CrossMark

REGULAR ARTICLE

On empirical cumulative residual entropy and a goodness-of-fit test for exponentiality

V. Zardasht · S. Parsi · M. Mousazadeh

Abstract The cumulative residual entropy (CRE) is a new measure of information and an alternative to the Shannon differential entropy in which the density function is replaced by the survival function. This new measure overcomes deficiencies of the differential entropy while extending the Shannon entropy from the discrete random variable cases to the continuous counterpart. Some properties of the cumulative residual entropy, its estimation and applications has been studied by many researchers. The objective of this paper is twofold. In the first part, we give a central limit theorem result for the empirical cumulative residual entropy based on a right censored random sample from an unknown distribution. In the second part, we use the CRE of the comparison distribution function to propose a goodness-of-fit test for the exponential distribution. The performance of the test statistic is evaluated using a simulation study. Finally, some numerical examples illustrating the theory are also given.

Keywords Censored data · Comparison distribution function · Cumulative entropy

1 Introduction

From a probabilistic point of view, to study a stochastic phenomenon, we try to measure how much chance a spatial outcome of the phenomena has in order for it to occur. A different viewpoint is adopted from an information theoretic point of view, which tries to answer how much we are able to predict the outcome of the phenomenon. In other words, we try to measure the amount of uncertainty or entropy contained in the

V. Zardasht (✉) · S. Parsi · M. Mousazadeh
Department of Statistics, Faculty of Mathematical Sciences, University of Mohaghegh Ardabili,
Ardabil, Iran
e-mail: zardasht@uma.ac.ir

🌀 Springer

outcome. Shannon (1948) was able to formulate the measurement of this uncertainty contained in a single event. The uncertainty contained in a discrete random variable is then considered as the weighted average of the uncertainty of each single event. Formally, for a discrete random variable X with probability mass function $p(x) = P(X = x)$, the Shannon entropy is defined as

$$H(X) = -\sum_x p(x) \log p(x). \tag{1}$$

Note that $H(X) = -E[\log p(X)]$ and an immediate extension leads us to its continuous analog called the differential entropy. That is, for a non-negative continuous random variable with density function $f(x)$ the differential entropy, which we denote by $H^c(X)$, is defined as

$$H^c(X) = -E[\log f(X)] = -\int_0^\infty f(x) \log f(x) dx. \tag{2}$$

Nevertheless, it is well-known (cf. Di Crescenzo and Longobardi (2009) and references therein) that this extension does not preserve some basic properties of an information measure; for instance, the differential entropy can take on negative values. Recently, among various attempts to define possible alternative information theoretic measures, Rao et al. (2004) proposed the cumulative residual entropy (CRE) and studied its properties. This measure replaces density function by the survival function. For a non-negative random variable X with distribution function F and survival function $\bar{F} = 1 - F$, the CRE is defined as follows:

$$\mathcal{E}(X) = -\int_0^\infty \bar{F}(x) \log \bar{F}(x) dx. \tag{3}$$

Properties of the CRE can be found in Rao (2005), Di Crescenzo and Longobardi (2009), and Navarro et al. (2010). Di Crescenzo and Longobardi (2009) also introduced and studied the Cumulative Entropy, denoted by $\mathcal{CE}(X)$, as an analog to CRE by using distribution function in (3) instead of survival function. That is,

$$\mathcal{CE}(X) = -\int_0^\infty F(x) \log F(x) dx.$$

Asadi and Zohrevand (2007) considered the corresponding dynamic properties of the CRE corresponding to the residual lifetime variable. Applications of CRE to image alignment and measurements of similarity between images can also be found in Wang and Vemuri (2007) and references therein.

