© Statistical Research and Training Center

مجلهٔ پژوهش‌های آماری ایران

J. Statist. Res. Iran **10** (2013): 147–161

دوره‌ی ۱۰، شماره‌ی ۲، پاییز و زمستان ۱۳۹۲، صص ۱۴۷–۱۶۱

Test for Exponentiality Based on the Sample Covariance

Narges H. Montazeri and Hamzeh Torabi*

Yazd University

Abstract. This paper proposes a simple goodness-of-fit test based on the sample covariance. It is shown that this test is preferable for alternatives of increasing and unimodal failure rate. Critical values for various sample sizes are determined by means of Monte Carlo simulations. We compare the test based on the sample covariance with tests based on Hoeffding's maximum correlation. The usefulness of the proposed test is shown for a real example. An empirical power study shows that the new test has the same level or upper level of performance than the best exponentiality tests in the statistical literature.

Keywords. Hoeffding's maximum correlation; goodness-of-fit test; Monte Carlo simulation; sample covariance; test for exponentiality.

MSC 2010: 62F03, 62F10.

1 Introduction

The exponential distribution is probably one of the most used in statistical researches after the normal distribution. It often appears in problems dealing with life testing and reliability. Hence goodness-of-fit testing for exponentiality is an attractive problem.

Let $X_1, \ldots, X_n$ be a random sample from a continuous population with cumulative distribution function (CDF), $F(\cdot)$, and probability density function (PDF), $f(\cdot)$. Based on the observations of the random sample, i.e.,

* Corresponding author

$x_1, \ldots, x_n$, we want to test

$$\begin{cases} H_0: \; F = F_0 \\ H_1: \; F \neq F_0, \end{cases} \qquad (1)$$

where F_0 is the CDF of the exponential distribution with unknown scale parameter $\theta > 0$, $F_0(x;\theta) = 1 - e^{-x/\theta}$, $x > 0$.

Van-Soest (1969), Finkelstein and Schafer (1971), Ebrahimi et al. (1992, 1994), Corea (1995), Choi et al. (2004), Gurevich and Davidson (2008), Alizadeh and Arghami (2011, 2011a, 2011b), Abbasnejad (2011), Abbasnejad et al. (2012), Fortiana and Grane (2002), Grane and Fortiana (2009) and many authors proposed different test statistics for exponentiality. For the first time, Cuadras and Fortiana (1993) introduced a goodness-of-fit test based Hoeffding's maximum correlation.

Let F_1 and F_2 be two CDF's having second-order moments. The Hoeffding's maximum correlation $\rho^+(F_1, F_2)$ is defined as

$$\rho^+(F_1, F_2) = \frac{\int_0^1 F_1^-(p)F_2^-(p)dp - \mu_1\mu_2}{\sigma_1\sigma_2},$$

where F_i^- is the is the left-continuous pseudo-inverse of F_i, $\mu_i = E(F_i)$ and $\sigma^2 = \mathtt{Var}(F_i)$, $i = 1, 2$. Here the notation $E(F)$ represents the expected value of any random variable whose CDF is F, and analogously for $\mathtt{Var}(F)$. $\rho^+(F_1, F_2)$ is a measure of agreement between F_1 and F_2, since it equals 1 if $F_1 = F_2$.

Cuadras and Fortiana (1993) proposed the statistic $\rho^+(F_n, F_0)$ as a qualitative measure of goodness-of-fit of a sample, with empirical CDF F_n, to a given distribution F_0.

Fortiana and Grane (2002) implemented this idea and proposed following statistic for exponentiality test based on the Hoeffding's maximum correlation:

$$Q_n = \frac{s_n}{\bar{x}_n}\rho^+(F_n, F_0) = \frac{\sum_{i=1}^n l_i x_{(i)}}{\sum_{i=1}^n x_{(i)}}$$

where $\bar{x}_n$, s_n^2 and F_n are the empirical mean, variance and CDF, respectively, of the observed sample and $l_i = (n-i)\log(n-i) - (n-i+1)\log(n-i+1) + \log(n)$, $i = 1, \ldots, n$, with $0 \log 0 = 0$. The H_0 in (1) is rejected for small or large values of the Q_n.

Also Grane and Fortiana (2009) proposed following statistic for exponentiality test based on the Hoeffding's maximum correlation:

$$Q_n^* = \frac{s_n \rho^+(F_n, F_0)}{\frac{1}{n}\sum_{i=1}^n b_i x_{(i)}} = \frac{\sum_{i=1}^n l_i x_{(i)}}{\sum_{i=1}^n b_i x_{(i)}}$$

where $b_i = \{i/n\} - \{(n+1)/(2n)\}$.

