Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

A New Test for New Better Than Used in Expectation Lifetimes

Edgardo Lorenzo, Ganesh Malla & Hari Mukerjee

To cite this article: Edgardo Lorenzo, Ganesh Malla & Hari Mukerjee (2015) A New Test for New Better Than Used in Expectation Lifetimes, Communications in Statistics - Theory and Methods, 44:23, 4927-4939, DOI: 10.1080/03610926.2013.824101

To link to this article: https://doi.org/10.1080/03610926.2013.824101

Accepted author version posted online: 09 Feb 2015.
Published online: 30 Nov 2015.

Communications in Statistics—Theory and Methods, 44: 4927–4939, 2015

ISSN: 0361-0926 print / 1532-415X online
DOI: 10.1080/03610926.2013.824101

A New Test for New Better Than Used in Expectation Lifetimes

EDGARDO LORENZO,[1] GANESH MALLA,[2] AND HARI MUKERJEE[3]

[1]Department of Mathematics, University of Puerto Rico at Mayagüez, Mayagüez, Puerto Rico
[2]Department of Mathematics, Mount St. Joseph University, Cincinnati, Ohio, USA
[3]Department of Mathematics and Statistics, Wichita State University, Wichita, Kansas, USA

The mean residual life of a non negative random variable X *with a finite mean is defined by $M(t) = E[X - t|X > t]$ for $t \geq 0$. A popular nonparametric model of aging is new better than used in expectation (NBUE), when $M(t) \leq M(0)$ for all $t \geq 0$. The exponential distribution lies at the boundary. There is a large literature on testing exponentiality against NBUE alternatives. However, comparisons of tests have been made only for alternatives much stronger than NBUE. We show that a new Kolmogorov-Smirnov type test is much more powerful than its competitors in most cases.*

Keywords New better than used in expectation; Hypothesis test; Asymptotic properties.

Mathematics Subject Classification Primary 62G10; Secondary 62N05, 62N01.

1. Introduction

Let X be a strictly positive random variable (RV) with the distribution function (DF) F, the survival function (SF) $S = 1 - F$, and a finite mean μ. The mean residual life (MRL) of X is defined by $M(t) = E[X - t|X > t],\ t \geq 0$, i.e.,

$$M(t) = \frac{\int_t^\infty S(u)du}{S(t)} I(S(t) > 0).$$

The inversion formula

$$S(t) = \frac{M(0)}{M(t)} e^{-\int_0^t \frac{1}{M(u)}du} I[M(t) > 0]$$

Received January 15, 2013; Accepted July 8, 2013.
Address correspondence to Ganesh Malla, Department of Mathematics, Mount St. Joseph University, 5701 Delhi Rd, Cincinnati, OH 45233 USA; E-mail: ganesh.malla@msj.edu

recovers the SF. Note that $M(0) = \mu$. Yang (1978) considered the empirical estimate of M given by

$$M_n(t) = \frac{\int_t^\infty S_n(u)du}{S_n(t)} I(S_n(t) > 0),$$

where S_n is the empirical SF based on the sample $X_1 \leq X_2 \leq \cdots \leq X_n$, and analyzed its asymptotic properties. Note that $M_n(0)$ is simply the sample mean $\bar{X}_n$.

There are various models of aging using MRL. The lifetime is said to have the new better than used in expectation (NBUE) property if $M(t) \leq M(0)$ for $t \geq 0$. The NBUE assumption relaxes the more stringent model of decreasing mean residual life (DMRL) when $M(t)$ is decreasing (we use decreasing (increasing) for nonincreasing (nondecreasing)) for all t. When the density $f = F'$ exists, an even more stringent model is the increasing failure rate (IFR) where the hazard rate $\lambda = f/S$ is increasing. It is well known that IFR $\Rightarrow$ DMRL $\Rightarrow$ NBUE. The exponential distribution lies at the boundary of all of these models.

The NBUE imposes the weakest condition on the lifetime of the three models described above. In particular, it allows for jumps in the SFs whereas the SFs are continuous, except possibly for jumps at the endpoints of the supports, in the other two models. The NBUE model is appropriate for reparable systems in reliability. In particular, Hollander and Proschan (1972) show that age replacement policy is better than no planned replacement policy if and only if the system SF is NBUE. One can also find applications in survival analysis. In a study of resistance developed by guinea pigs with different doses of a bacterium, the life distribution appeared to be DMRL initially, and then, as resistance developed, it appeared to have an increasing MRL, followed eventually by the DMRL pattern (Bjerkedal, (1960). However, at almost all dosages, the life distribution seemed to be NBUE.

