Communications for Statistical Applications and Methods
2014, Vol. 21, No. 2, 125–134

DOI: http://dx.doi.org/10.5351/CSAM.2014.21.2.125
ISSN 2287-7843

Kullback-Leibler Information of the Equilibrium Distribution Function and its Application to Goodness of Fit Test

Sangun Park[1,a], Dongseok Choi[b], Sangah Jung[a]

[a]Department of Applied Statistics, Yonsei University, Korea;
[b]Oregon Health and Science University, USA

Abstract

Kullback-Leibler (KL) information is a measure of discrepancy between two probability density functions. However, several nonparametric density function estimators have been considered in estimating KL information because KL information is not well-defined on the empirical distribution function. In this paper, we consider the KL information of the equilibrium distribution function, which is well defined on the empirical distribution function (EDF), and propose an EDF-based goodness of fit test statistic. We evaluate the performance of the proposed test statistic for an exponential distribution with Monte Carlo simulation. We also extend the discussion to the censored case.

Keywords: Cumulative residual KL information, exponential distribution, Fisher information, Goodness of fit test.

1. Introduction

Suppose that a random variable X has a distribution function, $F(x)$ with a continuous density function $f(x)$. Shannon entropy is defined as

$$H(f) = -\int_{-\infty}^{\infty} f(x)\log f(x)dx.$$

Shannon entropy is a measure of uncertainty and the distribution function maximizing the entropy (under some constraints) is called maximum entropy distribution (Jayens, 1957). The entropy of order statistics has been studied by Wong and Chen (1990), Park (1995), Abo-Eleneen (2011) and Mosayeb and Borzadaran (2013).

The Kullback-Leibler (KL) information is defined for $f(x)$ and $g(x)$ being the reference distribution as

$$K(g : f) = \int_{-\infty}^{\infty} g(x)\log\frac{g(x)}{f(x)}dx.$$

KL information is nonnegative and the equality to zero holds iff $f(x) = g(x)$. The sample estimates of $H(f)$ and $K(g : f)$ can be simply obtained as $H(f_n)$ and $K(g_n : f_n)$, respectively, but are not

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korea government(MEST) (No. 2011-0029104).

[1] Corresponding author: Professor, Department of Applied Statistics, Yonsei University, Shinchon Dong 134, Seoul 120-749, Korea. E-mail: sangun@yonsei.ac.kr

Published online 31 March 2014 / journal homepage: http://csam.or.kr

attainable for $f_n = dF_n$ where F_n is the empirical distribution function. Hence, several nonparametric density function estimators (Theil, 1980; Dudewicz and van der Meulen, 1981; Bowman, 1992; Park and Park, 2003) have been considered in estimating KL information, but we have to determine the bandwidth or the gap of order statistics.

In this paper, we consider the equilibrium distribution (or integrated tail distribution) whose probability density function is defined to be $f^*(x) = \bar{F}(x)/\int_0^\infty \bar{F}(x)dx$ (Andrews and Andrews, 1962; Nakamura, 2009) and study its entropy and KL information where $\bar{F}(x) = 1 - F(x)$. Since $f_n^*(x) = F_n(x)/\bar{x}$, $H(f^*)$ is well-defined on the empirical distribution function as

$$H(f_n^*) = -\frac{1}{\bar{x}} \sum_{i=0}^{n} \frac{n-i}{n} \log \frac{n-i}{n} (x_{i+1:n} - x_{i:n}) + \log \bar{x},$$

where $x_{i:n}$ is i^{th} ordered value from a sample of size n and $x_{0:n} = 0$.

KL information of $f^*(x)$ and $g^*(x)$ being the reference distribution can be written as

$$K(g^* : f^*) = \int_0^\infty \frac{\bar{G}(x)}{\int_0^\infty \bar{G}(x)dx} \log \frac{\bar{G}(x)}{\bar{F}(x)} dx + \log \frac{\int_0^\infty \bar{F}(x)dx}{\int_0^\infty \bar{G}(x)dx}$$

and is also well-defined on the empirical distribution function. $K(g^* : f^*)$ is nonnegative, and the equality to zero holds iff $F(x) = G(x)$ under the equal first moment condition. We consider this equal moment condition in parameter estimation and can provide the estimate of $K(g^* : f^*)$ as an EDF-based goodness fit test statistic. Monte Carlo simulation study has been done to evaluate the performance of the proposed test statistic. We also extend the result to the censored case by taking the equilibrium distribution function of the censored variable.

