Probability in the Engineering and Informational Sciences, **34**, 2020, 605–625.
doi:10.1017/S0269964819000196

ON CUMULATIVE RESIDUAL EXTROPY

S. M. A. Jahanshahi and H. Zarei
Department of Statistics, University of Sistan and Baluchestan, Zahedan, Iran
E-mail: mjahan@math.usb.ac.ir

A. H. Khammar
Department of Statistics, University of Birjand, Iran

Recently, an alternative measure of uncertainty called extropy is proposed by Lad et al. [12]. The extropy is a dual of entropy which has been considered by researchers. In this article, we introduce an alternative measure of uncertainty of random variable which we call it cumulative residual extropy. This measure is based on the cumulative distribution function F. Some properties of the proposed measure, such as its estimation and applications, are studied. Finally, some numerical examples for illustrating the theory are included.

Keywords: cumulative residual extropy, estimation, gini index, independence, proportional hazard model, risk measurs, stop-loss transform, stochastic orders

1. INTRODUCTION

Let X be a non-negative and absolutely continuous random variable (rv) with probability density function (pdf) f. To measure the uncertainty contained in X, the entropy was defined by Shannon [24] as follows:

$$H(X) = -\int_0^{+\infty} f(x) \log f(x) dx.$$

where "log" is the natural logarithm with the convention $0 \log 0 = 0$. Among various intensions to define possible alternative information theoretic measures, Rao et al. [20] proposed the cumulative residual entropy (CRE) and studied its properties. This measure replaces density function by the survival function. For a non-negative rv X with cumulative distribution function (cdf) F and survival function (sf) $\bar{F} = 1 - F$ the CRE was defined as follows:

$$\xi(X) = -\int_0^{+\infty} \bar{F}(x) \log \bar{F}(x) dx.$$

Lad et al. [12] provided a completion to theories of information based on entropy, resolving a longstanding question in its axiomatization as proposed by Shannon [24] and followed by Jaynes [11]. They showed that Shannon's entropy function has a complementary dual

 0269-9648/19 $25.00

function which is called “extropy”. They also analyzed the extropy function for densities, showing that relative extropy constitutes a dual to the Kullback–Leibler divergence, widely recognized as the continuous entropy measure. For non-negative rv X, its extropy was defined as

$$J(X) = -\frac{1}{2}\int_0^{+\infty} f^2(x)dx. \tag{1}$$

Several properties of this new information measure such as the maximum extropy distribution and its statistical applications were displayed in Lad et al. [12].

One statistical application of extropy is to score the forecasting distributions. For example, under the total log scoring rule, the expected score of a forecasting distribution equals the negative sum of the entropy and extropy of this distribution (see Gneiting and Raftery, [9]). In commercial or scientific areas such as astronomical measurements of heat distributions in galaxies, the extropy has been universally investigated (see Furuichi and Mitroi [8]; Vontobel [25]). Most recently, Qiu [13] further studied this new measure, exploring some characterization results, monotone properties and lower bounds of extropy of order statistics and record values. He also investigated the symmetric properties of extropy of order statistics. Qiu and Jia [14] proposed residual extropy to measure residual uncertainty of a non-negative rv as

$$J(X;t) = -\frac{1}{2\bar{F}^2(t)}\int_t^{+\infty} f^2(x)dx, \quad t \geq 0.$$

They studied monotone properties and characterization results of this measure and discussed similar properties of the proposed measure of order statistics. Qiu and Jia [15] proposed two estimators for extropy and they developed a goodness-of-fit test for standard uniform distribution. Yang et al. [28] studied the relations between extropy and variational distance and determined the distribution which attains the minimum or maximum extropy among these distributions within a given variation distance from any given probability distribution. Qiu et al. [16] explored an expression of the extropy of a mixed systems lifetime.

