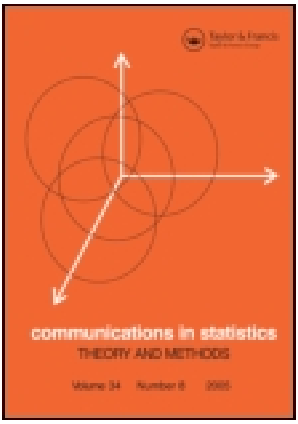

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

A Dynamic Discrimination Information Based On Cumulative Residual Entropy And Its Properties

A. Chamany [a] & S. Baratpour [a]
[a] Department of Statistics , School of Mathematical Sciences, Ferdowsi University of Mashhad , Mashhad , Iran
Published online: 14 Feb 2014.

Communications in Statistics—Theory and Methods, 43: 1041–1049, 2014

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610926.2012.729639

A Dynamic Discrimination Information Based On Cumulative Residual Entropy And Its Properties

A. CHAMANY AND S. BARATPOUR

Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

In this article, we introduce a measure of discrepancy between two life-time distributions based on cumulative residual entropy. The dynamic form of this measure is considered and some of its properties are obtained. The relations between dynamic form and some well-known concepts in reliability such as mean residual life-time, hazard rate order, and new better (worse) than used are studied.

Keywords Cumulative residual entropy; Hazard rate order; Kullback-Liebler divergence; Mean residual life-time function; New better (worse) than used; Shannon entropy.

Mathematics Subject Classification Primary 94A17; Secondary 62N05, 60E15.

1. Introduction

As is well known, Shannon (1948) established an important part of information theory that today has many applications in various fields. He proposed a way of achieving uncertainty associated with a probability distribution. For continuous and non negative random variable X, Shannon's differential entropy is defined by

$$H(X) = E(-\log f(X)) = -\int_0^\infty f(x)\log f(x)\mathrm{d}x,$$

where f is the probability density function (pdf) of X. As can be seen, $H(X)$ equals to the expectation of $-\log f(X)$.

Moreover, Kullback and Leibler (1951) proposed a non symmetric measure of distance between two distributions. K-L divergence is defined by

$$I(X, Y) = \int_0^\infty f(x)\log\frac{f(x)}{g(x)}\mathrm{d}x,$$

Received November 20, 2011; Accepted September 04, 2012
Address correspondence to S. Baratpour, Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran; E-mail: baratpur@math.um.ac.ir

where f and g are pdfs of X and Y, respectively. Ebrahimi and Kirmani (1996) defined a time-dependent (dynamic) form of Kullback-Leibler discrimination by comparing residual life distributions at each age $t \geq 0$, and they derived important properties of it. This discrepancy measure is the K-L divergence of random variables $[X - t|X > t]$ and $[Y - t|Y > t]$ and is defined as

$$I(X, Y; t) = \int_t^\infty f_t(x) \log \frac{f_t(x)}{g_t(x)} \mathrm{d}x,$$

where $f_t(x) = \frac{f(x)}{\overline{F}(t)}$.

However, the Shannon entropy has certain disadvantages. For example, it requires the knowledge of density function for non-discrete random variables, the discrete Shonnon entropy dose not converge to its continuous analogous, and in order to estimate the Shannon entropy for a continuous density, one has to obtain the density estimation, which is not a trivial task. Recently, Rao et al. (2004) developed a new measure of information called cumulative residual entropy (CRE) that is defined as

$$\mathscr{E}(X) = -\int_0^\infty \overline{F}(x) \log \overline{F}(x) \mathrm{d}x,$$

where $\overline{F} = 1 - F$ is the survival function of X. The basic idea in their definition was to replace the density function by the survival function in Shannon's definition. CRE is more general than the Shannon entropy and possesses more general mathematical properties. It can easily be estimated from sample data and this estimation asymptotically converges to the true value. An analougus definition can be considered for discrete distributions as

$$\mathscr{E}(X) = -\sum_x P(X > x) \log P(X > x).$$

CRE has applications in reliability engineering and computer vision, for more details see Rao (2005). Baratpour (2010) studied conditions under which the CRE of the first-order statistics can uniquely determine the parent distribution. Asadi and Zohrevand (2007) considered the dynamic version of CRE and called it dynamic cumulative residual entropy (DCRE). They studied the relations between DCRE and well-known reliability measures. DCRE is defined by

$$\mathscr{E}(X; t) = -\int_t^\infty \frac{\overline{F}(x)}{\overline{F}(t)} \log \frac{\overline{F}(x)}{\overline{F}(t)} \mathrm{d}x.$$

This measure is the CRE of random variable $[X - t|X > t]$.