2. KL Information of the Equilibrium Distribution Function

We consider only a nonnegative random variable so that $E(X)$ can be expressed as $\int_0^\infty \bar{F}(x)dx$. The equilibrium density function is defined as $f^*(x) = \bar{F}(x)/\int_0^\infty \bar{F}(x)dx$, and it can be easily shown that the exponential density function is the only continuous density function iff $f^*(x) = f(x)$. Suppose that the distribution functions of the random variables X, Y are F and G, respectively. Then the KL information of the equilibrium density functions $f^*(x) = \bar{F}(x)/E(X)$ and $g^*(y) = \bar{G}(y)/E(Y)$ can be written as

$$K(g^* : f^*) = \int_0^\infty \frac{\bar{G}(x)}{E(Y)} \log \frac{\bar{G}(x)}{\bar{F}(x)} dx + \log E(X) - \log E(Y).$$

$K(g^* : f^*)$ is location and scale invariant, and is well defined on the empirical distribution function. $K(g^* : f^*)$ is nonnegative, however, but the equality to zero means $g^*(x) = f^*(x)$, which is $\bar{F}(x)/E(X) = \bar{G}(x)/E(Y)$, not $F(x) = G(x)$. Hence, if we like to preserve the characterization property, the equal first moment condition should hold, which can be stated as follows.

Lemma 1. *For two nonnegative random variables where their first moments are finite and equal, $K(g^* : f^*)$ is nonnegative and the equality to zero holds iff $F(x) = G(x)$.*

Remark 1. Rao *et al.* (2004) introduced a cumulative residual entropy (CRE) as

$$\text{CRE}(F) = -\int_0^\infty \bar{F}(x) \log \bar{F}(x) dx$$

and Baratpour and Rad (2012) further suggested the cumulative residual KL information (CRKL) as

$$\mathrm{CRKL}(G : F) = \int_0^\infty \bar{G}(x) \log \frac{\bar{G}(x)}{\bar{F}(x)} dx + E(X) - E(Y),$$

which is well-defined on the empirical distribution function but is not scale invariant.

Because

$$K(g^* : f^*) - \frac{1}{E(Y)} \mathrm{CRKL}(G : F) = E(X) \left\{ \frac{E(Y)}{E(X)} - \log \frac{E(Y)}{E(X)} - 1 \right\},$$

we have

$$K(g^* : f^*) \approx \frac{1}{E(Y)} \mathrm{CRKL}(G : F)$$

if $E(X)/E(Y) \approx 1$.

$f^*(x)$ is different from $f(x)$ except for the exponential distribution. To measure the difference between $f^*(x)$ and $f(x)$, we can consider the approximation of KL information in terms of the Fisher information in Kullback (1959) as

$$K(f(x : \theta); f(x; \theta + \Delta\theta)) \approx \frac{1}{2} (\Delta\theta)^2 I_f(\theta),$$

where $I_f(\theta)$ is the Fisher information about θ in X.

Then we have the approximation of $K(f^*(x;\theta) : f^*(x; \theta + \Delta\theta))$ as follow.

$$K(f^*(x;\theta) : f^*(x; \theta + \Delta\theta)) \approx \frac{1}{2} (\Delta\theta)^2 I_{f^*}(\theta),$$

where

$$I_{f^*}(\theta) = -\int_0^\infty \frac{\bar{F}(x;\theta)}{E(X)} \frac{\partial^2}{\partial\theta^2} \log \bar{F}(x;\theta) dx + \frac{\partial^2}{\partial\theta^2} \log E(X).$$

Hence, $I_f(\theta) - I_{f^*}(\theta)$ can be considered in studying the departure of $f^*(x)$ from $f(x)$. It is evidently 0 for the exponential density function.

Example 1. If we consider the Weibull distribution, $F(x;\theta) = 1 - \exp(-x^\theta)$, $f^*(x)$ becomes the one-parameter generalized gamma (GGD) distribution, $\theta \exp(-x^\theta)/\Gamma(1/\theta)$. We note that the GGD in Stacy (1962) is actually a three parameters distribution as

$$f(x; a, d, p) = \frac{1}{\Gamma\left(\frac{d}{p}\right)} \frac{p}{a^d} x^{d-1} \exp\left(-\left(\frac{x}{a}\right)^p\right).$$

If $p = 1$, the GGD becomes the gamma distribution. For $a = d = 1$, we have one-parameter GGD, $p \exp(-x^p/\Gamma(1/p))$.

