson (1983)) which represents the survival times (in days) after diagnosis of 43 patients with a certain kind of leukemia. For such a data set, IFR may be too restrictive. Hopefully the treatment, applied after diagnosis, will (at least for a period) decrease the failure rate.

Table III: Survival times in days after diagnosis.

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					

If the Hollander and Proschan (1972) test against NBU alternatives is applied, it obtains a p value of 0.07. For the tests considered in our simulations we have $KS_{43} = 0.1617$ with

an estimated p - value of 0.053 and $L_{43} = 1.2742$ with a p - value of 0.072. Del Castillo and Puig (1999b) apply to this data set a likelihood ratio test against singly truncated normal alternatives which are recommended for lifetime data whose nature suggests an IFR distribution with hazard rate not vanishing at 0. Their statistics reject the hypothesis with a p value 0.033. In this case we may try to apply our test statistic with a large number of last residual means discarded. Choosing $\gamma = 0.9$ we obtain $T_{43}^{0.9} = 0.2229$ which rejects the hypothesis of exponentiality with a p value of 0.042.

4 Summary and conclusion

We propose a new test statistic for exponentiality which exploits mean residual life. Although there are other tests in the literature which use characterizations based on the MRL, our procedure seems more natural in the way it removes the unknown scale-factor, relating the test to uniform spacings.

As we have seen, practical applications require one to choose the amount of trimming via the choice of γ in order to carry out the test. However, it is not uncommon to find general as well as specialized test procedures which require the choice of some kind of parameter; see eg. Gail and Gaswirth (1978), Lee et al. (1980), Bandyopadhyay and Basu (1990), Henze (1993), Grzegorzewski and Wieczorkowski (1999), Klar (2000). This may indeed work out to be an advantage, since the fine-tuning involved in choosing such a parameter may allow the test procedure to be more sensitive towards specific classes of alternatives under consideration.

We find through extensive simulations that values of $\gamma = 0.8$ or $\gamma = 0.9$ are appropriate in most cases, although the situation maybe less certain for small samples. Indeed, for moderate or large sample sizes, unless one suspects to be in the case of DFR alternatives, we see that there is a clear indication of using a value of γ of 0.9 or even higher, restricting considerably the uncertainty over γ and obtaining very good performances of T_n which outperforms KS_n and L_n . In extensive simulations we ran for other common alternative distributions such as the Gamma, the Half Normal, the Half Cauchy, the Pareto, and the Uniform, we observed similar results.

We find that T_n works well even in situations where traditional tests of exponentiality fail to detect departures from the null hypothesis, as eg. in particular for IFR. It is gratifying to observe in our extensive simulations, that our test performs as well as some of the specialized tests that have been proposed for general classes of alternatives to exponentiality.

Remark 4. *The reasoning of Section 2.1 lends itself to develop a test based on a “quadratic distance” (similar to Cramer-von Mises statistic) between the sample mean and the residual sample means, so that one might consider a “trimmed” quadratic test statistic viz.*

$$\sum_{1 \leq k \leq n - [n^\gamma]} \left[\frac{\bar{X} - \bar{X}_{>k}}{\bar{X}} \right]^2 \quad (4.16)$$

The simulations we ran for this quadratic test show that one runs into similar problems at the upper tail as those encountered for T_n , and indeed their power performances are nearly identical with values which may be slightly higher or lower depending on the alternative considered.

5 Proofs

Recall that $U_{(k)}$ is the k -th order statistics from a uniform random sample of size $n + 1$ and that, as $n \rightarrow \infty$, $\max_{1 \leq k \leq n+1} |U_{(k)} - k/(n + 1)| \stackrel{a.s.}{\rightarrow} 0$.