In this paper, we implement this idea based on the sample covariance between the empirical distribution F_n and a given distribution F_0 in the form of a specific test. Our aim is to obtain a very simpler and powerful test for exponetiality using this idea.

The paper is organized as follows:
In Section 2, we introduce a new statistic of goodness-of-fit tests for exponentiality based on sample covariance. Then we discuss some of their features. To facilitate comparisons of the power of the present test with the powers of the previous tests, we select two series of alternatives listed in Ebrahimi et al. (1992) and Henz and Mentains (2005). The results of a simulation study are presented in Section 3. The last section includes some conclusions.

2 Definition of the Test Statistic

Most goodness-of-fit test statistics can be interpreted as measures of proximity between two distributions: empirical and hypothesized distributions. Consider the goodness-of-fit testing problem (1) based on a random sample $X_1, \ldots, X_n$.

Our test is based on the sample covariance between the empirical distribution function (EDF) of the scaled data $Y_i = X_i/\hat{\theta}$,

$$F_n(y) = \frac{1}{n}\sum_{i=1}^n \mathrm{I}(Y_i \leqslant y),$$

and the CDF of the standard exponential distribution,

$$F_0(Y_{(i)}) = 1 - e^{-Y_{(i)}} =: Z_{(i)},$$

$i = 1, \ldots, n$, where $\hat{\theta}$ is the maximum likelihood estimate (MLE), $\hat{\theta} = \bar{X}_n$ and $\mathrm{I}_A(x)$ denotes the indicator function of an event A. Also $X_{(i)}$, $Y_{(i)}$ and $Z_{(i)}$ are order statistics of X_i, Y_i and Z_i, respectively. Set $\boldsymbol{y} = (y_{(1)}, \ldots, y_{(n)})$, $F_0(\boldsymbol{y}) := \big(F_0(y_{(1)}), \ldots, F_0(y_{(n)})\big)$ and $F_n(\boldsymbol{y}) := \big(F_n(y_{(1)}), \ldots, F_n(y_{(n)})\big)$.

Define

$$\begin{aligned}
\mathtt{H}_n &= s(F_0(\boldsymbol{y}), F_n(\boldsymbol{y})) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(z_{(i)}\frac{i}{n}\right) - \left(\frac{1}{n}\sum_{i=1}^{n} z_{(i)}\right)\cdot\left(\frac{1}{n}\sum_{i=1}^{n}\frac{i}{n}\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} i z_{(i)} - \frac{(n+1)\bar{z}}{2n} \qquad (2)
\end{aligned}$$

where $s(F_0(\boldsymbol{y}), F_n(\boldsymbol{y}))$ is the sample covariance of $F_0(\boldsymbol{y})$ and $F_n(\boldsymbol{y})$. Note that the statistic $\mathtt{H}_n$ is non-negative, since F_0 and F_n are non-decreasing and $\mathbf{y}$ is an increasing vector. Also, using the Cauchy-Schwartz inequality,

$$\mathtt{H}_n \leqslant s_1 s_2,$$

where s_1 and s_2 are the sample standard deviations of $F_0(\mathbf{y})$ and $F_n(\mathbf{y})$, respectively.

Under the null hypothesis, we expect a large value for H_n near to its upper bound, i.e., s_1s_2. Conversely, if H_n is near to s_1s_2, we expect a linear relation between two vectors $F_0(\boldsymbol{y})$ and $F_n(\boldsymbol{y})$, say $F_0(\boldsymbol{y}) = a + bF_n(\boldsymbol{y})$. But $F_0(0) = F_n(0) = 0$ and $F_0(\infty) = F_n(\infty) = 1$, thus $a = 0$ and $b = 1$, i.e., the null hypothesis is concluded. Hence, it is justifiable that H_0 versus H_1 must be rejected if $\mathtt{H}_n$ has a small value. Therefore, we reject H_0 in favour of H_1 at the significance level α if $\mathtt{H}_n \leqslant q_n(\alpha)$, where the critical point $q_n(\alpha)$ is determined by the left α-quantile of the distribution of the $\mathtt{H}_n$-statistic under null hypothesis.

Since the statistic $\mathtt{H}_n$ considered for goodness-of-fit test (1), is a function of the scaled observations Y_i or their transformed values $Z_i = 1 - \exp(Y_i)$, $i = 1, \ldots, n$, it is scale invariant. As a consequence, the null distribution of $\mathtt{H}_n$ does not depend on the parameter θ.