If we let $f_0(x)$ be the standard exponential distribution, we have

$$K(f_0(x) : f(x;\theta)) - K(f_0(x) : f^*(x;\theta)) \approx \frac{1}{2} (\theta - 1)^2 \left(I_f(\theta) - I_{f^*}(\theta) \right).$$

The Fisher information about θ in $f(x)$ and $f^*(x)$ can be obtained as

$$I_f(\theta) = \int_0^\infty \left(\frac{1}{\theta} + \log x\right)^2 \theta x^{\theta-1} \exp\left(-x^\theta\right) dx$$
$$= \frac{1}{\theta^2}\left\{(\gamma - 1)^2 + \frac{\pi^2}{6}\right\}$$

and

$$I_{f^*}(\theta) = \frac{1}{\Gamma(1 + 1/\theta)} \int_0^\infty \exp\left(-x^\theta\right) (\log x)^2 x^\theta dx + \frac{2\psi^{(0)}(1 + 1/\theta)}{\theta^3} + \frac{\psi^{(1)}(1 + 1/\theta)}{\theta^4},$$

where γ is the Euler's constant, and $\psi^{(i)}$ is the polygamma function of order i.

3. Sample Estimate and Application to a Goodness of Fit Test

In this section, we consider the application of $K(g^* : f^*)$ to the goodness of fit by letting $F(x) = F_\theta(x)$ and $G(x) = F_n(x)$ where θ is an unknown parameter and F_n is the empirical distribution function. The equilibrium density function can be estimated with the empirical distribution function as $f_n^*(x) = \bar{F}_n(x)/\bar{x}$, and $H(f_n^*)$ can be written as

$$H\left(f_n^*\right) = -\frac{1}{\bar{x}} \sum_{i=0}^{n} \frac{n-i}{n} \log \frac{n-i}{n} \left(x_{i+1:n} - x_{i:n}\right) + \log \bar{x}.$$

We also have

$$K\left(f_n^* : f_\theta^*\right) = \int_0^\infty \frac{\bar{F}_n(x)}{\bar{x}} \log \frac{\bar{F}_n(x)}{\bar{F}_\theta(x)} dx + \left(\log E_\theta(X) - \log \bar{x}\right).$$

To preserve the characterization, we need to estimate the unknown parameter in $F_\theta(x)$ by using the moment condition, $E_{F_n}(X) = E_\theta(Y)$. For example, $\hat{\theta} = \bar{x}$ satisfies the moment condition for an exponential distribution, $f(x;\theta) = \exp(-x/\theta)/\theta$. If we have other parameter η than θ, we can consider any consistent estimator η. For example, we can consider the additional criterion of minimum discriminant information (MDI) loss (see Soofi, 2000), which choose the parameter value minimizing the KL information, as

$$\hat{\eta} = \arg\min_\eta K\left(f_n^* : f_{\eta,\hat{\theta}}^*\right).$$

Then $K(f_n^* : f_{\hat{\theta}}^*)$ can be considered as a goodness of fit test statistic. Suppose that f_n^* is a consistent estimator of the true density function f^*, we have for a consistent estimator $\hat{\theta}$, by Slutsky's theorem,

$$K\left(f_n^* : f_{\hat{\theta}}^*\right) \to K\left(f^* : f_\theta^*\right) \quad \text{as } n \to \infty.$$

Under the null hypothesis $f^* = f_\theta^*$, $K(f_n^* : f_{\hat{\theta}}^*)$ is a consistent estimate of 0.