Proof of Theorem 1. . For $t \in [X_{(k)}, X_{(k+1)})$ we write

$$\bar{X}_{>t} = \frac{1}{n - k + 1} \sum_{k+1}^{n+1} (X_{(i)} - t) = [\bar{F}_{n+1}(t)]^{-1} \left[\bar{X} - \frac{D_{n+1}(t)}{n + 1} \right]. \quad (5.17)$$

where $D_{n+1}(t)$ has been defined in (2.11). Note that from Koul (1978) we have that $D_{n+1}(t)/(n + 1) \stackrel{a.s.}{\rightarrow} \int_0^t \bar{F}(x) dx$ and by the Glivenko-Cantelli theorem $\bar{F}_{n+1}(t) \stackrel{a.s.}{\rightarrow} \bar{F}(t)$ uniformly in t . Then, as $n \rightarrow \infty$, we have

$$\sup_{0 < t < \infty} |\bar{X} - \bar{X}_{>t}| \stackrel{a.s.}{\rightarrow} \sup_{0 < t < \infty} |\theta - m(t)|.$$

Uniform convergence of $|\bar{X} - \bar{X}_{>k}|$ can be now shown by using standard arguments if we first note that the difference $\bar{X}_{>t} - \bar{X}_{>k}$ can be made arbitrarily small for all $t > 0$ in the interval $[X_{(k)}, X_{(k+1)})$. To this end, note that, for $t \in [X_{(k)}, X_{(k+1)})$,

$$\bar{X}_{>t} - \bar{X}_{>k} = t - X_{(k)}.$$

From continuity of F this difference can be made as small as desired, in fact we have

$$|F(t) - F(X_{(k)})| \leq |F(X_{(k+1)}) - F(X_{(k)})| \leq \left| U_{(k+1)} - \frac{k+1}{n+1} \right| + \left| U_{(k)} - \frac{k}{n+1} \right| + \frac{1}{n+1}$$

and the *r.h.s.* of the above expression converges to 0 a.s. □

The proof of Theorem 2 relies on the representation of T_n in terms of a uniform quantile process in a weighted metric. Note in fact that by using the same representation as in (2.12) and after some manipulation we have

$$\frac{\bar{X} - \bar{X}_{>k}}{\bar{X}} \stackrel{D}{=} \frac{\frac{i}{n+1} - U_{(i)}}{\frac{i}{n+1}}$$

for $k = 1, \dots, n$, $i = n - k + 1$. The relevant asymptotic theory can be obtained by exploiting results on uniform empirical processes in weighted metrics. To this end we define

$$U_n(t) = \begin{cases} U_{(i)} & \frac{i}{n+2} < t \leq \frac{i+1}{n+2}, \quad i = 1 \dots n, \\ 0 & \text{otherwise} \end{cases}$$

$$q_n(t) = \begin{cases} \frac{i}{n+1} & \frac{i}{n+2} < t \leq \frac{i+1}{n+2}, \quad i = 1 \dots n, \\ 0 & \text{otherwise} \end{cases}$$

The function $[q_n(t) - U_n(t)]/q_n(t)$ is a step function with jump points in $i/(n+2)$, $i = 1, \dots, n$, hence, if we define a process

$$\tilde{u}_n(t) = \sqrt{n}[q_n(t) - U_n(t)]$$

then it holds that

$$\sqrt{n}T_n \stackrel{D}{=} \sup_{\frac{n\gamma}{n+2} < t \leq \frac{n+1}{n+2}} \left| \frac{\tilde{u}_n(t)}{q_n(t)} \right|.$$

Next we define a continuous version of the uniform quantile process as

$$u_n(t) = \sqrt{n}[t - U_n(t)].$$

It is clear from the definition that $u_n(\frac{i}{n+1}) = \tilde{u}_n(t)$, $\frac{i}{n+2} < t \leq \frac{i+1}{n+2}$, $i = 1 \dots n$. The asymptotic distributions of T_n will then be obtained by those of the corresponding functionals

of the weighted uniform quantile process $u_n(t)/t$. We will adapt the methods discussed in Csörgő and Horváth (1993) by providing suitable approximations by Gaussian processes. Since the weak convergence of $u_n(t)$ in weighted metric does not imply the convergence in distribution of supremum functionals for the weight function $q(t) = t$, we will study the distributional properties of T_n on its own. Before doing so we need some preliminary results that, for convenience, we recall here in the form of lemmas.

Lemma 1. *We can define a sequence of Brownian Bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that*

$$\sup_{0 \leq t \leq 1} |\tilde{u}_n(t) - B_n(t)| \stackrel{a.s.}{=} O(n^{-1/2} \log n) \quad (5.18)$$

and

$$n^{1/2-v} \sup_{\lambda/n \leq t \leq 1} \frac{|\tilde{u}_n(t) - B_n(t)|}{t^v} = O_p(1) \quad (5.19)$$

for all $0 < v \leq 1/2$ and $0 < \lambda < \infty$.