In order to obtain the critical points of the test, we use a Monte Carlo simulation with $B = 10000$ replications:
For given n, first generate x_i from the standard exponential distribution and compute y_i for $i = 1, \ldots, n$. Then compute the statistics $\mathtt{H}_n$. Repeat these steps B times and compute $\mathtt{H}_{n_j}$ for $j = 1, 2, \ldots, B$. Sort the computed $\mathtt{H}_{n_j}$, values and then determine the order $[\alpha B]$ of $\mathtt{H}_{n_j}$. Table 1 gives the critical values of $\mathtt{H}_n$ for some various sample sizes. Note that the critical values do not depend on the unknown parameter, because we explained that $\mathtt{H}_n$ under null hypothesis is invariant under scale transformation of the observations.

Table 1. Critical values of H_n for $\alpha = 0.01$ and $\alpha = 0.05$.

n	5	6	7	8	9	10	15	20	25	30	35	40
$q_n(.01)$	.026	.031	.038	.040	.043	.047	.055	.060	.063	.065	.066	.068
$q_n(.05)$	.039	.044	.048	.052	.054	.056	.063	.066	.068	.070	.071	.072

3 Power Comparison

To compare the power of our proposed test, we consider the following tests:

1. van-Soest (1969) proposed the test statistics W^2 of the form

$$\mathrm{W}^2 = \frac{1}{12n} + \sum_{j=1}^{n} \left\{ Z_{(j)} - \frac{2j-1}{2n} \right\}^2 .$$

2. Finkelstein and Schafer (1971) introduced the statistics S of the form

$$\mathrm{S} = \sum_{i=1}^{n} \max \left\{ \left| Z_{(i)} - \frac{i}{n} \right|, \left| Z_{(i)} - \frac{i-1}{n} \right| \right\} .$$

3. Ebrahimi et al. (1992) introduced the statistics TV_{mn} of the form

$$\mathrm{TV}_{mn} = \frac{\exp\left(\mathrm{HV}_{mn}\right)}{\exp\{\ln(\bar{X}) + 1\}},$$

where

$$\mathrm{HV}_{mn} = \frac{1}{n} \sum_{i=1}^{n} \ln \left[\frac{n}{2m} \left\{ X_{(i+m)} - X_{(i-m)} \right\} \right], \qquad (3)$$

is Vasicek entropy estimator, Vasicek (1976), and window size m is a positive integer smaller than $n/2$ and $X_{(i)} = X_{(1)}$ if $i < 1$ and $X_{(i)} = X_{(n)}$ if $i > n$.

4. Correa (1995) introduced the statistics TC_{mn} of the form

$$\mathrm{TC}_{mn} = \frac{\exp\left(\mathrm{HC}_{mn}\right)}{\exp\{\ln(\bar{X}) + 1\}},$$

where

$$\mathrm{HC}_{mn} = \frac{1}{n} \sum_{i=1}^{n} \ln \left[\frac{\sum_{j=i-m}^{i+m} \left\{ X_{(j)} - \tilde{X}_{(i)} \right\} (j-i)}{n \sum_{j=i-m}^{i+m} \left\{ X_{(j)} - \tilde{X}_{(i)} \right\}^2} \right],$$

is Correa entropy estimator, Correa (1995) and $\tilde{X}_{(i)} = \sum_{j=i-m}^{i+m} \frac{X_{(j)}}{(2m+1)}$.

5. Correa (1995) introduced the statistics $\mathtt{TVE}_{mn}$ of the form

$$\mathtt{TVE}_{mn} = \frac{\exp\left(\mathrm{HVE}_{mn}\right)}{\exp\{\ln(\bar{X})+1\}},$$

where

$$\mathrm{HVE}_{mn} = \frac{1}{n-m} \sum_{i=1}^{n-m} \left[\frac{n+1}{m} \ln\left\{X_{(i+m)} - X_{(i)}\right\} + \sum_{k=m}^{n} \frac{1}{k} + \ln(m) - \ln(n+1) \right],$$

is Van Es entropy estimator, Van Es (1992).

6. Alizadeh and Arghami (2011) introduced the statistics $\mathtt{TA}_{mn}$ of the form

$$\mathtt{TA}_{mn} = \frac{\exp\left(\mathrm{HA}_{mn}\right)}{\exp\{\log(\bar{X})+1\}},$$

where

$$\mathrm{HA}_{mn} = -\frac{1}{n} \sum_{i=1}^{n} \log \left\{ \frac{\hat{f}(X_{(i+m)}) + \hat{f}(X_{(i-m)})}{2} \right\},$$

is Alizadeh and Arghami entropy estimator (2010), and $\hat{f}(x)$ is

$$\hat{f}(X_i) = \frac{1}{nh} \sum_{j=1}^{n} k\left(\frac{X_i - X_j}{h}\right).$$

The kernel function $k(\cdot)$ is chosen to be the standard normal density function and the bandwidth h is chosen to be the normal optimal smoothing formula, $h = 1.06sn^{-1.5}$, where s is the sample standard deviation.