Because of the popularity of the NBUE model, there is a large literature on testing exponentiality against the NBUE alternative:

$$H_0 : S \text{ is exponential} \quad vs \quad H_1 : S \text{ is NBUE but not exponential},$$

which is equivalent to testing

$$H_0 : M(t) \equiv M(0) \ vs \ H_1 : M(t) \leq M(0) \text{ for all } t \text{ with } M(t) < M(0) \text{ for some } t > 0.$$

The article by Anis and Basu (2012) gives an excellent review. All of the tests assume that S is continuous (which is assumed in our discussion of all tests in the literature without explicit mention), and all but one of the proposed test statistics are sample equivalents of integrals of various measures of directed deviations of H_0 to H_1. Hollander and Proschan (HP) (1975) were the first to provide such a test, based on the measure

$$\gamma = \int_0^\infty S(t)[M(0) - M(t)]dF(t); \tag{1}$$

$\gamma = 0$ if and only if S is exponential, and $\gamma > 0$ under H_1. The weight $S(t)$ is a natural one in the sense that the proportion of subjects alive at time t is $S(t)$. HP (1975) first define the statistic K_n and derive its form by

$$K_n = \int_0^\infty S_n(t)[M_n(0) - M_n(t)]dF_n(t) = \frac{1}{n^2}\sum_{i=1}^{n}\left(\frac{3n}{2} - 2i + \frac{1}{2}\right)X_i. \tag{2}$$

Since the problem is scale invariant, they define their test statistic by

$$K_n^* = \frac{K_n}{\bar{X}_n} = \frac{K_n}{M_n(0)}, \tag{3}$$

where $\bar{X}_n = M_n(0)$ is the sample mean. The test statistic turns out to be identical to that used by Barlow (1968) for testing exponentiality against the much narrower IFR alternatives. HP (1975) show that K_n^* is consistent against NBUE alternatives and that $\sqrt{12n}K_n^* \xrightarrow{d} N(0, 1)$ under H_0.

Anis and Mitra (AM) (2011) generalized the HP measure γ in (1) by the class of measures,

$$\gamma_j = \int_0^\infty [S(t)]^j [M(0) - M(t)]dF(t), \quad j > 0. \tag{4}$$

The larger the value of j, the more emphasis is placed on large values of $M(0) - M(t)$ occurring at smaller t's and vice versa. Lacking such pre-knowledge of the shapes of the MRLs, we prefer the more natural choice of $j = 1$ that corresponds to the HP measure (1). However, AM (2011) first integrate out the first term in the rhs of (1) to get

$$\begin{aligned}\gamma_1 &= \frac{M(0)}{2} - \int_0^\infty \int_t^\infty S(t)dF(t) = \frac{M(0)}{2} - \left[M(0) - \int_0^\infty S(t)(1 - S(t))dt\right] \\ &= -\frac{M(0)}{2} + \int_0^\infty S^2(t)dt.\end{aligned}$$

Then they derive the form of its empirical estimate by

$$\gamma_{1n} = -\frac{M_n(0)}{2} + \int_0^\infty S_n^2(t)dt = \frac{1}{n^2}\sum_{i=1}^n \left(\frac{3n}{2} - 2i + 1\right) X_i. \tag{5}$$

Their test statistic is $\gamma_{1n}^* = \gamma_{1n}/\bar{X}_n$ which is scale invariant. They thought that there was a mistake in the HP test statistic even though they are asymptotically equivalent. But, as our analysis shows, both tests are valid. In fact, since $\gamma_{1n}^* - K_n^* = 1/(2n)$, the critical values of $\sqrt{12n}K_n^*$ are exactly $\sqrt{3/n}$ less than those of $\sqrt{12n}\gamma_{1n}^*$. The difference $\gamma_{1n}^* - K_n^* = 1/(2n)$ holds for all continuous alternatives. However, the main purpose of this paper is to develop a good test against all NBUE alternatives, including those with jumps. Since the HP and AM test statistics are invalid for the case when the SF has jumps, we needed to derive their statistics for these cases. Now, the difference between the HP and AM statistics could be very large for such alternatives, and they are very dependent on the specific alternatives as opposed to the continuous case. At the end of the paragraph before (17), we give some heuristic reasons why this is the case. The dramatic differences in powers may be seen in the simulations in Table 3.

However, we will show that the AM test is much superior to the HP test when S has jumps, even though the AM test performs poorly relative to the new test we propose.

Recently, Basu and Anis (2012) derived the exact distribution of the AM statistic under H_0 and provided a table of critical values.

Anis and Basu (2012) made power comparisons of most of the known tests using several parametric families of SFs. Unfortunately, all well-known parametric families of SFs that are NBUE happen to be not only DMRL, but IFR. Nothing is known about the performance of these tests for NBUE but not DMRL, or even DMRL but not IFR

Table 1
Simulated critical values of $\sqrt{n}T_n^*$

n	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
5	1.1968	1.3239	1.5786
6	1.1810	1.3211	1.5745
7	1.1761	1.3148	1.5818
8	1.1716	1.3122	1.5779
9	1.1663	1.3077	1.5853
10	1.1640	1.3074	1.5788
11	1.1621	1.3017	1.5798
12	1.1584	1.3009	1.5801
13	1.1555	1.3019	1.5708
14	1.1550	1.2961	1.5760
15	1.1484	1.2950	1.5763
16	1.1460	1.2916	1.5623
17	1.1458	1.2900	1.5740
18	1.1454	1.2890	1.5629
19	1.1381	1.2867	1.5618
20	1.1379	1.2851	1.5741
21	1.1417	1.2833	1.5627
22	1.1392	1.2830	1.5766
23	1.1325	1.2818	1.5725
24	1.1350	1.2837	1.5638
25	1.1323	1.2790	1.5594
26	1.1290	1.2777	1.5577
27	1.1282	1.2796	1.5591
28	1.1314	1.2759	1.5622
29	1.1266	1.2774	1.5667
30	1.1296	1.2770	1.5707
35	1.1268	1.2745	1.5651
40	1.1200	1.2705	1.5655
45	1.1175	1.2691	1.5594
50	1.1163	1.2671	1.5600
60	1.1160	1.2657	1.5486
70	1.1108	1.2627	1.5560
80	1.1059	1.2550	1.5516
90	1.1036	1.2510	1.5478
100	1.1050	1.2556	1.5377
∞	1.0730	1.2239	1.5174