For the exponential distribution $f_\theta(x) = \exp(-x/\theta)/\theta$, the resulting parameter estimator is $\bar{x}$ so that $E_{F_n}(Y) = E_\theta(X)$. Then $K(f_n^* : f_{\hat{\theta}}^*)$ can be written as

$$K(f_n^* : f_{\hat{\theta}}^*) = \frac{1}{\bar{x}} \int_0^\infty \bar{F}_n(x) \log \bar{F}_n(x) dx + \frac{1}{\bar{x}} \int_0^\infty \frac{x\bar{F}_n(x)}{\bar{x}} dx$$
$$= \frac{1}{\bar{x}} \sum_{i=0}^{n} \frac{n-i}{n} \log \frac{n-i}{n} (x_{i+1:n} - x_{i:n}) + \frac{1}{\bar{x}^2} \sum_{i=0}^{n} \frac{n-i}{n} \int_{x_{i:n}}^{x_{i+1:n}} x dx.$$

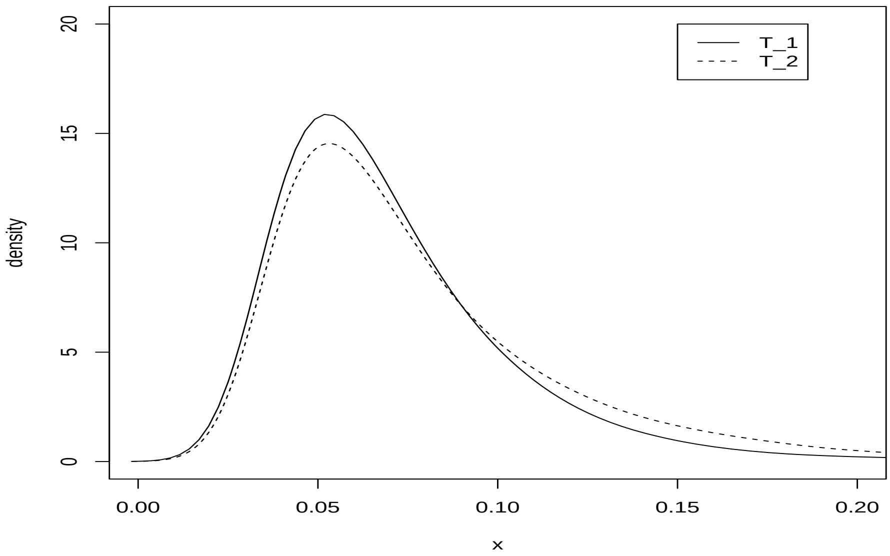

Figure 1: *Sampling distributions of T_1 and T_2 for a sample of size 20*

Table 1: The critical values of each statistic for $\alpha = 0.1, 0.05, 0.01$ with sample size $n = 20, 30, 40,$ and 50

	$\alpha = 0.1$		$\alpha = 0.05$		$\alpha = 0.01$	
	T_1	T_2	T_1	T_2	T_1	T_2
$n = 20$	0.1147	0.1329	0.1399	0.1607	0.2379	0.2237
$n = 30$	0.0877	0.0970	0.1080	0.1186	0.1859	0.1680
$n = 40$	0.0719	0.0780	0.0894	0.0958	0.1574	0.1390
$n = 50$	0.0610	0.0653	0.0764	0.0801	0.1361	0.1181

Then we have a test statistic, which can be simply written as

$$T_1 = \frac{1}{\bar{x}} \sum_{i=0}^{n} \frac{n-i}{n} \log \frac{n-i}{n} (x_{i+1:n} - x_{i:n}) + \frac{1}{2n\bar{x}^2} \sum_{i=1}^{n} x_i^2.$$

Baratpour and Rad (2012) considered the scaled CRKL as a goodness of fit test statistic as

$$T_2 = \frac{1}{\frac{\sum_{i=1}^n x_i^2}{2\sum_{i=1}^n x_i}} \left(\frac{\sum_{i=1}^n x_i^2}{2\sum_{i=1}^n x_i} - \mathrm{CRE}(F_n) \right). \tag{3.1}$$

The critical values of T_1 and T_2 were obtained with Monte Carlo simulation where the simulation size is 100000, and tabulated in Table 1. For a sample of size 20, the sampling distributions of T_1 and T_2 are displayed in Figure 1. We can see from Table 1 and Figure 1 that T_1 is closer to zero.

We obtain the powers for both T_1 and T_2 against gamma and Weibull alternatives with various shape parameter values. The results are presented in Figure 2, and we see that T_1 performs better than T_2 against alternatives with shape parameter value less than 1. Hence, it is recommended to use T_1 against decreasing failure rate (DFR) alternatives and to use T_2 against increasing failure rate (IFR) alternatives.