Due to the extensive applications of various information criteria in studying biological and engineering systems, it is incumbent on practitioners to estimate

CRE when no prior information are available on the underlying distribution of X. Rao et al. (2004) consider empirical CRE as a plug-in estimator of CRE through replacing survival function by the empirical survival function and show that it is a strongly consistent estimator for CRE. Similarly, Di Crescenzo and Longobardi (2009) use the empirical cumulative entropy to estimate the cumulative entropy. They are also able to prove the strong consistency of the empirical cumulative entropy and provided a central limit theorem based on a random sample from an exponential distribution. It is also worthy to mention that similar results have been obtained for other information measures. For instance, Abraham and Sankaran (2006) introduced and studied Renyi's information measure for residual lifetime distributions. Maya et al. (2013) proposed several nonparametric estimators for the Renyi's information measure for the residual lifetime distribution based on complete and censored data and established their asymptotic properties under suitable regularity conditions.

Let $X_1, \ldots, X_n$ be independent positive random variables with continuous distribution function $F(t)$, survival Function $\bar{F}(t) = 1 - F(t)$, and cumulative hazard function $\Lambda(t) = -\log \bar{F}(t)$. Assume that X_is are censored on the right by independent and identically distributed positive random variables T_i (with survival function $\bar{C}(x)$) which are also independent of X_i. Define $Z_i = \min\{X_i, T_i\}$ and $\delta_i = 1$ or 0 according as to whether $X_i \leq T_i$ or $X_i > T_i$ respectively. Then the available data are $\{(Z_1, \delta_1), \ldots, (Z_n, \delta_n)\}$. A well-known estimate of $\bar{F}$ is the Kaplan-Meier estimator, $\hat{\bar{F}}$ (Kaplan and Meier,1958) which is given by

$$\hat{\bar{F}}(t) = \prod_{i:z_i \leq t} \left(1 - \frac{\delta_i}{n - i + 1}\right). \tag{4}$$

In this paper, we replace $\bar{F}$ and Λ with their corresponding Kaplan-Meier and Nelson-Aalen estimators, respectively. Observe that $\mathcal{E}(X)$ can also be written as

$$\mathcal{E}(X) = \int_0^\tau \bar{F}(x)\Lambda(x)dx, \tag{5}$$

where $\tau = \sup\{x : \bar{F}(x) > 0\}$, and due to this, we propose the following estimator of the CRE:

$$\mathcal{E}(\hat{\bar{F}}) = \int_0^\tau \hat{\bar{F}}(x)\hat{\Lambda}(x)dx. \tag{6}$$

In this paper, we will prove that this estimator is a consistent estimator and its asymptotic distribution is normal.

Testing for exponentiality has involved a great deal of current statistical research recently, and is of some importance in statistical inference. The tests are usually constructed by using the characterization results from reliability theory and also by using different information measures such as similarity or discrimination measures for comparing between distribution functions(cf. Baringhaus and Henze 2000; Baratpour and

Habibi Rad 2012, and references therein). Let X and Y be non-negative random variables with distribution functions F and G, respectively. To compare between X and Y, the comparison distribution function (Parzen 1998) is defined as $D(u) = F(G^{-1}(u))$, for $0 \leq u \leq 1$ (note that if $G = F$, then $D(u)$ will be the cumulative distribution function of the uniform distribution). Our test statistic is motivated by considering the CRE for the comparison distribution function, that is

$$\mathcal{C}(X, Y) = -\int_0^\infty \bar{F}(x) \log \bar{F}(x) dG(x). \tag{7}$$

If Y is a non-negative random variable with distribution function G and X is distributed as exponential distribution with mean λ, then

$$\mathcal{C}(exp, Y) = \int_0^\infty \frac{x}{\lambda} e^{-\frac{x}{\lambda}} dG(x), \tag{8}$$

will compare distribution function G with the exponential distribution. If Y is distributed as exponential distribution, then $\mathcal{C}(exp, Y) = \frac{1}{4}$, which is indeed, the value of CRE for a standard uniform random variable. Viewing the difference $\mathcal{C}(exp, Y) - \frac{1}{4}$ as a measure of the deviation of the distribution of Y from the exponential distribution, we give another goodness-of-fit test for the exponential distribution.