alternatives. We propose a new Kolmogorov-Smirnov (KS) type sup-test that allows for ties in the observations, using the statistic

$$T_n^* = \frac{T_n}{\bar{X}_n}, \quad \text{where} \quad T_n = \sup_{t \geq 0} S_n(t)[M_n(0) - M_n(t)]. \tag{6}$$

Table 2
Critical values (% difference from asymptotic values) of $\sqrt{12n}\gamma_{1n}^*$ and $\sqrt{n}T_n^*$

	$\sqrt{12n}\gamma_{1n}^*$			$\sqrt{n}T_n^*$		
n	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
10	1.77 (38.3)	2.11(27.9)	2.73 (17.2)	1.16 (8.4)	1.31 (7.4)	1.58 (3.9)
20	1.64 (28.1)	1.99 (20.6)	2.64 (13.3)	1.14 (6.5)	1.29 (5.7)	1.57 (3.3)
30	1.58 (23.4)	1.94 (17.6)	2.60 (11.6)	1.13 (5.6)	1.28 (4.9)	1.57 (3.3)
40	1.54 (20.3)	1.90 (15.2)	2.57 (10.3)	1.12 (4.7)	1.27 (4.1)	1.57 (3.3)
50	1.52 (18.8)	1.88 (13.9)	2.55 (09.4)	1.12 (4.7)	1.27 (4.1)	1.56 (2.6)
60	1.49 (16.4)	1.85 (12.1)	2.52 (08.2)	1.12 (4.7)	1.27 (4.1)	1.55 (2.0)
70	1.48 (15.6)	1.84 (11.5)	2.52 (08.2)	1.11 (3.7)	1.26 (3.3)	1.56 (2.6)
80	1.48 (15.6)	1.82 (10.3)	2.50 (07.3)	1.11 (3.7)	1.26 (3.3)	1.55 (2.0)
90	1.46 (14.1)	1.82 (10.3)	2.49 (06.9)	1.11 (3.7)	1.26 (3.3)	1.54 (1.3)
100	1.45 (13.3)	1.81 (09.7)	2.48 (06.4)	1.11 (2.8)	1.26 (3.3)	1.54 (1.3)
∞	**1.28**	**1.65**	**2.33**	**1.07**	**1.22**	**1.52**

Koul (1978), using the fact that S is NBUE if and only if $(1/\mu)\int_0^t F(u)du \le F(t)$ for all $t \ge 0$, defined a different KS-type test statistic,

$$D_n^* = \max_{1\le i\le n} D_i, \text{ where } D_i = \frac{\int_0^{X_i} F_n(t)dt}{\bar{X}_n} - F_n(X_i) = \frac{\sum_{j=1}^{i} X_j + (n-i)X_i}{n\bar{X}_n} - \frac{i}{n}. \quad (7)$$

Although $\sqrt{n}D_n^*$ and $\sqrt{n}T_n^*$ have the same asymptotic distributions under H_0, they can behave very differently under H_1 as we will show.

In Sec. 2, we describe our test, derive its asymptotic properties, and provide a table of its critical values for small and moderate sample sizes by simulations. In Sec. 3, we

Table 3
Percent rejected for the extreme NBUE alternatives in (12)

	$n = 30$				$n = 50$			
(h, Δ)	K_n^*	γ_{1n}^*	D_n^*	T_n^*	K_n^*	γ_{1n}^*	D_n^*	T_n^*
$(Q_{.00} = 0, 0.2)$	05.63	8.01	4.74	25.54	05.52	10.22	4.56	81.20
$(Q_{.00} = 0, 0.3)$	06.58	18.98	4.11	94.32	06.65	28.02	3.95	99.99
$(Q_{.00} = 0, 0.4)$	07.95	46.28	3.44	99.97	07.73	65.71	3.25	100.0
$(Q_{.00} = 0, 0.5)$	09.14	77.92	2.49	100.0	09.17	93.36	2.30	100.0
$(Q_{.10} = .105, 0.2)$	05.43	8.00	4.63	28.55	05.54	09.90	4.56	55.56
$(Q_{.10} = .105, 0.3)$	06.25	16.81	4.04	72.42	06.25	22.95	4.03	93.81
$(Q_{.10} = .105, 0.4)$	07.12	35.01	3.24	93.86	06.92	49.08	3.06	99.62
$(Q_{.10} = .105, 0.5)$	07.56	60.05	2.31	99.10	07.51	77.84	2.08	99.99
$(Q_{.25} = .288, 0.2)$	05.14	07.57	4.41	23.48	05.20	08.82	4.37	38.20
$(Q_{.25} = .288, 0.3)$	05.56	13.12	3.66	51.85	05.41	16.76	3.61	75.05
$(Q_{.25} = .288, 0.4)$	05.90	23.72	2.72	76.98	05.73	31.53	2.63	93.77
$(Q_{.25} = .288, 0.5)$	05.65	39.10	1.68	91.32	05.70	52.33	1.71	96.98