7. Gurevich and Davidson (2008) introduced the statistics $\mathtt{MKL}_n^1$ of the form

$$\mathtt{MKL}_n^1 = \max_{1 \leqslant m \leqslant \frac{n}{2}} \left\{ \frac{n\left[\prod_{j=1}^{n}\{(X_{(i+m)} - X_{(i-m)}\}\right]^{\frac{1}{n}}}{2me\bar{X}} \right\}.$$

8. Gurevich and Davidson (2008) introduced the statistics MKL_n^2 of the form

$$\mathrm{MKL}_n^2 = \max_{1 \leqslant m \leqslant n^{1-\delta}} \left\{ \frac{n \left[\prod_{j=1}^{n} \{X_{(i+m)} - X_{(i-m)}\} \right]^{\frac{1}{n}}}{2me\bar{X}} \right\},$$

   where $0 < \delta < 1$. Set up $\delta = 0.5$ in the definition MKL_n^2.

9. Abbasnejad (2011) proposed a test statistics ED_r^V based on Renyi Information of the form

$$\mathrm{ED}_r^V = \log(\bar{X}) + \frac{1}{1-r} \log \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[\left\{ \frac{\frac{2m}{n}}{X_{(i+m)} - X_{(i-m)}} \right\} e^{\frac{X_{(i)}}{\bar{X}}} \right]^{r-1} \right\}.$$

10. Abbasnejad et al. (2012) proposed a test statistics based on Lin-Wong divergence measure. The test statistics L_V is defined as

$$\mathrm{L}_V = -\frac{1}{n} \sum_{i=1}^{n} \log \left[\frac{1}{2} + \frac{n}{4m\bar{X}} \left\{ X_{(i+m)} - X_{(i-m)} \right\} e^{\frac{-X_{(i)}}{\bar{X}}} \right],$$

   where $m = [\sqrt{n} + 0.5]$.

The exponentiality hypothesis of the data are rejected for small value of the TV_{mn}, TC_{mn} , TVE_{mn}, TA_{mn}, MKL_n^1, and MKL_n^2 and also for large value of the W^2, S, ED_r^V and L_V.

To facilitate comparisons of the power of the present test with the powers of the mentioned tests, we select two series of alternatives:

(i) Alternatives listed in Ebrahimi et al. (1992) and their choices of parameters: Gamma, Weibull, log-normal distributions.

(ii) Alternatives listed in Henz and Mentains (2005) and their choices of parameters: Gamma, Weibull, log-normal, Half-normal, Uniform, Chen, modified extreme value and Dhillon distributions.

3.1 Alternatives Listed in Ebrahimi et al. (1992)

We select the same three alternatives listed in Ebrahimi et al. (1992) and their choices of parameters which are:

- the Gamma distribution, $\mathtt{G}(\theta, \lambda)$, with density

$$f(x) = \frac{\lambda^{\theta} x^{\theta-1} e^{-\lambda x}}{\Gamma(\theta)}, \qquad x > 0;$$

- the Weibull distribution, $\mathtt{W}(\theta, \lambda)$, with density

$$f(x) = \theta \lambda^{\theta} x^{\theta-1} e^{-(\lambda x)^{\theta}}, \qquad x > 0;$$

- the Log-Normal distribution, $\mathtt{LN}(\nu, \theta)$, with density

$$f(x) = \frac{1}{x\sqrt{2\pi\theta^2}} \exp\left\{ -\frac{(\log x - \nu)^2}{2\theta^2} \right\}, \qquad x > 0.$$

We also choose the parameters so that $E(X) = 1$, i.e, $\lambda = \theta$ for the gamma, $\lambda = \Gamma(1 + 1/\theta)$ for the Weibull and $\nu = -\theta^2/2$ for the log-normal family of distributions.

For comparison of powers , under each alternative, we compute the powers of tests based on listed statistics in Section 3 by Monte Carlo simulations. Tables 2, 3 and 4 show the estimated powers at significance level $\alpha = 0.01$ and $\alpha = 0.05$ according to 10000 simulations of sizes 10 and 20, for alternatives gamma, Weibull and log-normal, respectively. For each alternative, bold type in theses tables indicate the maximal power. To compute the estimated powers of the new test, we use R software. Reported powers of the other tests are based on Ebrahimi (1992), Alizadeh and Arghami (2011) and Abbasnejad et al. (2012).

For the test statistics TV, TVE, TC and TA, the value of m which maximizes the power of the test for each alternative are given in parentheses. For the test statistic L_V, choosing of m is based on formula $m = [\sqrt{n} + 0.5]$.