The rest of this article is organized as follows. In Sec. 2, the aim is to present the dynamic version of Kullback-Leibler divergence between two distributions based on cumulative residual entropy and obtain some of its properties. In Sec. 3, the relations between dynamic form and some well-known concepts in reliability such as the mean residual life-time, hazard rate order and new better than used (NBU), and new worth than used (NWU) classes, are studied.

2. A Dynamic Form of Divergence Between Two Distributions

Let X and Y be two non negative continuous random variables with cumulative distribution functions F and G, and survival functions $\overline{F} = 1 - F$ and $\overline{G} = 1 - G$, respectively. We first introduce a new measure of distance between two distributions that is similar to the Kullback-Leibler divergence, but using the survival function rather than the density function and call it cumulative Kullback-Leibler (CKL) divergence.

CKL between two distributions F and G is defined by

$$\mathscr{C}\mathscr{E}(X, Y) = \int_0^\infty \overline{F}(x) \log \frac{\overline{F}(x)}{\overline{G}(x)} \mathrm{d}x - \big(E(X) - E(Y)\big),$$

where $E(X)$ is the expectation of X. It can be shown that $\mathscr{C}\mathscr{E}(X, Y) \geq 0$, and equality holds if and only if $F = G, \quad a.e$, that is concluded easily by using the log-sum inequality and the inequality $a \log \frac{a}{b} \geq a - b$, for all $a, b > 0$ (Baratpour and Habibi Rad, 2012). Now, we consider the dynamic version of CKL and call it dynamic cumulative Kullback-Leibler (DCKL) divergence that is defined by

$$\mathscr{C}\mathscr{E}(X, Y; t) = \int_t^\infty \left[\overline{F}_t(x) \log \frac{\overline{F}_t(x)}{\overline{G}_t(x)} + \overline{G}_t(x) - \overline{F}_t(x) \right] \mathrm{d}x,$$

where $\overline{F}_t(x) = \frac{\overline{F}(x)}{\overline{F}(t)}$. Moreover, let $\lambda_F(x) = f(x)/\overline{F}(x)$ and $\lambda_G(x) = g(x)/\overline{G}(x)$ be the hazard rate functions of X and Y, respectively, whereas $m_F(t) = E(X - t|X > t) = \int_t^\infty \frac{\overline{F}(x)}{\overline{F}(t)} \mathrm{d}x$ and similarly $m_G(t)$ denote their MRL functions. It can be shown that

$$\begin{aligned}
\mathscr{C}\mathscr{E}(X, Y; t) &= \int_t^\infty \left[\frac{\overline{F}(x)}{\overline{F}(t)} \log \frac{\overline{F}(x)/\overline{F}(t)}{\overline{G}(x)/\overline{G}(t)} + \frac{\overline{G}(x)}{\overline{G}(t)} - \frac{\overline{F}(x)}{\overline{F}(t)} \right] \mathrm{d}x \\
&= m_G(t) - m_F(t) - \mathscr{E}(X; t) - \int_t^\infty \frac{\overline{F}(x)}{\overline{F}(t)} \log \frac{\overline{G}(x)}{\overline{G}(t)} \mathrm{d}x.
\end{aligned}$$

Also, we can show that $\lim_{t \to 0^+} \mathscr{C}\mathscr{E}(X, Y; t) = \mathscr{C}\mathscr{E}(X, Y)$.

CKL and DCKL can be used for goodness of fit testing. Baratpour and Habibi Rad (2012) based on CKL, developed a consistent test statistic for testing the hypothesis of exponentiality against some alternatives.

Note that $\mathscr{C}\mathscr{E}(X, Y; t)$ is not symmetric, i.e., $\mathscr{C}\mathscr{E}(X, Y; t) \neq \mathscr{C}\mathscr{E}(Y, X; t)$. A symmetrized version can easily be constructed as

$$\hat{\mathscr{C}\mathscr{E}}(X, Y; t) = \frac{1}{2}(\mathscr{C}E(X, Y; t) + \mathscr{C}E(Y, X; t)).$$

Therefore,

$$\hat{\mathscr{C}\mathscr{E}}(X, Y; t) = \frac{1}{2} \int_t^\infty (\overline{F}_t(x) - \overline{G}_t(x)) \log \frac{\overline{F}_t(x)}{\overline{G}_t(x)} \mathrm{d}x.$$

As can be seen, for the symmetrized version of DCKL, the difference between MRLs is omitted.