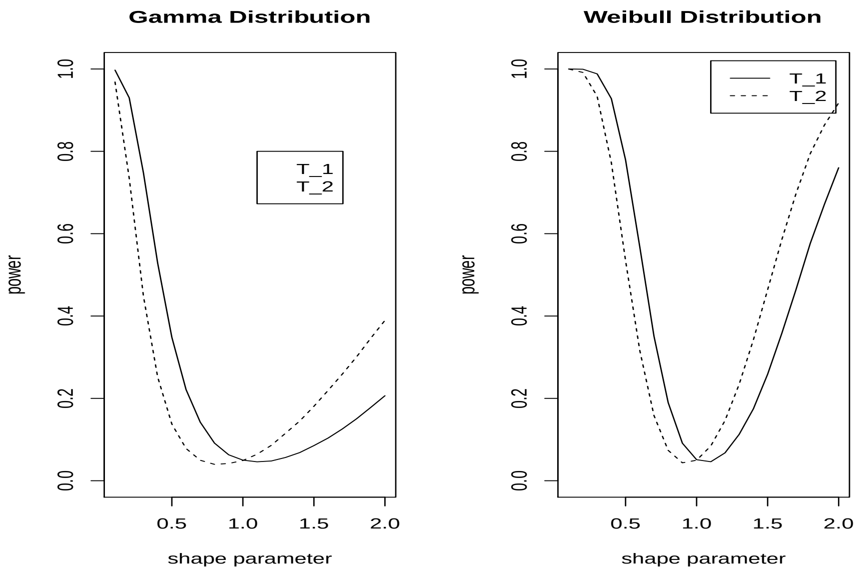

Figure 2: *Powers of* T_1 *and* T_2 *against gamma and Weibull alternatives* $(n = 20, \alpha = 0.05)$

4. Extension to the Censored Data

In this section, we consider the Type I *right* censored variable, $\min(X, C)$, where C is the censoring point assumed to be a constant. The density function of $\min(X, C)$ can be written as

$$f_C(x) = \begin{cases} f(x), & \text{if } x < C, \\ 1 - F(C), & \text{if } x = C, \\ 0, & \text{elsewhere.} \end{cases}$$

Then the KL information of $f_C(x)$ and $g_C(x)$ in terms of $g_C(x)$ can be obtained as

$$K(g_C : f_C) = \int_0^C g(x) \log \frac{g(x)}{f(x)} dx + (1 - G(C)) \log \frac{1 - G(C)}{1 - F(C)},$$

but is also not well-defined on the empirical distribution function.

For this censored variable, we can consider the corresponding equilibrium density function to $f_C(x)$, which can be obtained as

$$f_C^*(x) = \begin{cases} \dfrac{\bar{F}(x)}{\int_0^C \bar{F}(x) dx}, & \text{if } x < C, \\ 0, & \text{elsewhere.} \end{cases}$$

Then the KL information of the equilibrium density functions $f_C^*(x)$ and $g_C^*(x)$ can be written as

$$\begin{aligned} K\left(g_C^* : f_C^*\right) &= K_C(g_C^* : f_C^*) \\ &= \int_0^\infty \frac{\bar{G}(x)}{\int_0^C \bar{G}(x) dx} \log \frac{\bar{G}(x)}{\bar{F}(x)} dx + \log \int_0^C \bar{F}(x) dx - \log \int_0^C \bar{G}(x) dx. \end{aligned}$$

Table 2: Critical value estimates ($\alpha = 0.05$) of T based on 200,000 simulations

Censoring point (r)	$n = 10$	$n = 20$	$n = 30$	$n = 40$	$n = 50$
4	0.1307	0.0659	0.0441	0.0332	0.0265
5	0.1431	0.0722	0.0481	0.0363	0.0290
6	0.1538	0.0776	0.0519	0.0389	0.0312
7	0.1635	0.0827	0.0554	0.0414	0.0332
8	0.1721	0.0873	0.0583	0.0438	0.0351
9	0.1786	0.0915	0.0612	0.0459	0.0369
10	0.2094	0.0954	0.0639	0.0479	0.0385
15		0.1110	0.0753	0.0565	0.0454
20		0.1397	0.0837	0.0637	0.0511
25			0.0890	0.0694	0.0561
30			0.1078	0.0738	0.0599
35				0.0753	0.0630
40				0.0893	0.0655
45					0.0653
50					0.0765

$K(g_C^* : f_C^*)$ is location and scale invariant, and is well-defined on the empirical distribution function. $K(g_C^* : f_C^*)$ is nonnegative, and the equality to zero means $F(x) = G(x)$ under the condition that $\int_0^C \bar{F}(x)dx = \int_0^C \bar{G}(x)dx$.