Table 2. Monte Carlo power estimates against the gamma distribution

n	θ	α	TV(m)	TVE(m)	TC(m)	TA(m)	ED_r^V	MKL_n^1	MKL_n^2	L_V	W^2	S	Q_n^*	Q_n	H_n
10	2	.01	.120(5)	.086(5)	.116(5)	.122(3)	.068	.118	.103	**.137**	.069	.072	.013	.052	.119
		.05	.344(4)	.284(5)	.341(4)	.362(3)	.295	.324	.300	.365	.250	.253	.063	.190	**.363**
	3	.01	.332(5)	.247(5)	.329(5)	.345(3)	.245	.300	.266	.355	.202	.207	.020	.167	**.352**
		.05	.661(4)	.556(5)	.652(4)	.692(3)	.584	.601	.564	**.698**	.531	.544	.081	.424	**.698**
	4	.01	.562(5)	.420(5)	.558(5)	.563(3)	.434	.501	.455	.590	.378	.385	.027	.322	**.606**
		.05	.860(4)	.768(5)	.853(4)	.882(3)	.788	.790	.759	**.885**	.761	.757	.102	.631	.881
20	2	.01	.275(6)	.217(10)	.268(6)	**.360**(5)	.150	.281	.238	.342	.220	.223	.019	.139	.345
		.05	.560(8)	.510(10)	.557(8)	.646(5)	.421	.550	.493	.629	.464	.489	.086	.356	**.652**
	3	.01	.696(8)	.622(10)	.686(8)	**.817**(5)	.513	.707	.651	.791	.663	.669	.033	.465	.803
		.05	.909(7)	.876(10)	.887(7)	**.961**(5)	.817	.902	.870	.942	.887	.889	.131	.762	**.961**
	4	.01	.916(6)	.873(10)	.900(4)	**.972**(5)	.787	.921	.889	.958	.918	.911	.044	.750	.970
		.05	.989(6)	.981(10)	.983(6)	**.998**(6)	.962	.988	.981	.995	.989	.990	.165	.935	**.998**

Table 3. Monte Carlo power estimates against the Weibull distribution.

n	θ	α	TV(m)	TVE(m)	TC(m)	TA(m)	ED_r^V	MKL_n^1	MKL_n^2	L_V	W^2	S	Q_n^*	Q_n	H_n
10	2	.01	.422(5)	.289(5)	**.425**(5)	.421(3)	.320	.364	.326	**.425**	.255	.264	.038	.250	.395
		.05	.747(5)	.604(5)	.753(5)	.759(1)	.668	.662	.622	**.759**	.613	.641	.130	.546	.718
	3	.01	.905(5)	.774(5)	**.911**(5)	.904(3)	.832	.831	.798	.906	.765	.807	.082	.775	.879
		.05	.990(4)	.957(5)	.991(5)	.992(2)	.982	.962	.951	**.993**	.969	.974	.238	.958	.983
	4	.01	.995(5)	.964(5)	**.996**(5)	.995(3)	.985	.978	.968	.995	.970	.978	.113	.978	.989
		.05	1.00(4)	.998(5)	.999(5)	1.00(3)	1.00	.999	.998	1.00	.999	.999	.310	1.00	**1.00**
20	2	.01	.824(10)	.682(10)	.827(10)	**.860**(5)	.650	.789	.737	.854	.758	.776	.089	.681	.827
		.05	.962(9)	.893(10)	.958(9)	**.977**(1)	.901	.941	.918	.969	.926	.946	.283	.911	.960
	3	.01	.999(10)	.995(9)	.999(10)	**1.00**(2)	.996	.999	.998	1.00	1.00	1.00	.234	.999	.999
		.05	1.00(9)	1.00(10)	1.00(9)	1.00(1)	1.00	1.00	.999	1.00	1.00	1.00	.532	1.00	**1.00**
	4	.01	1.00(10)	1.00(10)	1.00(9)	1.00(1)	1.00	1.00	1.00	1.00	1.00	1.00	.241	.999	**1.00**
		.05	1.00(9)	1.00(10)	1.00(9)	1.00(1)	1.00	1.00	1.00	1.00	1.00	1.00	.535	1.00	**1.00**

Table 4. Monte Carlo power estimates against the log-normal distribution.