In the following example, we obtain DCKL for some common distributions.

Example 2.1.

(a) Let X and Y be distributed as exponentials with mean λ and β, respectively, It is well known that $m_F(t) = \mathscr{E}(X) = \mathscr{E}(X; t) = \lambda$. Then, it can be easily shown that $\mathscr{CE}(X, Y; t) = \mathscr{CE}(X, Y) = \frac{(\lambda-\beta)^2}{\beta}$. As can be seen, $\mathscr{CE}(X, Y; t)$ does not depend on t.

(b) If X and Y have Pareto distributions with survival functions $\overline{F}(x) = (\frac{\beta}{x+\beta})^{\lambda}$ and $\overline{G}(x) = (\frac{\beta}{x+\beta})^{\gamma}$, $\beta > 0$, $\lambda > 1, \gamma > 1, x \geq 0$, respectively. Then, we can easily show that $m_F(t) = \frac{1}{\lambda-1}(\beta + t), \mathscr{E}(X, t) = \frac{\lambda}{(\lambda-1)^2}(\beta + t)$ and $\mathscr{CE}(X, Y; t) = (\beta + t)(\frac{(\lambda-\gamma)^2}{(\lambda-1)^2(\gamma-1)}$. As you see, $m_F(t), \mathscr{E}(X; t)$, and $\mathscr{CE}(X, Y; t)$ are increasing linear functions of t.

(c) Let survival functions of X and Y be $\overline{F}(x) = (1-x)^{\lambda}$ and $\overline{G}(x) = (1-x)^{\beta}, 0 < x < 1, \lambda > 0, \beta > 0$, respectively. It is not hard to see that $m_F(t) = \frac{1-t}{\lambda+1}, \mathscr{E}(X; t) = \frac{\lambda}{(1+\lambda)^2}(1-t)$ and $\mathscr{CE}(X, Y; t) = \frac{(1-t)(\beta-\lambda)^2}{(1+\lambda)^2(\beta+1)}$. As you see again, $m_F(t), \mathscr{E}(X; t)$, and $\mathscr{CE}(X, Y; t)$ are decreasing linear functions of t.

Some identities of DCKL are:

- $\mathscr{CE}(X, Y; t) \geq 0$ and equality holds, if and only if $\overline{F}(x) = \overline{G}(x)$. The proof is similar to the proof for *CKL*;
- $\mathscr{CE}(X, Y; t)$ is convex. The proof is analogous to the proof for convexity of the relative entropy (see Cover and Thomas, 2006, p. 32, Theorem 2.7.2).

In the following theorem, the monotonicity of DCKL is studied.

Theorem 2.1. *$\mathscr{CE}(X, Y; t)$ is non decreasing (non increasing), if and only if*

$$\mathscr{CE}(X, Y; t) \geq (\leq) \left(1 - \frac{\lambda_G(t)}{\lambda_F(t)}\right)(m_G(t) - m_F(t)), \ \ t \geq 0. \tag{1}$$

Proof. By differentiating $\mathscr{CE}(X, Y; t)$ with respect to t, we have

$$\frac{\partial}{\partial t}\mathscr{C}E(X, Y; t) = \lambda_F(t)[\mathscr{C}E(X, Y; t)] + (\lambda_G(t) - \lambda_F(t))(m_G(t) - m_F(t)), \ \ t \geq 0,$$

which implies the assertion of (1).

Nanda et al. (2006) showed that in the proportional mean residual life (PMRL) models which is presented as $m_G(t) = cm_F(t)$, the relation between $\lambda_G(t)$ and $\lambda_F(t)$ is

$$\lambda_G(t) = \lambda_F(t) + \frac{1-c}{cm_F(t)}.$$

By using Theorem 2.1, we have the following corollary for the PMRL models.

Corollary 2.1. *In PMRL model, $\mathscr{CE}(X, Y; t)$ is increasing (decreasing), if and only if*

$$\mathscr{CE}(X, Y; t) \leq (\geq) \frac{(1-c)^2}{c\lambda_F(t)}.$$

3. Some Results on $\mathscr{CE}(X, Y; t)$

Stochastic orders and inequalities have been used in many areas of probability and statistics such as reliability theory, survival analysis, economics, management, etc. One of this stochastic orders is hazard rate order, that is usually applied to random life times. X is said to be smaller than Y in the hazard rate order (denoted as $X \leq_{hr} Y$), if $\lambda_X(t) \geq \lambda_Y(t),\ \ t \geq 0$. In the following theorem, by using the concept of the hazard rate order, we show that under some conditions, triangle inequality is satisfied .