compare the power of our estimator with that of the HP, AM, and Koul tests after deriving their test statistics in the case of ties; we do not make comparisons with other tests since the HP and the AM tests appear to be among the best in the study of Anis and Basu (2012). Intuitively, if $M(0)$ and $\Delta \equiv \sup_{t\geq 0}[M(0) - M(t)]$ in the alternative are held fixed, we would expect the sup-test to be more powerful than an integral test if $M(t)$ is substantially less than $M(0)$ over a small range, while the opposite will be true if $M(t) \approx M(0) - \Delta$ over a large range. Using simulations, we show that our sup-test has much higher power than the others for alternatives that are NBUE with discontinuous SFs. We also provide explanations as to why the HP test performs so poorly relative to the AM test and why Koul's test performs so poorly relative to our test. For DMRL, but not IFR alternatives, the sup-tests are still substantially more powerful than the others. However, for some IFR alternatives, the integral tests perform better. In Sec. 4, we give two examples of application of our procedure. In Sec. 5, we make some concluding remarks.

2. The T_n^* Test and Its Asymptotic Properties

We consider the test described in the Introduction using our test statistic T_n^*. We first derive the form of T_n. Let $d_0 = 0$ and let

$$0 = d_0 < d_1 < d_2 < \cdots < d_m = X_n$$

be the m distinct observation points with r_i observations at d_i for $1 \leq i \leq m$. Let $R_i = \sum_{j=1}^{i} r_j,\ \ 1 \leq i \leq m$ and define $r_0 = R_0 = 0$. Thus, $S_n(d_i) = (n - R_i)/n,\ \ 0 \leq i \leq m$, and $M_n(0) = \bar{d}_n = \sum_{i=1}^{m} r_i d_i/n$. Note that M_n is right continuous, jumps up at each d_i for $1 \leq i \leq m-1$, and $M_n'(t) = -1$ between jump points. (Note that $M'(t) \geq -1$ where it exists since $M(t) + t$ is increasing.) Thus, T_n will be realized at one of the d_i^-'s and (Throughout, we integrate step functions by adding areas of horizontal strips.)

$$\begin{aligned}
T_n &= \max_{1\leq i\leq m} S_n(d_i^-)[M_n(0) - M_n(d_i^-)] \\
&= \max_{1\leq i\leq m} \left[\frac{n - R_{i-1}}{n}\bar{d}_n - \int_{d_i}^{d_m} S_n(u)du\right] \\
&= \max_{1\leq i\leq m} \left[\frac{n - R_{i-1}}{n}\bar{d}_n - \sum_{j=i+1}^{m} r_j(d_j - d_i)/n\right] \\
&= n^{-1} \max_{1\leq i\leq m} \left[(n - R_{i-1})\bar{d}_n - \left(n\bar{d}_n - \sum_{j=1}^{i} r_j d_j\right) + (n - R_i)d_i\right] \\
&= n^{-1} \max_{1\leq i\leq m} \left[\sum_{j=1}^{i} r_j d_j - R_{i-1}\bar{d}_n + (n - R_i)d_i\right].
\end{aligned} \tag{8}$$

Our scale invariant test statistic is $T_n^* = T_n/\bar{X}_n$. Since $\bar{d}_n \to \mu$ a.s., this normalization is equivalent to testing

$$H_0 : S(t) = e^{-t} \quad \text{vs.} \quad H_1 : S \text{ is NBUE with } M(0) = 1, \text{ but not exponential.} \tag{9}$$

2.1. Asymptotic Distribution of $\sqrt{n}T_n^*$

The following theorem provides the asymptotic distribution of $\sqrt{n}T_n^*$ under H_0 in (9).

Theorem 2.1 *Assume that* X *has a finite variance. Under H_0 in (9), when $S(t) = e^{-t}$, we have*

$$\sqrt{n}T_n^* \xrightarrow{d} \sup_{0\le u\le 1} B(e^{-u}),$$

where B *is a standard Brownian bridge.*

Proof. Let $Z_n = \sqrt{n}[M_n - M]$ and let $\sigma^2(t) = Var(X-t|X > t)$ be the conditional variance of X given survival till time t; $\sigma^2(0)$ is simply the variance of X. Yang (1978) and Hall and Wellner (1979) showed that

$$\frac{Z_n S_n}{\psi_n} \xrightarrow{w} W \circ U \quad \text{on } [0, \tau), \tag{10}$$

where ψ_n is any consistent estimator of $\sigma(0)$, W is a standard Brownian motion, $f \circ g$ denotes the composition of f and g, $U(t) = \dfrac{S(t)\sigma^2(t)}{\sigma^2(0)}$ that can be shown to be a SF, and τ is the right endpoint of the support of S that could be ∞. Now assume that H_0 holds. Then $S(t) = e^{-t}$ so that we can and do choose ψ_n to be the sample mean $\bar{X}_n = M_n(0)$, $Z_n(t) = \sqrt{n}[M_n(t) - M(0)]$, and $U(t) = S(t)$ for $t \ge 0$, and $\tau = \infty$. From (10),