n	ν	α	TV(m)	TVE(m)	TC(m)	TA(m)	ED_r^V	MKL_n^1	MKL_n^2	L_V	W^2	S	Q_n^*	Q_n	H_n
10	-.3	.01	.091(3)	.083(5)	.090(5)	.097(3)	.073	.100	.090	**.117**	.055	.055	.019	.037	.114
		.05	.269(4)	.290(5)	.281(3)	.302(4)	.272	.298	.280	.317	.217	.255	.074	.133	**.355**
	-.2	.01	.254(4)	.205(5)	.246(5)	.267(3)	.186	.254	.227	.280	.148	.155	.015	.105	**.300**
		.05	.570(4)	.533(5)	.543(4)	.593(3)	.511	.554	.520	.606	.451	.456	.062	.289	**.668**
	-.1	.01	.736(5)	.633(5)	.718(5)	.775(3)	.637	.704	.657	.796	.604	.570	.017	.457	**.813**
		.05	.947(4)	.921(5)	.935(4)	.968(3)	.921	.918	.900	.961	.922	.903	.077	.736	**.980**
20	-.3	.01	.220(4)	.255(7)	.204(4)	.276(7)	.128	.236	.220	.278	.169	.161	.030	.066	**.340**
		.05	.482(4)	.568(5)	.463(3)	.569(8)	.397	.499	.490	.534	.406	.402	.100	.195	**.660**
	-.2	.01	.592(4)	.593(10)	.563(4)	.727(5)	.378	.621	.581	.679	.525	.499	.019	.251	**.781**
		.05	.837(4)	868(10)	.813(4)	.921(8)	.735	.848	.834	.872	.803	.784	.080	.490	**.954**
	-.1	.01	.988(5)	.987(10)	.982(4)	.999(6)	.938	.991	.987	.993	.993	.985	.028	.832	**1.00**
		.05	.999(4)	1.00(10)	1.00(4)	1.00(6)	.997	.999	.999	1.00	1.00	1.00	.111	.953	**1.00**

According to Tables 2, 3 and 4, in the cases of log-normal distributions, $\mathtt{H}_n$ is preferable. We also observe that the proposed test performs well compared with the other tests for gamma and weibull alternatives.

The goodness-of-fit test based on entropy involves choosing the best integer parameter m, where there is no choosing criterion for m and in general it depends on the alternatives, but our proposed test is very simple and is powerful for all alternatives.

3.2 Alternatives Listed in Henz and Mentains (2005)

To illustrate that the proposed test is sensitive with respect to other alternatives, we consider alternatives listed in Henz and Mentains (2005). Inspired by works of Henz and Mentains (2005), we have choosen the following families of distribution, defined either by its PDF or CDF:

- the Weibull distribution, $\mathtt{W}(\theta, 1)$ with density $f(x) = \theta x^{\theta-1} e^{-x^{\theta}}$, $x > 0$;
- the Gamma distribution, $\mathtt{G}(\theta, 1)$, with density $f(x) = \frac{x^{\theta-1}e^{-x}}{\Gamma(\theta)}$, $x > 0$;
- the Log-Normal distribution, $\mathtt{LN}(\theta, 1)$, with density

$$f(x) = \frac{1}{x\sqrt{2\pi\theta^2}} \exp\left\{-\frac{(\log x)^2}{2\theta^2}\right\}, \qquad x > 0.$$

- the half-Normal distribution, $\mathtt{HN}$, with density

$$f(x) = \sqrt{\frac{2}{\pi}} \exp\left\{\frac{-x^2}{2}\right\}, \qquad x > 0.$$

- the uniform distribution, $\mathtt{U}$, with density 1, $0 \leqslant x \leqslant 1$
- the Chen distribution, Chen (2000), $\mathtt{CH}(\theta)$, with CDF as

$$F(x) = 1 - \exp\left\{2\left(1 - e^{x^{\theta}}\right)\right\}, \qquad x > 0.$$

- the modified extreme value distribution, $\mathtt{EV}(\theta)$, with CDF as

$$F(x) = 1 - \exp\left\{\theta^{-1}\left(1 - e^{x}\right)\right\}, \qquad x > 0.$$

- the Dhillon distribution Dhillon (1981), $\mathtt{DL}(\theta)$, with CDF as

$$F(x) = 1 - \exp\left[-\{\log(x+1)\}^{\theta+1}\right], \qquad x > 0.$$

These distributions comprise widely used as alternatives to the exponential model and included densities f with decreasing failure rates (DFR), increasing failure rates (IFR) as well as models with unimodal (increasing-decreasing) failure rate (UFR) functions and bathtub (decreasing-increasing) failure rate (BFR) function.

The empirical scaled TTT (Total Time on Test) transform, introduced by Aarset (2004), can be used to identify the shape of the hazard function. The scaled TTT transform is convex (concave) if the hazard rate is decreasing (increasing), and for bathtub (unimodal) hazard rates, the scaled TTT transform is first convex (concave) and then concave (convex).

According to Table 5, in cases of IFR and UFR alternatives (except U(0,1), CH(1.5) and EV(1.5)), $\mathtt{H}_n$ is preferable.