This lemma is just a straightforward modification of Theorem 4.4.2 in Csörgő and Horváth (1993) which can be obtained by noting that $\sup_{0 < t < 1} |u_n(t) - \tilde{u}_n(t)| \leq \sqrt{n}/(n+2)$.

Lemma 2. *(Csörgő and Horváth, 1993, p.302 Theorem 2.3, iv). Let $a(n) \rightarrow 0$, $na(n) \rightarrow \infty$ as $n \rightarrow \infty$. Then*

$$a(n)^{1/2} \sup_{a(n) \leq t \leq 1-a(n)} \frac{|u_n(t)|}{t} \xrightarrow{D} \sup_{0 \leq t \leq 1} |W(t)|. \quad (5.20)$$

Proof of Theorem 2. i) we will first show that

$$\sup_{\frac{n\gamma}{n+2} < t \leq \frac{n+1}{n+2}} |n^{\gamma/2} T_n(t) - n^{\gamma/2-1/2} \frac{u_n(t)}{t}| = o_p(1). \quad (5.21)$$

Note that by the definition of $\tilde{u}_n(t)$ and $u_n(t)$ we can rewrite the above expression as

$$\begin{aligned}
n^{\gamma/2} \sup_{\frac{n^\gamma}{n+2} < t \leq \frac{n+1}{n+2}} \left| \frac{U_n(t)}{t} - \frac{U_n(t)}{q_n(t)} \right| &= n^{\gamma/2} \max_{n^\gamma < i \leq n} U_{(i)} \sup_{\frac{i}{n+2} < t \leq \frac{i+1}{n+2}} \left| \frac{1}{t} - \frac{n+1}{i} \right| \\
&= n^{\gamma/2} \max_{n^\gamma < i \leq n} U_{(i)} \left[\max \left(\frac{n-i+1}{i(i+1)}, \frac{1}{i} \right) \right] \\
&= n^{\gamma/2} \max \left[\max_{n^\gamma < i \leq n/2} U_{(i)} \frac{n-i+1}{i(i+1)}, \max_{n/2 < i \leq n} U_{(i)} \frac{1}{i} \right]
\end{aligned}$$

and the above expression can be seen to be $n^{\gamma/2} O_p(n^{-\gamma}) = o_p(1)$. Next, from the fact that $\sup_{0 \leq t \leq 1} |u_n(t)|$ is bounded in probability we easily see that

$$n^{\gamma/2-1/2} \sup_{1-\frac{n^\gamma}{n} \leq t \leq 1} \frac{|u_n(t)|}{t} = o_p(1),$$

hence result (2.15) follows from applying Lemma 2 with $a(n) = n^\gamma/n$.

□

Acknowledgments. The authors would like to thank an anonymous referee who suggested new distributions for comparisons and simulations, and suggested the streamlined classification given in Table II.

6 References

- Ahmad, I.A. and Alwasel, I. (1999). A goodness-of-fit test for exponentiality based on the memoryless property. *J. R. Stat. Soc. Ser. B* **61**, 681-689.
- Alwasel, I. (2001). On goodness of fit testing for exponentiality using the memoryless property. *J. of Nonparametric Statistics* **13**, 569-581.

- Ascher, S. (1990). A survey of tests for exponentiality. *Comm. Statist. Theory Methods* **19**, 1811-1825.
- Bandyopadhyay, D. and Basu, A.P. (1990). A class of tests for exponentiality against decreasing mean residual life alternatives. *Comm. in Stat. - Theory and Methods* **19**, 905-920.
- Baringhaus, L. and Henze, N. (1991). A class of consistent tests for exponentiality based on the empirical Laplace transform. *Ann. Inst. Statist. Math.* **43**, 551-564.
- Baringhaus, L. and Henze, N. (1992). An Adaptive Omnibus Test for Exponentiality. *Communications in Statistics - Theory and Methods* **21**, 969-978.
- 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.
- Bergman, B. and Klefsjö, B. (1989). A family of test statistics for detecting monotone mean residual life. *J. Stat. Plann. and Infer.* **21**, 161-178.
- Bhattacharjee, M.C. and Sen, P.K. (1995). Kolmogorov -Smirnov type tests for NB(W)UE alternatives under censoring schemes. In Koul, H.L. and Deshpande, J.V. (Eds.) *Analysis of Censored Data*, 25-38.
- Csörgő, M. and Horváth, L. (1993). *Weighted approximations in probability and statistics*. Wiley, New York.
- Del Castillo, J. and Puig, P. (1999a). Invariant exponential models applied to reliability theory and survival analysis. *J. Amer. Statist. Assoc.*, **94**, 522-528.