$$\begin{aligned}
&\sqrt{n}\frac{S_n(t)}{\bar{X}_n}[M_n(t) - M(0)] \xrightarrow{w} W(e^{-t}), \quad t \in [0, \infty), \Longrightarrow \\
&\sqrt{n}\frac{1}{\bar{X}_n}[M_n(0) - M(0)] \xrightarrow{d} W(1) \Longrightarrow \\
&\sqrt{n}\frac{S_n(t)}{\bar{X}_n}[M_n(t) - M_n(0)] = \sqrt{n}\frac{S_n(t)}{\bar{X}_n}[[M_n(t) - M(0)] - [M_n(0) - M(0)]] \\
&\xrightarrow{w} W(e^{-t}) - e^{-t}W(1), \quad t \in [0, \infty),
\end{aligned} \tag{11}$$

the last convergence following from the fact that $S_n \xrightarrow{a.s.} S$ uniformly. Using the continuous mapping theorem, the facts that $W \stackrel{d}{=} -W$ and $B \stackrel{d}{=} -B$, and the fact that

$$\{W(u) - uW(1) : 0 \le u \le 1\} \stackrel{d}{=} \{B(u) : 0 \le u \le 1\},$$

we have

$$\begin{aligned}
\sqrt{n}T_n^* &= \sup_{t\ge 0} \frac{S_n(t)}{\bar{X}_n}[M_n(0) - M_n(t)] \xrightarrow{d} \sup_{t\ge 0}[W(e^{-t}) - e^{-t}W(1)] \\
&\stackrel{d}{=} \sup_{0\le u\le 1} B(u).
\end{aligned}$$

□

Yang (1978) showed that M_n is strongly uniformly consistent on $[0, b]$ for all $b < \tau$. Thus, under H_1, $\sup_{t\ge 0} \sqrt{n}[M(0) - M_n(t)] = \infty$. From the derivation of (11), it is clear that the test is consistent.

The asymptotic p-values of our test are given by the well known formula,

$$P(\sup_{0\le t\le 1} B(t) > t) = e^{-2t^2}, \quad t \ge 0.$$

The critical values for small to moderate sample sizes were obtained by simulations with 10^5 replications. The results are given in Table 1. The table shows that the critical values are close to the asymptotic values even for small sample sizes.

3. Comparison of Powers

In this section we compare the tests D_n^*, K_n^*, γ_{1n}^* and T_n^* for some chosen alternatives. Let $\mathcal{M}$ be the class of NBUE SFs with $M(0) = 1$. Following Birnbaum (1953), Chapman (1958), Doksum (1966), and others, we consider the subclass of alternatives $\mathcal{M}_\Delta$ that are at a fixed Kolmogorov distance $\Delta > 0$ from the null, i.e., $\mathcal{M}_\Delta = \{M \in \mathcal{M} : \sup_{t\ge 0}[M(0) - M(t)] = \Delta\}$. It can be seen that $\mathcal{M}_\Delta$ is a closed (in the sup norm) convex set. It contains the smaller set $\{M \in \mathcal{M}_\Delta : M \text{ is DMRL}\}$, that is also a closed convex set, and it contains the even smaller set of IFR distributions in $\mathcal{M}_\Delta$. The set $\mathcal{M}_\Delta^h$, defined by those M of the form (the corresponding SFs are given in square brackets)

$$M(t)\ [S(t)] = \begin{cases} 1 \quad [e^{-t}] & \text{if } 0 \le t < h \\ 1 - t + h \quad [e^{-h}] & \text{if } h \le t < h + \Delta \\ 1 \quad [(1-\Delta)e^{-t+\Delta}] & \text{if } t \ge h + \Delta \end{cases} \tag{12}$$

for $h \ge 0$, are extremes of $\mathcal{M}_\Delta$ in the sense that they cannot be expressed as a convex combination of two distinct M's in $\mathcal{M}_\Delta$. This can be seen from the fact that, for a given h, an $M \in \mathcal{M}_\Delta$ cannot be more than 1 on $[0, h] \cup [h+\Delta, \infty)$ and it cannot be less than $1-t+h$ for $h < t < h + \Delta$ since $M'(t) \ge -1$. We compare the simulated powers of the various tests against these alternatives. We also compare them against the DMRL alternatives with continuous SFs, given by (SFs in square brackets)

$$M(t)\ [S(t)] = \begin{cases} 1 \quad [e^{-t}] & \text{if } 0 \le t \le h \\ 1 - t + h \quad [e^{-h}] & \text{if } h \le t \le h + \Delta \\ 1 - \Delta \quad [e^{-\frac{t-\Delta(1+h)}{1-\Delta}}] & \text{if } t \ge h + \Delta \end{cases}. \tag{13}$$

To make these comparisons, we have to first derive the the test statistics for the other tests when there are ties since these are not known. We use the same notation used in Sec. 2.