Theorem 3.1. *Let us consider three random lifetimes* X, Y *and* Z. *If* (i) $Y \geq_{hr} X$ *and* $Y \geq_{hr} Z$, *or* (ii) $Y \leq_{hr} X$ *and* $Y \leq_{hr} Z$, *then the triangle inequality is satisfied*:

$$\mathscr{CE}(X, Y; t) + \mathscr{C}E(Y, Z; t) \geq \mathscr{C}E(X, Z; t).$$

Proof.

$$\begin{aligned}
&\mathscr{CE}(X, Y; t) + \mathscr{CE}(Y, Z; t) \\
&= \int_t^{\infty} \left[\overline{F}_t(x) \log \frac{\overline{F}_t(x)}{\overline{G}_t(x)} + \overline{G}_t(x) - \overline{F}_t(x) \right] \mathrm{d}x + \int_t^{\infty} [\overline{G}_t(x) \log \frac{\overline{G}_t(x)}{\overline{H}_t(x)} \\
&\quad + \overline{H}_t(x) - \overline{G}_t(x)]\mathrm{d}x \\
&= \int_t^{\infty} \left[\overline{F}_t(x) \log \frac{\overline{F}_t(x)}{\overline{H}_t(x)} + \overline{H}_t(x) - \overline{F}_t(x) \right] \mathrm{d}x + \int_t^{\infty} (\overline{F}_t(x) - \overline{G}_t(x)) \log \frac{\overline{H}_t(x)}{\overline{G}_t(x)} \mathrm{d}x \\
&= \mathscr{CE}(X, Z; t) + \int_t^{\infty} (\overline{F}_t(x) - \overline{G}_t(x)) \log \frac{\overline{H}_t(x)}{\overline{G}_t(x)} \mathrm{d}x, \qquad (2)
\end{aligned}$$

where $\overline{F}$, $\overline{G}$ and $\overline{H}$, are survival functions of X, Y, and Z, respectively. By noting that if X and Y be two arbitrary random variables with survival functions $\bar{F}$ and $\overline{G}$, respectively, and $X \leq_{hr} Y$, then $\frac{\overline{F}_t(x)}{\overline{G}_t(x)} \leq 1$. By using (i) or (ii) it follows that the second term of (2) is non-negative which completes the assertion.

In the next theorem, it has been shown that ordering in hazard rate concludes that DCKL is less than the difference between MRLs.

Theorem 3.2. *If* $X \leq_{hr} Y$, *then*

$$\mathscr{CE}(X, Y; t) \leq m_G(t) - m_F(t). \qquad (3)$$

Proof. From $X \leq_{hr} Y$, we have $\frac{\overline{F}_t(x)}{\overline{G}_t(x)} \leq 1$ for all $t > 0$. Therefore,

$$\mathscr{CE}(X, Y; t) = \int_t^{\infty} \left[\overline{F}_t(x) \log \frac{\overline{F}_t(x)}{\overline{G}_t(x)} + \overline{G}_t(x) - \overline{F}_t(x) \right] \mathrm{d}x \leq \int_t^{\infty} \left[\overline{G}_t(x) - \overline{F}_t(x) \right] \mathrm{d}x,$$

which gives (3).

The non negative random variable X is said to be new better (worse) than used in expectation (denoted by NBUE (NWUE)), if

$$\frac{\int_t^\infty \overline{F}(x)dx}{\overline{F}(t)} \leq (\geq) E(X),$$

for all $t \geq 0$ (see Shaked and Shanthikumar, 2007). The following corollary is obtained from Theorem 3.3.

Corollary 3.1. *If $X \leq_{hr} Y$, and X and Y be NWUE and NBUE, respectively, then*

$$\mathscr{CE}(X, Y; t) \leq E(Y) - E(X).$$

In the literature, a good number of models have been introduced for modeling and analyzing failure time data. proportional hazard rate model introduced by Lehmann (1953) gained wide spread extensions after the rationale behind the model was explained by Cox (1972).

Remark 3.1. For the proportional hazard rate model which is defined as

$$\overline{G}(x) = [\overline{F}(x)]^\theta, \quad x > 0, \quad \theta > 0,$$

if $\theta > 1$, then

$$\mathscr{CE}(X, Y; t) \leq (\theta - 1)\mathscr{E}(X, t).$$

The following theorem gives some conditions for $\mathscr{C}E(X, Y; t)$ to be more than $\mathscr{CE}(Z, Y; t)$.