If we let $f_{C,n}^*(x)$ and $f_{C,\theta}^*(x)$ be $\bar{F}_n(x)/\int_0^C \bar{F}_n(x)dx$ and $\bar{F}_\theta(x)/\int_0^C \bar{F}_\theta(x)dx$, respectively, we can establish a goodness of fit test statistic. For an exponential distribution, $f_\theta(x) = \exp(-x/\theta)/\theta$, $K(f_{C,n}^* : f_{C,\hat{\theta}}^*)$ can be obtained as

$$K\left(f_{C,n}^* : f_{C,\hat{\theta}}^*\right) = \frac{1}{\int_0^C \bar{F}_n(x)dx}\int_0^C \bar{F}_n(x)\log\bar{F}_n(x)dx + \frac{1}{\hat{\theta}\int_0^C \bar{F}_n(x)dx}\int_0^C x\bar{F}_n(x)dx,$$

where $\hat{\theta}$ is chosen so that $\int_0^C \bar{F}_{\hat{\theta}}(x)dx = \int_0^C \bar{F}_n(x)dx$.

Suppose that we have $r-1$ observations less than or equal to C so that $x_{r-1:n} \le C < x_{r:n}$ where $x_{i:n}$ is the i^{th} smallest value. Then $K(f_{C,n}^* : f_{C,\hat{\theta}}^*)$ can be written as

$$\begin{aligned} K\left(f_{C,n}^* : f_{C,\hat{\theta}}^*\right) &= \frac{1}{\sum_{i=0}^{r-1}\frac{n-i}{n}(x_{i+1:n}-x_{i:n})}\sum_{i=0}^{r-1}\frac{n-i}{n}\log\frac{n-i}{n}(x_{i+1:n}-x_{i:n}) \\ &+ \frac{1}{\hat{\theta}\sum_{i=0}^{r-1}\frac{n-i}{n}(x_{i+1:n}-x_{i:n})}\sum_{i=0}^{r-1}\frac{n-i}{2n}\left(x_{i+1:n}^2 - x_{i:n}^2\right), \end{aligned}$$

where $x_{r:n} = C$.

Because the censoring time depends on the unknown scale parameter, the critical value of $K(f_{C,n}^* : f_{C,\hat{\theta}}^*)$ under the exponential distribution is not free of θ for a given censoring time. Hence, we consider here the Type II censored case where C is taken to be $x_{r:n}$.

We made 200,000 Monte Carlo simulations to determine the 5% critical value under the exponential null distribution for $n = 10(10)50$, and the results are tabulated in Table 2. We compare the performance of the proposed test statistic K as a goodness of fit test statistic with two competing test statistics as follow.

Table 3: Power estimate (%) of .05 tests for Type-II Censoring against eleven alternatives of the exponential distributions based on 100,000 simulations; $n = 20$

Censoring points	$r = 5$			$r = 10$			$r = 15$		
Alternatives	K	T_p	W	T	T_p	W	T	T_p	W
Exp(1)	5.12	5.07	5.14	5.08	5.14	5.06	4.95	4.89	5.03
Gamma(0.5)	5.96	1.48	18.22	25.30	4.95	24.38	42.78	13.92	29.19
Gamma(1.5)	11.22	11.66	7.07	12.27	17.08	10.88	7.85	19.14	13.97
Gamma(2)	18.88	19.50	12.16	25.05	34.25	23.45	19.02	41.38	32.36
Log normal(0.5)	69.79	71.45	56.83	89.31	97.16	85.50	86.95	99.19	91.20
Log normal(1)	18.59	20.28	11.99	14.31	26.59	13.76	6.35	22.31	10.39
Log normal(1.5)	6.20	6.98	4.78	5.22	5.32	6.98	22.85	7.60	19.96
Weibull(0.5)	7.38	1.52	21.37	38.28	10.07	36.20	69.41	34.90	54.16
Weibull(1.5)	14.03	14.49	8.90	19.69	25.12	17.85	17.05	33.67	28.71
Weibull(2)	27.13	27.21	18.13	47.65	55.95	45.95	53.36	74.55	71.64
Uniform	6.49	6.46	5.33	10.50	10.80	8.51	16.49	23.90	24.29