The gamma and the Weibull distributions are DFR for $0 < \theta < 1$ and IFR for $\theta > 1$. Figure 1 contains the power curves of the 5% significance level test for exponentiality based on statistics Q_n and H_n for $n = 10$ for gamma and Weibull distributions. From Figure 1 we can see that $\mathtt{H}_n$ is more powerful than Q_n for $\theta > 1$. Therefore $\mathtt{H}_n$ is preferable for most IFR alternatives.

4 Conclusions

In this paper, we proposed a simple test for exponentiality and compared this new test with some tests for exponentiality. The usefulness of the new test illustrated through a simulation study. In the cases of gamma and Weibull distribution, the statistics $\mathtt{TC}(m)$ and $\mathtt{TA}(m)$ had maximal power, but these tests involve choosing the best integer parameter m, where there was no criterion for choosing m and in general it depended on the alternatives. We showed that, the new test had the same level or higher level of performance than the best tests in the statistical literature for alternatives of IFR and UFR. Therefore, the new test are serious competitors and powerful.

Table 5. Monte Carlo power estimates against the various distributions.

Dist.	α	Type Hazard	$n = 10$			$n = 20$		
			Q_n^*	Q_n	H_n	Q_n^*	Q_n	H_n
W(.8)	.01	DFR	.025	**.041**	.003	.030	**.071**	.000
	.05		.083	**.119**	.011	.101	**.187**	.006
W(1.4)	.01	IFR	.014	.30	**.069**	.027	.112	**.196**
	.05		.065	.152	**.263**	.104	.315	**.466**
G(.4)	.01	DFR	.058	**.177**	.000	.082	**.353**	.000
	.05		.153	**.342**	.010	.207	**.558**	.000
G(1)	.01		.010	.008	.010	.010	.010	.010
	.05		.049	.051	.050	.050	.050	.050
G(2)	.01	IFR	.014	.043	**.107**	.022	.156	**.337**
	.05		.061	.185	**.364**	.089	.384	**.663**
LN(.8)	.01	UFR	.022	.027	**.078**	.038	.060	**.274**
	.05		.074	.121	**.312**	.106	.175	**.611**
LN(1.5)	.01	DFR	.156	**.221**	.000	.298	**.470**	.012
	.05		.268	**.362**	.015	.455	**.613**	.032
HN	.01	IFR	.0167	.021	**.039**	.035	.073	**.083**
	.05		.070	.107	**.158**	.126	.234	**.252**
U(0,1)	.01	IFR	.138	**.144**	.130	.567	**.602**	.307
	.05		.363	**.420**	.349	.840	**.855**	.564
CH(.5)	.01	BFR	.042	**.118**	.000	.052	**.236**	.000
	.05		.116	**.267**	.002	.155	**.432**	.000
CH(1)	.01	IFR	.014	.012	**.027**	.027	.046	**.055**
	.05		.059	.077	**.115**	.115	.170	**.183**
CH(1.5)	.01	IFR	.041	.142	**.215**	.134	**.578**	.567
	.05		.154	.437	**.529**	.370	**.852**	.826
EV(.5)	.01	IFR	.013	.014	**.029**	.028	.047	**.052**
	.05		.064	.077	**.119**	.110	.169	**.184**
EV(1.5)	.01	UFR	.032	.057	**.086**	.108	**.246**	.194
	.05		.125	.220	**.263**	.315	**.535**	.440
DL(1)	.01	UFR	.013	.021	**.053**	.022	.050	**.161**
	.05		.066	.103	**.221**	.083	.154	**.432**
DL(1.5)	.01	UFR	.015	.071	**.173**	.022	.239	**.540**
	.05		.065	.253	**.492**	.095	.501	**.843**

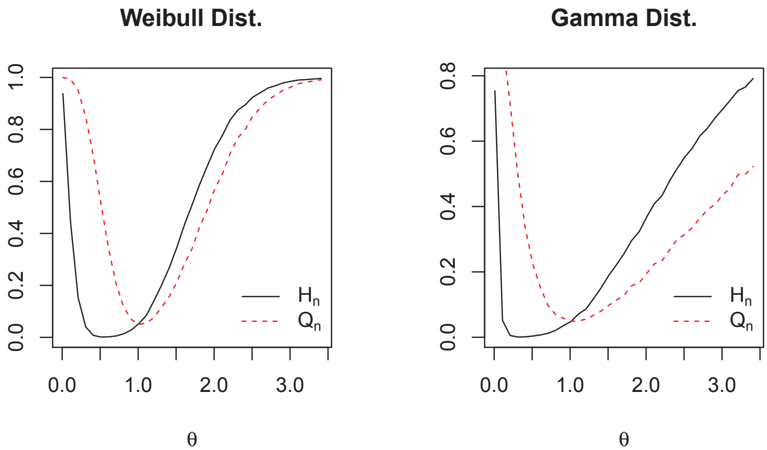

Figure 1. Comparison of the tests based on Q_n and H_n for Weibull and gamma distribution for $n = 10$ and $\alpha = 0.05$.