Theorem 3.3. *Let us consider three random lifetimes X, Y, and Z with survival functions $\overline{F}$, $\overline{G}$ and $\overline{H}$, respectively. If $X \leq_{hr} Z \leq_{hr} Y$, then*

$$\mathscr{CE}(Y, X; t) \geq max\{\mathscr{CE}(Y, Z; t), \mathscr{CE}(Z, X; t)\}.$$

Proof.

$$\begin{aligned}
\mathscr{CE}(Y, X; t) - \mathscr{CE}(Y, Z; t) &= \int_t^\infty [\overline{G}_t(x) \log \frac{\overline{H}_t(x)}{\overline{F}_t(x)} \mathrm{d}x + m_F(t) - m_H(t) \\
&\geq \int_t^\infty [\overline{H}_t(x) \log \frac{\overline{H}_t(x)}{\overline{F}_t(x)} \mathrm{d}x + m_F(t) - m_H(t) \\
&= \mathscr{CE}(Z, X; t), \qquad (4)
\end{aligned}$$

where inequality is obtained by $\overline{H}_t(x) \leq \overline{G}_t(x)$. By (4) and non negativity of DCKl, we conclude that $\mathscr{CE}(Y, X; t) \geq \mathscr{C}E(Y, Z; t)$ and $\mathscr{CE}(Y, X; t) \geq \mathscr{E}(Z, X; t)$. Hence, the proof is complete.

Mixture distributions play an important role in nearby every field of statistics. Let X and Y be two random variables with distribution functions F and G, respectively, and Z be a random variable with distribution function $pF + (1 - p)G$, for some $p \in (0,1)$. Now, if $X \leq_{hr} Y$, then $X \leq_{hr} Z \leq_{hr} Y$ (see Shaked and Shanthikumar, 2007, p. 27, Theorem 1.B.22), that consequently satisfies the condition of Theorem 5. Thus, we have the following corollary.

Corollary 3.2. *If $X \leq_{hr} Y$ and Z be a mixture of X and Y, then*

$$\mathscr{CE}(Y, X; t) \geq max\{\mathscr{CE}(Y, Z; t), \mathscr{C}E(Z, X; t)\}.$$

The next theorem represents that the difference between $\mathscr{CE}(X, Y; t)$ and $\mathscr{CE}(X, Z; t)$ is less than the MRLs of Y and Z.

Theorem 3.4. *Let X, Y, and Z be three random lifetimes with survival functions $\overline{F}$, $\overline{G}$, and $\overline{H}$, and mean residual life-time functions m_F, m_G, and m_H, respectively. If $Z \leq_{hr} Y$, then*

$$\mathscr{CE}(X, Y; t) - \mathscr{CE}(X, Z; t) \leq m_G(t) - m_H(t).$$

Proof. The proof is analogous to that of Theorem 3.2 and hence omitted.

In the next theorem, we find an upper bound for the difference between $\mathscr{CE}(X, Y)$ and $\mathscr{CE}(X, Y; t)$. We first recall that the non negative random variable X is said to be new better (worse) than used (denoted by NBU (NWU)), if $\overline{F}(x + y) \leq (\geq)\overline{F}(x)\overline{F}(y)$ for all $x, y > 0$.

Theorem 3.5. *If X and Y be NWU and NBU, respectively, then*

$$\mathscr{CE}(X, Y) - \mathscr{CE}(X, Y; t) \leq [E(Y) - E(X)] - [m_G(t) - m_F(t)].$$

Proof. By definitions of NWU and NBU, we have

$$\frac{\overline{F}(x + t)}{\overline{F}(t)} \geq \overline{F}(x)$$

and

$$\frac{\overline{G}(x + t)}{\overline{G}(t)} \leq \overline{G}(x).$$

Therefore,

$$\int_0^\infty \frac{\overline{F}(x + t)}{\overline{F}(t)} \log \frac{\overline{F}(x + t)/\overline{F}(t)}{\overline{G}(x + t)/\overline{G}(t)} \mathrm{d}x \geq \int_0^\infty \overline{F}(x) \log \frac{\overline{F}(x)}{\overline{G}(x)} \mathrm{d}x.$$

Thus,

$$\mathscr{CE}(X, Y; t) - [m_G(t) - m_F(t)] \geq \mathscr{CE}(X, Y) - [E(Y) - E(X)],$$

which completes the proof.