The rest of the paper is organized as follows. In Sect. 2 we give the large sample properties of the empirical CRE. In Sect. 3 we apply the comparison CRE to construct a goodness-of-fit test for the exponential distribution. Section 4 is devoted to the simulation results and a couple of numerical examples and finally, some concluding remarks are given in Sect. 5.

2 Asymptotic properties of $\mathcal{E}(\hat{\bar{F}})$

In this section, we investigate the consistency and asymptotic normality of $\mathcal{E}(\hat{\bar{F}})$. We first recall some notations from standard counting process methods. Let $N(t) = \sum_{i=1}^n I(Z_i \leq t, \delta_i = 1)$ be the number of failures or deaths up to time t i.e, the number of uncensored samples, and $Y(t) = \sum_{i=1}^n I(Z_i \geq t)$ be the number of at risk process. The Nelson-Aalen estimator of the cumulative hazard function is given by (cf. Kalbfleisch and Prentice 2002, p. 168) $\hat{\Lambda}(t) = \int_0^t dN(u)/Y(u)$, where the reciprocal of $Y(u)$ is defined to be 0 whenever $Y(u)$ is 0. It is also well-known that the process $M(t) = N(t) - \int_0^t Y(u)d\Lambda(u)$ is a square integrable martingale with respect to the natural filtration.

Theorem 2.1 *Let* $y(t) = \bar{F}(t)\bar{C}(t)$. *Then, as* $n \to \infty$,

(i) $\mathcal{E}(\hat{\bar{F}}) \xrightarrow{p} \mathcal{E}(X)$,

(ii) $\sqrt{n}(\mathcal{E}(\hat{\bar{F}}) - \mathcal{E}(X))$ *converges in distribution to a Gaussian random variable Z with mean zero and variance*

$$\sigma^2 = \int_0^\tau \int_0^\tau \bar{F}(t)\bar{F}(u)v(t \wedge u)dtdu, \tag{9}$$

where,

$$v(t) = \int_0^t \frac{d\Lambda(u)}{y(u)},$$

and $\xrightarrow{p}$ *represents convergence in probability.*

Proof First, one can easily show by using the Glivenko-Cantelli Theorem that

$$\sup_{0 \le t \le \tau} |Y(t)/n - y(t)| \xrightarrow{p} 0.$$

This and the Rebolledo's Theorem (see Kalbfleisch and Prentice 2002, pp. 166–168) imply that $\sqrt{n}(\hat{\Lambda}(t) - \Lambda(t))$ converges to a Gaussian random variable with mean zero and variance $v(t)$. The result now follows from Theorem 3.1 in Sengupta et al. (1998), as its one dimensional case, by replacing $K(t)$ and $X(t)$ by $\hat{\bar{F}}(t)$, the Kaplan-Meier estimator of $\bar{F}$, and $\sqrt{n}(\hat{\Lambda}(t) - \Lambda(t))$, respectively.

By the standard counting process method, an estimator of σ^2 can be given by

$$\hat{\sigma}^2 = \int_0^\tau \int_0^\tau \hat{\bar{F}}(t)\hat{\bar{F}}(u)V(t \wedge u)dtdu, \tag{10}$$

where,

$$V(t) = \int_0^t \frac{dN(u)}{Y^2(u)},$$

and $t \wedge u$ stands for $\min\{t, u\}$. $\square$

Remark 2.2 In the censored sample case, an analogue estimator of the cumulative entropy can also be given by

$$\mathcal{CE}(\hat{\bar{F}}) = -\int_0^\tau (1 - \hat{\bar{F}}(x)) \log\left(1 - \hat{\bar{F}}(x)\right) dx. \tag{11}$$