Acknowledgement

The authors are highly grateful to referees for their valuable comments and suggestions for improving the paper.

References

Aarset, M. (2004). How to Identify Bathtub Hazard Rate. *IEEE Transactions on Reliability*, **36**, 106-108.

Abbasnejad, M. (2011). Some Goodness-of-fit Tests Based on Renyi Information.*Applied Mathematical Sciences*, **5**, 1921-1934.

Abbasnejad, M., Arghami, N.R. and Tavakoli, M. (2012). A Goodness-of-fit Tests for Exponentiality bBsed on Lin-Wong Information. *Journal of the Iranian Statistical Society*, **11**, 191-202.

Alizadeh Noughabi, H. and Arghami, N.R. (2010). A New Estimator of Entropy and Its Application in Testing Normality. *Journal of Statistical Computation and Simulation*, **80**, 1151-1162.

Alizadeh Noughabi, H. and Arghami, N.R. (2011). Monte Carlo Comparison of Five Exponentiality Tests Using Different Entropy Estimates. *Journal of Statistical Computation and Simulation*, **81**, 1579-1592.

Alizadeh Noughabi, H. and Arghami, N.R. (2011a). Testing Exponentiality Using Transformed Data. *Journal of Statistical Computation and Simulation*, **81**, 511-516.

Alizadeh Noughabi, H. and Arghami, N.R. (2011b). Testing Exponentiality Based on Characterizations of the Exponential Distribution. *Journal of Statistical Computation and Simulation*, **81**, 1641-1651.

Chen, Z. (2000). A New Two-Parameter Lifetime Distribution with Bathtub Shape or Increasing Failure Rate Function. *Statistics and Probability Letters*, **49**, 155-161.

Choi, B., Kim, K. and Heun Song, S. (2004). Goodness-of-fit Test for Exponentiality Based on Kullback–Leibler Information. *Communications in Statistics - Simulation and Computation*, **33**, 525-536.

Correa, J. (1995). A New Estimator of Entropy. *Communications in Statistics -Theory and Methods*, **24**, 2439-2449.

Cuadras, C.M. and Fortiana, J. (1993). Continuous Metric Scaling and Prediction. In: Cuadras, C.M., Rao, C.R. (Eds.), Multivariate Analysis, Future Directions, Vol. 2. Elsevier Science Publishers B.V., North-Holland, Amsterdam, 47-66.

Dhillon, B. (1981). Life Distributions. *IEEE Transactions on Reliability*, **30**, 457-459.

Ebrahimi, N., Habibullah, M. and Soofi, S.E. (1992). Testing Exponentiality Based on Kullback-Leibler Information. *Journal of the Royal Statistical Society B*, **54**, 739-748.

Ebrahimi, N., Pflughoeft, K. and Soofi, S.E. (1994). Two Measures of Sample Entropy. *Statistics and Probability Letters*, **20**, 225-234.

Finkelstein, J. and Schafer, R.E. (1971). Imported Goodness-of-fit Tests. *Biometrika*, **58**, 641-645.

Fortiana, J. and Grane, A. (2002) A Scale-free Goodness-of-fit Statistic for the Exponential Distribution Based on Maximum Correlations. *J Stat Plann Infer*, **108**, 85-97.

Grane, A. and Fortiana, J. (2009). A Location and Scale Goodness-of-fit Statistic for the Exponential Distribution Based on Maximum Correlations. *Statistics* **43**, 1-12

Gurevich, G. and Davidson, A. (2008). Standardized Forms of Kullback-Leibler Information Based Statistics for Normality and Exponentiality. *Computer Modelling and New Technologies*, **12**, 14-25.

Henze, N. and Meintanis, S.G. (2005). Recent and Classical Tests for Exponentiality: A Partial Review with Comparisons. *Metrika*, **61**, 29-45.

Van Es, B. (1992). Estimating Functionals Related to a Density by Class of Statistics Based on Spacing. *Scand. J. Statist.*, **19**, 61-72.

Van-Soest, J. (1969). Some Goodness-of-fit Tests for the Exponential Distribution. *Statist. Neerland.*, **23**, 41-51.

Vasicek, O. (1976). A Test for Normality Based on Sample Entropy. *Journal of the Royal Statistical Society. Series B (Methodological)*, **38**, 54-59.

Narges H. Montazeri
Department of Statistics,
Yazd University,
Yazd, Iran.
email: *nmontazeri@stu.yazd.ac.ir*

Hamzeh Torabi
Department of Statistics,
Yazd University,
Yazd, Iran.
email: *htorabi@yazd.ac.ir*