For the HP test,

$$\begin{aligned} K_n &= \int_0^\infty S_n(t)\,[M_n(0) - M_n(t)]dF_n(t) \\ &= n^{-1}\sum_{i=1}^{m-1} S_n(d_i)[\bar{d}_n - M_n(d_i)]r_i \\ &= n^{-1}\sum_{i=1}^{m-1}[(n - R_i)\bar{d}_n/n - \int_{d_i}^{d_m} S_n(u)du]r_i \\ &= n^{-2}\sum_{i=1}^{m-1}[(nr_i - R_i r_i)\bar{d}_n - \sum_{j=i+1}^{m} r_j(d_j - d_i)r_i] \end{aligned}$$

$$= n^{-2}\left[(n^2 - \sum_{i=1}^{m} R_i r_i)\bar{d}_n - \sum_{j=2}^{m}\sum_{i=1}^{j-1} d_j r_j r_i + \sum_{i=1}^{m-1}\sum_{j=i+1}^{m} d_i r_i r_j\right.$$

$$= n^{-2}\left[(n^2 - \sum_{i=1}^{m} R_i r_i)\bar{d}_n - \sum_{i=1}^{m} d_i r_i R_{i-1} + \sum_{i=1}^{m} d_i r_i (n - R_i)\right.$$

$$= n^{-2}\left[(n^2 - \sum_{i=1}^{m} R_i\ r_i)\ \bar{d}_n + \sum_{i=1}^{m}(n - R_i - R_{i-1}) r_i d_i\right]. \tag{14}$$

When there are no ties, i.e., $m = n,\ r_i = 1,\ R_i = i,\ d_i = X_i$ for $1 \le i \le n$, and $\bar{d}_n = \bar{X}_n$, it reduces to K_n in (2).

For the AM test,

$$\gamma_{1n} = -M_n(0)/2 + \int_0^\infty S_n^2(t)dt$$

$$= \sum_{i=1}^{m} d_i[-r_i/(2n) + S_n^2(d_{i-1}) - S_n^2(d_i)]$$

$$= \frac{1}{n^2}\sum_{i=1}^{m} r_i d_i (3n/2 - R_i - R_{i-1}), \tag{15}$$

that reduces to γ_{1n} of (5) when there are no ties. AM (2009) suggested the test based on γ_j in (4), where j is chosen by simulations for the best power, instead of choosing $j = 1$ as we have. For simulations, they chose the parametric families of Gamma and Weibull with shape parameters $\theta > 1$, and the linear failure rate with shape parameter $\theta > 0$; these are all IFR. Even assuming that the alternative resembles one of these parametric IFR families, it is not possible to determine which family to choose for a given data set in a nonparametric setting. Moreover, if we suspect that the distribution is IFR, then the assumption of the vastly enlarged model of the NBUE class becomes moot.

For Koul's test,

$$D_n^* = \max_{1\le i\le n}\left[\frac{\int_0^{d_i} F_n(t)dt}{\bar{d}_n} - F_n(d_i)\right]$$

$$= \max_{1\le i\le n}\left[\frac{\sum_{j=1}^{m}(d_j \wedge d_i) r_j}{n\bar{d}_n} - \frac{R_i}{n}\right]$$

$$= \max_{1\le i\le n}\left[\frac{\sum_{j=1}^{i} r_j d_j + (n - R_i) d_i}{n\bar{d}_n} - \frac{R_i}{n}\right], \tag{16}$$

that reduces to (7) when there are no ties. By comparing (8) and (16), we note that $T_n^* - D_n^* = 1/n$ when there are no ties. Koul (1978) showed that D_n^* is consistent against NBUE alternatives and that $\sqrt{n}D_n^*$ has the same asymptotic distribution as $\sqrt{n}T_n^*$ under H_0.

We compare the rejection rates for the four tests with the alternatives given by (12), (13), and Weibull $(1, \theta)$ for selected values of the parameters with sample sizes 30 and 50, each based on 10^5 trials. We used the critical values for the AM test from the table given by

Basu and Anis (2012), for the HP test by subtracting $\sqrt{3/n}$ from those for the AM test, for our test from Table 1, and for Koul's test by subtracting $\sqrt{1/n}$ from those in Table 1. Before presenting the power comparisons, we would like to point out that the critical values of the sup-tests are much closer to the asymptotic values for small and moderate sample sizes than the integral tests; the comparison of the AM test and our test is presented in Table 2. We believe the main reason for the difference is the very jagged nature of the empirical MRL that jumps up at each order statistic, even making $M_n(0) - M_n(t) < 0$ at times.