By applying the same method, one can easily conclude that, as $n \to \infty$, $\mathcal{CE}(\hat{\bar{F}})$ converges in probability to $\mathcal{CE}(X)$. Furthermore, $\sqrt{n}(\mathcal{CE}(\hat{\bar{F}}) - \mathcal{CE}(X))$ converges in distribution to a zero mean Gaussian random variable with variance estimated by

$$\hat{\sigma}^2 = \int_0^\tau \int_0^\tau \log\left(1 - \hat{\bar{F}}(t)\right) \log(1 - \hat{\bar{F}}(u)) V^*(t \wedge u) dt du, \tag{12}$$

where,

$$V^*(t) = \hat{\bar{F}}^2(t) \int_0^t \frac{dN(u)}{Y(u)[Y(u) - \Delta N(u)]},$$

and $\Delta N(t) = N(t) - N(t^-)$. This is an extension for the result by Di Crescenzo and Longobardi (2009) in which they provide a central limit theorem for the empirical cumulative entropy based on random samples from the exponential distribution.

3 A goodness-of-fit test for the exponential distribution

Let $X_1, X_2, \ldots, X_n$ be a random sample from the population of a non-negative random variable X with continuous distribution function F. In this section, we apply the measure (8) to construct a test statistic for testing the hypothesis $H_0 : F(x) = 1 - e^{-x/\lambda}$ versus the alternative $H_a : F(x) \neq 1 - e^{-x/\lambda}$. Under the null hypothesis $\mathcal{C}(exp, X) = \frac{1}{4}$ (which is indeed, the value of CRE for a standard uniform random variable), then large or small value of the difference $\mathcal{C}(exp, X) - \frac{1}{4}$ will lead us to reject the null hypothesis in favor of the alternative H_a. Using the standard U-statistic theory (cf. Lee 1990), we propose the following statistic C_n, an estimator of $\mathcal{C}(exp, X)$, as our test statistic:

$$C_n = \frac{1}{n} \sum_{i=1}^n \frac{X_i}{\bar{X}} e^{-\frac{X_i}{\bar{X}}}, \tag{13}$$

where $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. The following theorem gives the asymptotic distribution of the test statistic.

Theorem 3.1 *Under the null hypothesis H_0, as $n \to \infty$*

$$\sqrt{n}\left(C_n - \frac{1}{4}\right) \xrightarrow{d} N\left(0, \frac{5}{382}\right),$$

where, $\xrightarrow{d}$ denotes convergence in distribution and $N(0, \frac{5}{382})$ stands for the normal random variable with mean zero and variance $\frac{5}{382}$.

Proof First, the central limit theorem gives that $\sqrt{n}(\bar{X} - \lambda) \xrightarrow{d} N(0, \lambda^2)$ which implies that $\sqrt{n}(\bar{X} - \lambda) = O_p(1)$. On the other hand, from the standard U-statistic theory (cf. Lehmann 1999, p. 369), under the null hypothesis we have

$$\sqrt{n}\left(U_n(\lambda) - \frac{1}{4}\right) \xrightarrow{d} N\left(0, \frac{5}{382}\right),$$

where $U_n(\lambda) = \frac{1}{n} \sum_{i=1}^{n} \frac{X_i}{\lambda} e^{-\frac{X_i}{\lambda}}$. Now the result immediately follows from Theorem 2.13 in Randles (1982). □

We reject H_0 in favor of H_a at the significant level α if $\sqrt{\frac{382n}{5}}|C_n - \frac{1}{4}| > Z_{1-\frac{\alpha}{2}}$, where $Z_{1-\frac{\alpha}{2}}$ is $100(1 - \frac{\alpha}{2})$- percentile of the standard normal distribution. In the next section, we use the Monte Carlo simulation to compare the power of our test statistic with some other statistics for fitting the exponential distribution to a random sample data.