Theorem 3.6.

(*i*) *For an increasing function* ϕ *on* $(0, \infty)$, *if* $a \leq \phi' \leq b, \quad a, b > 0$, *where* ϕ' *is the derivative of* ϕ, *then*

$$a\ \mathscr{CE}(X, Y; \phi^{-1}(t)) \leq \mathscr{CE}(\phi(X), \phi(Y); t) \leq b\ \mathscr{CE}(X, Y; \phi^{-1}(t)).$$

(*ii*) $\mathscr{CE}(bX, bY; t) = b\ \mathscr{CE}\big(X, Y; \frac{t}{b}\big), \ b > 0.$

Proof.

$$\begin{aligned}
\mathscr{CE}(\phi(X), \phi(Y); t) &= \int_t^\infty \left[\frac{\overline{F}(\phi^{-1}(x))}{\overline{F}(\phi^{-1}(t))} \log \frac{\overline{F}(\phi^{-1}(x))/\overline{F}(\phi^{-1}(t))}{\overline{G}(\phi^{-1}(x))/\overline{G}(\phi^{-1}(t))} \right. \\
&\quad \left. + \frac{\overline{G}(\phi^{-1}(x))}{\overline{G}(\phi^{-1}(t))} - \frac{\overline{F}(\phi^{-1}(x))}{\overline{F}(\phi^{-1}(t))} \right] \mathrm{d}x \\
&= \int_{\phi^{-1}(t)}^\infty \phi'(y) \left[\frac{\overline{F}(y)}{\overline{F}(\phi^{-1}(t))} \log \frac{\overline{F}(y)/\overline{F}(\phi^{-1}(t))}{\overline{G}(y)/\overline{G}(\phi^{-1}(t))} \right. \\
&\quad \left. + \frac{\overline{G}(y)}{\overline{G}(\phi^{-1}(t))} - \frac{\overline{F}(y)}{\overline{F}(\phi^{-1}(t))} \right] \mathrm{d}y.
\end{aligned} \tag{5}$$

Therefore, if $a \leq \phi'(y) \leq b$, then (i) results. If $\phi(x) = bx, \ b > 0$, then by (5), we have

$$\begin{aligned}
\mathscr{CE}(bX, bY; t) &= b \int_{\frac{t}{b}}^\infty \left[\frac{\overline{F}(y)}{\overline{F}(\frac{t}{b})} \log \frac{\overline{F}(y)/\overline{F}(\frac{t}{b})}{\overline{G}(y)/\overline{G}(\frac{t}{b})} + \frac{\overline{G}(y)}{\overline{G}(\frac{t}{b})} - \frac{\overline{F}(y)}{\overline{F}(\frac{t}{b})} \right] \mathrm{d}y \\
&= b\ \mathscr{CE}\left(X, Y; \frac{t}{b}\right).
\end{aligned}$$

Thus, (ii) is concluded.

Acknowledgment

The authors would like to thank the two referees for their careful reading and their useful comments which led to this considerably improved version.

References

Asadi, M., Zohrevand, Y. (2007). On the dynamic cumulative entropy. *JSPI* 137:1931–1941.

Baratpour, S. (2010). Characterizations based on cumulative residual entropy of first order statistics. *Commun. Statist. Theor. Meth.* 39:3645–3651.

Baratpour, S., Habibi Rad, A. (2012). Testing goodness-of-fit for exponential distribution based on cumulative residual entropy. *Commun. Statist. Theor. Meth.* 41:1387–1396.

Cox, D. R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. B* 34:187–202.

Cover, T. M., Thomas, Joy A. (2006). *Elements of Information Theory*. New York: Wiely.

Ebrahimi, N., Kirmani, S. N. U. A. (1996). A measure of discrimination between two residual life-time distributions and its applications. *Ann. Instit. Statist. Math.* 48: 257–265.

Kullback, S., Leibler, R. A. (1951). On information and sufficiency. *Ann. Math. Statist.* 22: 79–86.

Lehmann, E. L. (1953). The power of rank test. *Ann. Math. Statist.* 24:23–43.

Nanda, A. K., Bhattacharjee, S., Alam. S. S. (2006). Properties of proportional mean residual life model. *SPL* 76:880–890.

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

Rao, M., Chen, Y., Vemuri, B. C., Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE Trans. Inform. Theor.* 50(6):1220–12222.

Shaked, M., Shanthikumar, J. G. (2007). *Stochastic Orders.* New York: Springer.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27: 379–423.