1. We consider a censored version of a Shapiro-Wilk test (Samanta and Schwarz, 1988), which is known to be one of competing test statistics, as

$$W = \frac{(\sum_{i=1}^{r} y_i)^2}{r \sum_{i=2}^{r+1} \sum_{j=2}^{r+1} ((\min(i, j) - 1)/(r - \min(i, j) + 2)) y_{i-1} y_{j-1}},$$

where $y_i = (n - i + 1)(x_{(i)} - x_{(i-1)})$.

2. Some goodness of fit test statistics based on the Kullback-Leibler information or entropy difference have been proposed (Ebrahimi *et al.*, 1992). The censored version of Ebrahimi *et al.* (1992) has been proposed (Park, 2005) as

$$T_p = \frac{r}{n}\left(\log \hat{\theta}_{mle} + 1\right) - \frac{1}{n}\sum_{i=1}^{r} \log\left\{\frac{n}{2m}(x_{i+m:n} - x_{i-m:n})\right\} + \left(1 - \frac{r}{n}\right)\log\left(1 - \frac{r}{n}\right).$$

To compare the powers of the test statistics, we consider gamma (shape = 0.5, 1, 2), lognormal (scale = 0.5, 1, 1.5), Weibull (shape = 0.5, 1, 2) and uniform distributions as alternatives. We note that gamma(shape = 0.5) and Weibull(shape = 0.5) distributions are the DFR alternatives, while gamma(shape = 1.5, 2), Weibull(shape = 1.5, 2) and uniform distributions are the IFR alternatives. The power estimates are obtained with 100,000 Monte Carlo simulations, and are tabulated in Tables 3 and 4 for $n = 20, 40$. T_p shows best performances at increasing hazard alternatives, but we can notice from Table 3 that T_p is not an unbiased test. However K has comparable power to T_p in the case of light censoring. In general, K shows better powers than W ($r/n = 0.25, 0.5$) against increasing hazard alternatives in the case of heavy censoring and against decreasing hazard alternatives in the case of light censoring ($r/n = 0.75$).

5. Illustrated Example

In this section, we consider one real-life data analysis from Lawless (1982) illustrate the use of the proposed test statistic in a goodness of fit test for exponentiality. The data are given below, which consist of failure times for 36 appliances subjected to an automatic life test:

Data : 11, 35, 49, 170, 329, 381, 708, 958, 1062, 1167, 1594, 1925, 1990, 2223, 2327, 2400, 2451, 2471, 2551, 2565, 2568, 2694, 2702, 2761, 2831, 3034, 3059, 3112, 3214, 3478, 3504, 4329, 6367, 6976, 7846, 13403

Table 4: Power estimate (%) of .05 tests for Type-II Censoring against eleven alternatives of the exponential distributions based on 100,000 simulations; $n = 40$

Censoring points	$r = 10$			$r = 20$			$r = 30$		
Alternatives	K	T_p	W	T	T_p	W	T	T_p	W
Exp(1)	5.09	4.97	5.00	5.14	4.93	5.10	4.94	5.06	4.99
Gamma(0.5)	21.95	6.77	26.86	49.71	25.03	38.26	67.21	44.93	47.44
Gamma(1.5)	16.56	18.25	11.96	19.66	26.17	18.54	15.68	30.09	23.94
Gamma(2)	33.95	36.97	26.42	46.71	57.67	44.25	46.02	68.75	57.64
Log normal(0.5)	98.00	99.17	93.39	99.90	100.00	98.67	99.93	100.00	99.33
Log normal(1)	33.72	42.85	26.10	24.85	49.83	22.51	10.40	42.81	12.90
Log normal(1.5)	6.91	10.28	6.00	6.91	8.77	8.69	36.19	19.43	30.36
Weibull(0.5)	28.15	9.49	33.17	70.14	44.44	57.82	92.29	80.17	80.61
Weibull(1.5)	22.60	23.49	16.64	35.06	40.55	33.55	40.05	54.98	53.48
Weibull(2)	51.09	51.84	42.13	80.78	84.06	78.31	91.53	96.40	95.68
Uniform	7.52	7.00	5.85	14.74	13.25	13.05	35.96	38.50	48.58

Then we can calculate K to be 0.045719 whose corresponding p-value can be estimated from Monte Carlo simulations as 0.3940; therefore, the null hypothesis is not rejected. We note that T_p and W can be calculated to be 0.199325 and 0.031024, and the corresponding p-values can be estimated from Monte Carlo simulations as 0.1802 and 0.4092.