4 Simulation study

Recently, Baratpour and Habibi Rad (2012) provide a goodness-of-fit test statistic based on a discrimination measure arising from a version of the Kulback-Leibler information measure to test the hypothesis H_0 versus the alternative H_a. Their test statistic is given by

$$T_n = \frac{\sum_{i=1}^{n-1} \frac{n-i}{n} (\ln \frac{n-i}{n})(X_{(i+1)} - X_{(i)}) + \frac{\sum_{i=1}^{n} X_i^2}{2 \sum_{i=1}^{n} X_i}}{\frac{\sum_{i=1}^{n} X_i^2}{2 \sum_{i=1}^{n} X_i}},$$

where $X_{(i)}$ is the ith ordered statistic related to the sample and H_0 is rejected at significant level α if $T_n \geq T_{n,1-\alpha}$, where $T_{n,1-\alpha}$ is $100(1 - \alpha)$-percentile of T_n under H_0. They also provide a Monte Carlo simulation study to compare between the performance of T_n, the statistic introduced by Van-Soest (1969)

$$W^2 = \sum_{i=1}^{n} \left[F_0(X_{(i)}, \hat{\lambda}) - \frac{2i-1}{2n}\right]^2 + \frac{1}{12n},$$

the statistic introduced by Finkelstein and Schafer (1971)

$$S^* = \sum_{i=1}^{n} \max\left\{\left|F_0(X_{(i)}, \hat{\lambda}) - \frac{i}{n}\right|, \left|F_0(X_{(i)}, \hat{\lambda}) - \frac{i-1}{n}\right|\right\},$$

where $F_0(x, \lambda) = 1 - e^{-\frac{x}{\hat{\lambda}}}$, $\hat{\lambda} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ and the one introduced by Choi et al. (2004)

$$KLC_{mn} = \frac{\exp(C_{mn})}{\exp(\ln \bar{X} + 1)},$$

where $C_{mn} = -\frac{1}{n} \sum_{i=1}^{n} \log \frac{\sum_{j=i-m}^{i+m} (X_{(j)} - X_{(i)})(j-i)}{n \sum_{j=i-m}^{i+m} (X_{(j)} - \bar{X}_i)^2}$ and $\bar{X}_i = \frac{1}{2m+1} \sum_{j=i-m}^{i+m} X_{(j)}$, which are proposed for testing H_0 against H_a. In KLC_{mn} statistic, the window size m is a positive integer smaller than $\frac{n}{2}$, $X_{(j)} = X_{(1)}$, if $j < 1$, and $X_{(j)} = X_{(n)}$, if $j > n$. H_0 is rejected of large values of W^2, S^* and of small values of KLC_{mn}. We have undertaken a simulation exercise to investigate the performance of our test statistic comparing it with the above statistics T_n, W^2, S^*, and KLC_{mn}. In our simulation, we considered the following distribution functions and the empirical powers of the test statistics were compared for each of the distributions.

(i) a Weibull distribution with density function

$$f(x; , \lambda, \beta) = \frac{\beta}{\lambda^\beta} x^{\beta-1} \exp\left\{-\left(\frac{x}{\lambda}\right)^\beta\right\}, \quad \beta > 0, \ \lambda > 0, \ x > 0,$$

(ii) a gamma distribution with density function

$$f(x; , \lambda, \beta) = \frac{x^{\beta-1} \exp\left\{-\frac{x}{\lambda}\right\}}{\Gamma(\beta)\lambda^\beta}, \quad \beta > 0, \ \lambda > 0, \ x > 0,$$

(iii) a lognormal distribution with density function

$$f(x; \mu, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\ln x - \mu)^2\right\},$$
$$-\infty < \mu < \infty, \ \sigma > 0, \ x > 0,$$

(iv) an inverse Gaussian distribution with density function

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}, \quad \mu > 0, \ \lambda > 0, \ x > 0.$$