Table 3 presents the simulated powers for the alternatives in (12) for the selected values of h and Δ. It is clear that T_n^* is the hands down winner; γ_{1n}^*, although not very good, beats K_n^* soundly; and D_n^* is a disappointing last. The integral tests perform poorly because $M(0) - M(t) > 0$ only for a small range of t's. The following analysis shows why the HP test performs so poorly compared to the AM test. From (14) and (15), we can rewrite

$$K_n^* = \frac{1}{n^2\bar{d}_n}\sum_{i=1}^{m}(n - R_i - R_{i-1})r_i d_i + 1 - \frac{1}{n^2}\sum_{i=1}^{m} R_i r_i$$

$$\gamma_{1n}^* = \frac{1}{n^2\bar{d}_n}\sum_{i=1}^{m}(n - R_i - R_{i-1})r_i d_i + \frac{1}{2}, \tag{17}$$

and note that

$$\gamma_{1n}^* - K_n^* = \frac{1}{n^2}\sum_{i=1}^{n} i - \frac{1}{2} = \frac{1}{2n}$$

when there are no ties, as shown earlier. Now, assume that M is of the form given in (12). Note that $S((h+\Delta)^-) = e^{-h}$ and $S(h+\Delta) = (1-\Delta)e^{-h}$, implying the jump in the SF is Δe^{-h} at $h+\Delta$. Thus we should expect approximately $s \equiv n\Delta e^{-h}$ observations at $h+\Delta$, and there should be no other ties. For a fixed n, assuming there is an observation at $h+\Delta$, let (the n-dependent) $d_{i_0} = h+\Delta$. Then $R_{i_0} = i_0 - 1 + r_{i_0}$, $r_{i_0}/n = \hat{s}_n$, the empirical estimate of s, and

$$\begin{aligned}\frac{1}{n^2}\sum_{i=1}^{m} R_i r_i - \frac{1}{n^2}\sum_{i=1}^{n} i &= \frac{1}{n^2}\left[R_{i_0} r_{i_0} - \sum_{j=i_0}^{i_0+r_{i_0}-1} j\right] \\ &= \frac{1}{n^2}\left[r_{i_0}((r_{i_0}-1)+i_0) - i_0 r_{i_0} - \frac{r_{i_0}(r_{i_0}-1)}{2}\right] \\ &= \frac{r_{i_0}(r_{i_0}-1)}{2n^2} = \frac{\hat{s}_n^2}{2} - \frac{\hat{s}_n}{2n}.\end{aligned}$$

Thus,

$$\begin{aligned}\gamma_{1n}^* - K_n^* &= \frac{1}{n^2}\sum_{i=1}^{m} R_i r_i - \frac{1}{2} \\ &= \frac{1}{n^2}\sum_{i=1}^{m} R_i r_i - \frac{1}{n^2}\sum_{i=1}^{n} i + \frac{1}{n^2}\sum_{i=1}^{n} i - \frac{1}{2} \\ &= \frac{\hat{s}_n^2}{2} - \frac{\hat{s}_n}{2n} + \frac{1}{2n}.\end{aligned}$$

When multiplied by $\sqrt{12n}$, the extra difference from $1/(2n)$, that occurs under H_0, could be huge. The somewhat surprisingly poor performance of Koul's sup-test stems from the fact that, couched in terms of MRLs, his test statistic can be seen to be $D_n^* = \max_{1\le i\le m} S_n(d_i)[M_n(0) - M_n(d_i)]$ while our test statistic is $T_n^* = \sup_{t\ge 0} S_n(t)[M_n(0) - M_n(t)]$ with the supremum being at some d_i^-. Since M_n always jumps up at each d_i, D_n^* could be much smaller than T_n^*. Although we have not considered testing for new worse than used in expectation in this paper, a similar argument will show that the parallels of our test and Koul's test for this problem will be identical.

In Table 4, we present the rejection rates for some of the DMRL alternatives in (13). Since D_n^* and K_n^* differ by the constant $1/n$ for continuous SFs, these tests are equivalent. The sup-tests still outperform the integral tests substantially. Since the SFs are continuous, the advantage of the AM test over the HP test is not very much.

In Table 5, we consider the IFR alternatives, $Weibull\ (1,\theta)$, for selected values of $\theta > 1$. As expected, the sup-tests are essentially equivalent (the slight differences coming from the fact that the suprema in the two tests may occur at different locations), the integral

Table 4
Percent rejected for the DMRL alternatives in (13)

	$n = 30$				$n = 50$			
(h, Δ)	K_n^*	γ_{1n}^*	D_n^*	T_n^*	K_n^*	γ_{1n}^*	D_n^*	T_n^*
$(Q_{0.00} = 0, 0.2)$	27.48	27.06	38.56	38.56	38.54	38.90	56.30	56.30
$(Q_{0.00} = 0, 0.3)$	46.19	45.67	67.09	67.09	63.64	63.96	86.61	86.61
$(Q_{0.00} = 0, 0.4)$	64.83	64.38	87.92	87.92	83.10	83.31	98.03	98.03
$(Q_{0.00} = 0, 0.5)$	79.56	79.22	97.12	97.12	93.71	93.82	99.87	99.87
$(Q_{0.10} = 0.105, 0.2)$	74.52	74.07	87.89	87.89	63.21	63.56	77.50	77.50
$(Q_{0.10} = 0.105, 0.3)$	91.79	91.59	98.56	98.56	98.87	98.89	99.97	99.97
$(Q_{0.10} = 0.105, 0.4)$	91.79	91.59	98.55	98.55	98.87	98.89	99.97	99.97
$(Q_{0.10} = 0.105, 0.5)$	98.22	98.17	99.92	99.92	99.92	99.92	100.0	100.0
$(Q_{0.25} = 0.288, 0.2)$	61.76	61.20	66.27	66.27	82.94	83.24	97.63	97.63
$(Q_{0.25} = 0.288, 0.3)$	93.21	93.00	99.72	99.72	99.43	99.44	100.0	100.0
$(Q_{0.25} = 0.288, 0.4)$	99.76	99.76	100.0	100.0	99.99	99.99	100.0	100.0
$(Q_{0.25} = 0.288, 0.5)$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 5
Percent rejected for the $Weibull\ (1, \theta)$ alternatives