Now suppose that we have only first 18 failure times, which results in the Type II censored sample of $r = 18$. Then we can calculate K to be 0.029548 and estimate its p-value to be 0.3322. Hence, we cannot reject the null hypothesis. We note that T_p and W can be calculated to be 0.103372 and 0.069876, and the corresponding p-values can be estimated from Monte Carlo simulations as 0.5243 and 0.3390.

6. Conclusion

We suggest the KL information of the equilibrium density function to overcome the limitation of KL information that it is not well-defined on the empirical distribution. The KL information of the equilibrium density function has the nonnegativity property and characterization property under the equal first moment condition. Hence, we established an EDF-based goodness of fit test statistic, and studied its performance for an exponential distribution. We further extended the discussion to the censored case by considering the equilibrium density function of the censored variable. This EDF-based test is an omnibus test which is applicable to other distribution cases, but its sampling distribution is not tractable and the equal first moment condition should be considered in parameter estimation.

Acknowledgements

The authors are grateful to an anonymous referee and the associate editor for providing some useful comments on an earlier version of this manuscript.

References

Abo-Eleneen, Z. A. (2011). The entropy of progressively censored samples, *Entropy*, **13**, 437–449.

Andrews, F. C and Andrews, A. C. (1962). The form of the equilibrium distribution function, *Transactions of the Kansas Academy of Science.*, **65**, 247–256.

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

Bowman, A. W. (1992). Density based tests for goodness-of-fit, *Journal of Statistical Computation and Simulation*, **40**, 1–13.

Dudewicz, E. and van der Meulen, E. (1981). Entropy based tests of uniformity, *Journal of the American Statistical Association*, **76**, 967–974.

Ebrahimi, N., Habibullah, M. and Soofi, E. S. (1992). Testing exponentiality based on Kullback-Leibler information, *Journal of the Royal Statistical Society*, **54**, 739–748.

Jaynes, E. T. (1957). Information theory and statistical mechanics, *Physical Revies*, **106**, 620–630.

Kullback, S. (1959). *Information Theory and Statistics*, Wiley, New York.

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

Mosayeb, A. and Borzadaran, M. G. R. (2013). Kullback-Leibler information in view of an extended version of *k*-records, *Communications for Statistical Applications and Methods*, **20**, 1–13.

Nakamura, T. K. (2009). Relativistic equilibrium distribution by relative entropy maximization, *Europhysics letters*, **88**, 40009.

Park, S. (1995). The entropy of consecutive order statistics, *IEEE Transactions on Information Theory*, **41**, 2003–2007.

Park, S. (2005). Testing exponentiality based on Kullback-Leibler information with the type II cnesored data, *IEEE Transactions on Reliability*, **54**, 22–26.

Park, S. and Park, D. (2003). Correcting moments for goodness of fit tests based on two entropy estimates, *Journal of Statistical Computation and Simulation*, **73**, 685–694.

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

Samanta, M. and Schwarz, C. J. (1988). The Shapiro-Wilk test for exponentiality based on censored data, *Journal of the American Statistical Association*, **83**, 528–531.

Soofi, E. S. (2000). Principal information theoretic approaches, *Journal of the American Statistical Association*, **95**, 1349–1353.

Stacy, E. W. (1962). A generalization of the Gamma distribution, *Annals of Mathematical Statistics*, **33**, 1187–1192.

Theil, H. (1980). The entropy of the maximum entropy distribution, *Economics Letters*, **5**, 145–148.

Wong, K. M. and Chen, S. (1990). The entropy of ordered sequences and order statistics, *IEEE Transactions on Information Theory*, **36**, 276–284.

Received December 5, 2013; Revised January 22, 2014; Accepted March 14, 2014