As in Baratpour and Habibi Rad (2012), for each case we set the parameters such that $\frac{E(X_1^2)}{2E(X_1)} = 1$, That is, $\lambda = \frac{2\Gamma(1+\frac{1}{\beta})}{\Gamma(1+\frac{2}{\beta})}$ for the Weibull distribution, $\lambda = \frac{2}{1+\beta}$ for the gamma distribution, $\sigma^2 = \frac{2}{3}(\ln 2 - \mu)$ for the lognormal distribution and $\lambda = \frac{\mu^2}{2-\mu}$ for the inverse Gaussian distribution. The empirical power was computed for each statistic under a total of 100, 000 generated samples of sizes $n = 5, 10, 15, 20, 25$. The power was taken as the fractional number of times, out of 100, 000, the corresponding statistic exceeded the relevant threshold. Tables 1, 2, 3 and 4 summarize the results

Table 1 Power comparison between the tests C_n, T_n, W^2, S^* and KLC_{mn}, at the significance level $\alpha = 0.05$, when the alternative distribution is Weibull

n	β	C_n	T_n	W^2	S^*	KLC_{mn}
5	1	0.04224	0.05019	0.04969	0.04883	0.04988
	2	0.23628	0.34555	0.29159	0.29868	0.36816
	3	0.57310	0.70709	0.64287	0.65862	0.73240
	4	0.81875	0.91062	0.86633	0.88015	0.91904
10	1	0.04299	0.05148	0.05025	0.04977	0.05130
	2	0.61451	0.65573	0.60997	0.62598	0.75572
	3	0.96963	0.97992	0.96771	0.97515	0.99297
	4	0.99904	0.99966	0.99911	0.99938	0.99996
15	1	0.04512	0.05024	0.05025	0.05024	0.05054
	2	0.84462	0.82598	0.82438	0.84206	0.92164
	3	0.99878	0.99913	0.99845	0.99908	0.99986
	4	1.00000	1.00000	1.00000	1.00000	1.00000
20	1	0.04476	0.05063	0.05303	0.05200	0.05245
	2	0.94289	0.91759	0.93264	0.94039	0.97907
	3	0.99996	0.99994	0.99995	0.99995	1.00000
	4	1.00000	1.00000	1.00000	1.00000	1.00000
25	1	0.04614	0.05184	0.04845	0.05024	0.05098
	2	0.98272	0.96440	0.97593	0.98078	0.99408
	3	1.00000	1.00000	1.00000	1.00000	1.00000
	4	1.00000	1.00000	1.00000	1.00000	1.00000

of the simulation for each example. One can see from the tables that the power of all tests against any alternative show an increasing pattern with respect to sample size. This reveals the consistency of the tests. In general, there is no big difference between the power of the test statistics C_n and other tests, but it has the added advantages of having simple form and a known asymptotic distribution.

4.1 Data Analysis

In this section, we give a couple of numerical examples based on real life data set to illustrate the use of the test statistic C_n for validating the goodness of an exponential distribution fitting to a real data set.

Example 4.1 Proschan (1963) gave date on time, in hours of operation, between successive failures of air-conditioning equipment in 13 aircraft to study their aging properties. The data for plane number 3 are as follows:

90, 10, 60, 186, 61, 49, 14, 24, 56, 20, 79, 84, 44, 59, 29, 118, 25, 156, 310, 76, 26, 44, 23, 62, 130, 208, 70, 101, 208.

Table 2 Power comparison between the tests C_n, T_n, W^2, S^* and KLC_{mn}, at the significance level $\alpha = 0.05$, when the alternative distribution is Gamma