- Del Castillo, J. and Puig, P. (1999b). The best test of exponentiality against singly truncated normal alternatives. *J. Amer. Statist. Assoc.*, **94**, 529-532.
- Doksum, K.A. and Yandell, B.S. (1984). Tests for exponentiality. *Handbook of statistics*, Vol. 4, North Holland, 579-612.
- Ebrahimi, N., Soofi, E.S. and Habibullah, M. (1992). Testing Exponentiality Based on Kullback-Leibler Information. *J. of the Roy. Stat. Soc., Ser. B* **54**, 739-748.
- Gail, M.H. and Gastwirth, L. A scale free goodness of fit test for the Exponential distribution based on the Lorentz curve. *J. Amer. Stat. Assoc.* **73**, 787-793.
- Gatto, R. and Jammalamadaka, S. Rao (2002). A saddlepoint approximation for testing exponentiality against some increasing failure rate alternatives. *Statistics and Probab. Letters* **58**, 71-82.
- Grzegorzewski, P. and Wieczorkowski, R. (1999). Entropy-Based Goodness of Fit Test for Exponentiality. *Communications in Statistics - Theory and Methods* **28**, 1183-1202.
- Grubbs, F.E. (1971) Fiducial bounds on reliability for the two parameter negative exponential distribution. *Technometrics* **13**, 873-876.
- Guess, F. and Proschan, F. (1988). Mean residual life, theory and applications. In Krishnaiah, P.R. and Rao, C.R. (Eds.), *Handbook of Statistics Vol. 7, Reliability and Quality Control*, 215-224.
- Gupta, R.D. and Richards, D.St.P. (1997). Invariance properties of some classical tests for

- exponentiality. *J. Stat. Plann. Inference* **63**, 203-213.
- Henze, N. and Meintanis, S.G. (2002). Tests of fit for Exponentiality based on the empirical Laplace transform. *Statistics* **36**, 147-162.
- Henze, N. (1993). A New Flexible Class of Omnibus Tests for Exponentiality. *Communications in Statistics - Theory and Methods* **22**, 115-133.
- Hollander, M. and Proschan, F. (1972). Testing whether New is Better than Used. *Ann. Math. Stat.* **43**, 1136-1146.
- Jammalamadaka, S. Rao and Lee, E.S. (1998). Testing for Harmonic New Better than Used in Expectation. *Probab. in Engg. and Inform. Sci.* **12**, 409-416.
- Jammalamadaka, S. Rao 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**, 719-729
- Keating, J.P., Glaser, R.E. and Ketchum, N.S. (1990). Testing hypotheses about the shape of a Gamma distribution. *Technometrics*, **32**, 67-82.
- Klar, B. (2000). A class of tests for exponentiality against HNBUE alternatives. *Statistics and Prob. Letters* **47**, 199-207.
- Klar, B. (2001). Goodness-of-fit tests for the exponential and the normal distribution based on the integrated distribution function. *Ann. Inst. Statist. Math.* **53**, no. 2, 338-353.
- Klar, B. (2003). On a test for exponentiality against Laplace order dominance. *Statistics*

37, 505-515.

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

Koul, H.L. (1978). Testing for new is better than used in expectation. *Comm. Statist. - Theor. Meth.* **7**, 685-701.

Lawless, J.F. (1982). *Statistical Models and Methods for Lifetime Data*. Wiley, New York.

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

Patel, J.K. (1983). Hazard rate and other classifications of distributions. In Johnson, N.L. and Kotz, S, Eds. *Encyclopedia od Statistical Sciences*. Wiley, New York.

Shanbhag, D.N. (1970). The Characterizations for Exponential and Geometric Distributions. *J. Amer. Statist. Assoc.* , *65*, 1256-1259.

Shapiro, S.S. (1995). Goodness of Fit tests. In *The Exponential Distribution: Theory Methods and Applications*. Balakrishnan, N. and Basu, A.P. Eds, Gordon and Breach, Amsterdam.

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