	$n = 30$				$n = 50$			
θ	K_n^*	γ_{1n}^*	D_n^*	T_n^*	K_n^*	γ_{1n}^*	D_n^*	T_n^*
$\theta = 1.1$	13.22	13.71	6.79	6.79	18.98	18.34	10.66	10.66
$\theta = 1.2$	27.88	28.91	15.11	15.11	42.42	42.48	26.24	26.70
$\theta = 1.3$	46.58	47.90	27.23	27.33	68.92	70.82	48.48	48.92
$\theta = 1.4$	66.25	68.44	42.44	42.67	87.90	88.98	69.56	69.72
$\theta = 1.5$	81.67	81.78	58.46	58.52	96.33	96.74	85.01	85.11

tests are essentially equivalent, and the integral tests are somewhat superior to the sup-tests. However, one should note that the DMRL alternative in (13) with $h = 0$ is actually an IFR alternative (the failure rate is 0 on $[0, \Delta)$ and $1 - \Delta$ on $[\Delta, \infty)$). Thus, the superiority of the integral tests does not hold for all IFR alternatives.

4. Examples

Example 4.1. This data set consists of $n = 27$ observations of the intervals between successive failures of the air-conditioning systems of 7913 jet airplanes of a fleet of Boeing 720 jet airplanes as reported in Proschan (1963). The computed value of the statistic and p-value are 0.6211 and 0.4623, respectively. It is clear that the null hypothesis of exponentiality will be accepted and which agrees with the conclusions of Anis and Hoque (2008) and Anis and Mitra (2012).

Data set: 1,4,11,16,18,18,18, 24,31,39,46,51,54,63,68,77,80,82,97,106,111, 141,142, 163,191,206,216.

Example 4.2. This data is reported in Bryson and Siddiqui (1969). It represents the survival times, in days from diagnosis, of patients suffering from chronic granulocytic leukemia. The computed value of the statistic and p-value are 1.2742 and 0.0389, respectively. We observe that the null hypothesis of exponentiality is rejected and which agrees with the conclusions of Hollander and Proschan (1975) and Anis and Mitra (2012).

Data Set: 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.

5. Concluding Remarks

There is a large body of literature on testing exponentiality against NBUE alternatives. So far, power studies of these tests have been confined to IFR alternatives, a very narrow subset of the NBUE alternatives. In this paper we have initiated a study of powers for much broader sets of alternatives, including ones with discontinuous SFs. We have also introduced a new KS-type sup-test that is much more powerful against large classes of alternatives than some of the best tests known. Also, it is not much worse than the integral tests for IFR alternatives.

References

Anis, M. Z., Basu, K. (2012). Tests for exponentiality against NBUE alternatives: a Monte Carlo comparison; http://arxiv.org/pdf/1207.7311.pdf.

Anis, M. Z., Hoque, Z. (2008). On the performance of a new test of exponentiality against IFR alternatives. *Econ. Qual. Control* 23:181–195.

Anis, M. Z., Mitra, M. (2011). A generalized Hollander-Proschan type test of exponentiality against NBUE alternatives. *Statist. Probab. Lett.* 81(1):126–132.

Bandyopadhyay, D., Basu, A. P. (1990). A class of tests for exponentiality against decreasing mean residual life alternatives. *Commun. Statist. Theor. Meth.* 19:905–920.

Barlow, R. E. (1968). Likelihood ratio tests for restricted families of probability distributions. *Ann. Mathemat. Statist.* 39:547–560.

Basu, K., Anis, M. Z. (2012). The exact null distribution of the generalized Hollander-Proschan type test for NBUE alternatives; http://arxiv.org/pdf/1207.6763.pdf.

Birnbaum, Z. W. (1953). On the power of one-sided tests of fit for continuous distribution functions. *Ann. Mathemat. Statist.* 24:284–289.

Bjerkedal, T. (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Amer. J. Hygiene* 72:130–148.

Bryson, M. C., Siddiqui, M. M. (1969). Some criteria for aging. *J. Amer. Statist. Assoc.* 64:1472–1483.

Chapman, D. G. (1958). A comparative study of several one-sided goodness-of-fit tests. *Ann. Mathemat. Statist.* 29:655–674.

Doksum, K. (1966). Asymptotically minimax distribution-free procedures. *Ann. Mathemat. Statist.* 37:619–628.

Hall, W. J., Wellner, J. A. (1979). Estimation of mean residual life. Unpublished manuscript.

Hollander, M., Proschan, F. (1972). Testing whether new is better than used. *Ann. Mathemat. Statist.* 43:1136–1146.

Hollander, M., Proschan, F. (1975). Tests for the mean residual life. *Biometrika* 62:585–593.

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

Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics* 5:375–383.

Yang, G. L. (1978). Estimation of a biometric function. *Ann. Statist.* 6:112–116.