n	β	C_n	T_n	W^2	S^*	KLC_{mn}
5	1	0.04191	0.05012	0.05006	0.04940	0.04933
	5	0.40790	0.51620	0.48736	0.48864	0.56962
	6	0.50300	0.60581	0.58484	0.58413	0.66622
	7	0.58781	0.68259	0.66763	0.66545	0.74434
10	1	0.04359	0.04901	0.04989	0.05058	0.04968
	5	0.91530	0.83344	0.89267	0.89011	0.94032
	6	0.96863	0.90964	0.95473	0.95258	0.97765
	7	0.98961	0.95350	0.98275	0.98025	0.99189
15	1	0.04323	0.05104	0.05156	0.05002	0.05041
	5	0.99399	0.94619	0.98833	0.98622	0.99293
	6	0.99935	0.98234	0.99834	0.99774	0.99857
	7	0.99986	0.99393	0.99957	0.99949	0.99968
20	1	0.04241	0.05125	0.04915	0.04845	0.05056
	5	0.99968	0.98361	0.99891	0.99870	0.99865
	6	0.99999	0.99655	0.99998	0.99993	0.99982
	7	1.00000	0.99946	1.00000	1.00000	0.99998
25	1	0.04616	0.05006	0.05045	0.05118	0.04967
	5	0.99998	0.99546	0.99995	0.99992	0.99972
	6	1.00000	0.99944	1.00000	1.00000	0.99998
	7	1.00000	0.99995	1.00000	1.00000	1.00000

Applying the test statistic C_n gives $C_{29} = 0.269$, and the standard normal distribution approximation to $\sqrt{\frac{(29)(382)}{5}}(C_{29} - 0.25)$ gives a P-value of 0.379. Thus, the test does not reject the null hypothesis that the failure times follow an exponential distribution at significance level $\alpha = 0.05$. Using three other test statistics, Lawless (1982) obtained the same result for the above failure data.

Example 4.2 The following data are from Lawless (1982) and it consists of failure times for 36 appliances subjected to an automatic life test.

111, 351, 491, 1701, 3291, 3811, 7081, 9581, 10621, 11671, 15941, 19251, 19901, 22231, 23271, 24001, 24511, 24711, 25511, 25651, 25681, 26941, 27021, 27611, 28311, 30341, 30591, 31121, 32141, 34781, 35041, 43291, 63671, 69761, 78461, 13403.

For these data, we obtain $C_{36} = 0.28$ and the normal approximation gives a P-value of 0.115. Thus, the test accepts the null hypothesis at significance level $\alpha = 0.05$. That is, the test does not indicate any evidence against the exponential model for the failure times.

Table 3 Power comparison between the tests C_n, T_n, W^2, S^* and KLC_{mn}, at the significance level $\alpha = 0.05$, when the alternative distribution is Lognormal

n	μ	C_n	T_n	W^2	S^*	KLC_{mn}
5	0.4	0.43924	0.52026	0.52316	0.51504	0.60221
	0.5	0.65915	0.71358	0.73654	0.72025	0.80221
	0.6	0.94782	0.94778	0.96987	0.95878	0.98282
10	0.4	0.93266	0.79615	0.92483	0.90023	0.93280
	0.5	0.99586	0.94968	0.99498	0.98956	0.99318
	0.6	1.00000	0.99967	1.00000	0.99999	1.00000
15	0.4	0.99591	0.90456	0.99393	0.98985	0.98512
	0.5	1.00000	0.99045	0.99999	0.99989	0.99935
	0.6	1.00000	0.99999	1.00000	1.00000	1.00000
20	0.4	0.99989	0.95570	0.99966	0.99915	0.99317
	0.5	1.00000	0.99831	1.00000	1.00000	0.99987
	0.6	1.00000	1.00000	1.00000	1.00000	1.00000
25	0.4	1	0.97908	0.99999	0.99996	0.99696
	0.5	1	0.99972	1.00000	1.00000	0.99996
	0.6	1	1.00000	1.00000	1.00000	1.00000

Table 4 Power comparison between the tests C_n, T_n, W^2, S^* and KLC_{mn}, at the significance level $\alpha = 0.05$, when the alternative distribution is Inverse Guassian

n	μ	C_n	T_n	W^2	S^*	KLC_{mn}
5	0.1	0.37544	0.01270	0.36479	0.36068	0.03382
	0.2	0.27751	0.01565	0.27316	0.26882	0.02269
	0.3	0.20014	0.01995	0.20386	0.19872	0.02279
10	0.1	0.64859	0.38345	0.69273	0.68652	0.00660
	0.2	0.50078	0.25832	0.56465	0.55579	0.00445
	0.3	0.35766	0.17202	0.43788	0.42856	0.00702
15	0.1	0.80302	0.66484	0.85406	0.85299	0.01217
	0.2	0.66068	0.51416	0.74788	0.74583	0.00189
	0.3	0.48970	0.37546	0.61101	0.60813	0.00255
20	0.1	0.89120	0.82312	0.93295	0.93162	0.00184
	0.2	0.77133	0.68531	0.85553	0.85153	0.00039
	0.3	0.60069	0.54071	0.73889	0.73376	0.00111
25	0.1	0.94230	0.90814	0.96927	0.96874	0.00299
	0.2	0.85083	0.80146	0.91982	0.91854	0.00014
	0.3	0.69402	0.66931	0.82468	0.82200	0.00042

5 Conclusion

In this paper, we have considered the asymptotic behaviour of the empirical cumulative residual entropy. We were able to show that the empirical CRE converges in

distribution to a normal random variable. It was also shown that the same result holds for the empirical cumulative entropy which extends the result by Di Crescenzo and Longobardi (2009). We used the CRE entropy of the comparison distribution function to propose a new goodness-of-fit test for an exponential distribution. An extensive simulation exercise was undertaken to compare between the performance of this test statistic and four other test statistics and the results revealed the consistency and high power of the proposed test statistic. Finally, using a couple of numerical examples, the use of the test statistic for testing goodness-of-fit for exponential distribution was illustrated.

References

Abraham B, Sankaran PG (2006) Renyi's entropy for residual lifetime distribution. Stat Pap 47:17–29

Asadi M, Zohrevand Y (2007) On the dynamic cumulative residual entropy. J Stat Plan Inference 137:1931–1941

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

Baringhaus L, Henze N (2000) Tests of fot for exponentiality based on a characterization via the mean residual life function. Stat Pap 41:225–236

Choi B, Kim K, Song SH (2004) Goodness-of-fit test for exponentiality based on KullbackLeibier information. Simul Comput 33(2):525–536

Di Crescenzo A, Longobardi M (2009) On cumulative entropies. J Stat Plan Inference 139:4072–4087

Finkelstein J, Schafer RE (1971) Improved goodness of fit tests. Biometrika 58:641–645

Kalbfleisch JD, Prentice RL (2002) The statistical analysis of failure time data, 2nd edn. Wiley, New York

Lawless JF (1982) Statistical models and methods for life-time data. Wiley, New York

Lee AJ (1990) U-statistics: theory and practice. Marcel Dekker Inc., Florida

Lehmann EL (1999) Elements of large-sample theory. Springer, New York

Maya R, Abdul-Sathar EI, Rajesh G, Muraleedharan Nair KR (2013) Estimation of the Renyi's residual entropy of order α with dependent data. Stat Pap. doi:10.1007/s00362-013-0506-1

Navarro J, del Aguila Y, Asadi M (2010) Some new results on the cumulative residual entropy. J Stat Plan Inference 140:310–322

Parzen E (1998) Statistical methods mining, two sample data analysis, comparison distributions and quantile limit theorems. In: Szyszkowicz B (Ed) Asymptotic methods in probability and statistics, pp 611–617

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

Randles RH (1982) On the asymptotic normality of statistics with estimated parameters. Annal stat 10(2):462–474

Rao M (2005) More on a new concept of entropy and information. J Theorl Probab 18:967–981

Rao M, Chen Y, Vemuri BC, Wang F (2004) Cumulative residual entropy: a new measure of information. IEEE Trans Inf Theory 50:1220–1228

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:279–423

Van-Soest F (1969) Some goodness of fit tests for the exponential distribution. Stat Neerl 23:41–51

Wang F, Vemuri BC (2007) Non-rigid multi-modal image registration using cross-cumulative residual entropy. Intern J Comput Vis 74:201–215