In this paper we introduce a new measure of uncertainty that will be called cumulative residual extropy (CREX). The basic idea is to replace the pdf with the cdf in extropy definition (1). The cdf is more regular than the pdf, because the pdf is computed as the derivative of the cdf. Moreover, in practice what is of interest and/or measurable is cdf. For example, if the rv is the life span of a machine, then the event of interest is not whether the life span equals t, but rather whether the life span exceeds t.

The rest of the paper is organized as follows: Section 2 contains the definition of CREX and a description of its properties in the form of several theorems. We determine upper and lower bounds and inequalities concerning CREX, moreover we give a relationship between CREX and extropy. Also, we show that the exponentially distributed rv have maximum CREX. In Section 3, we discuss proportional hazard model and Gini index. Some stochastic order properties are discussed in Section 4. In Section 5, we consider the problem of estimating the CREX by means of the empirical CREX. In this regards, we use two different empirical estimators of cdf to estimate CREX. Finally, in Section 6, we derive two applications of the CREX to risk measure and independence. In particular, we study the absolute value of CREX as a risk measure. For comparing the absolute value of CREX with the standard deviation and the right-tail risk measure, several examples are also given. Moreover by using CREX, we introduce a new measure of independence between two rvs.

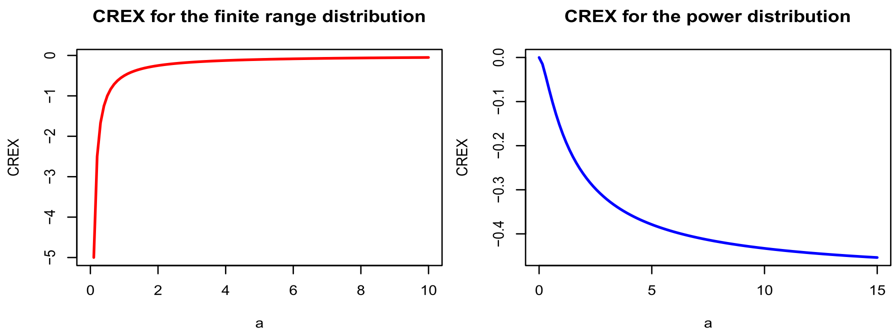

FIGURE **1.** Graph of $\xi J(X)$ for $0 \leq a \leq 10$, $b = 0.5$ (left panel) and $0 \leq a \leq 15$ (right panel).

2. CUMULATIVE RESIDUAL EXTROPY

For rv X, by analogy to Rao et al. [20], we propose the following definition of the CREX.

DEFINITION 2.1: *For a non-negative rv X with an absolutely continuous sf $\bar{F}$, we define the CREX as*

$$\xi J(X) = -\frac{1}{2} \int_0^{+\infty} \bar{F}^2(x) dx. \quad \textbf{(2)}$$

EXAMPLE 2.2: *To study the $\xi J(X)$ value for some distributions we have*

(I) If X is exponentially distributed with parameter $\lambda > 0$, then $\xi J(X) = -(1/4\lambda)$.
(II) If X is uniformly distributed in $[0, b]$, then $\xi J(X) = -(b/6)$.
(III) If X is power distribution with parameter $a > 0$, then $\xi J(X) = -(a^2/(((a + 1)(2a + 1)))$.
(IV) Let X is finite range distribution with sf $\bar{F}(x) = (1 - ax)^b$, $a > 0$, $b > 0$, $x \in (0, 1/a)$, then $\xi J(X) = -(1/(a(1 + 2b)))$.
(V) Let X has Pareto distribution with parameters (α, β), so that $\bar{F}(x) = (\beta^\alpha)/((x + \beta)^\alpha)$, $x \geq 0$, $\beta > 0$. If $0 < \alpha < 1$ then $\xi J(X) = -\infty$, where as if $\alpha \geq 1$, then $\xi J(X) = -(\beta/((2(2\alpha - 1)))$.
(VI) If X is Rayleigh distribution with parameter $\alpha > 0$, then $\xi J(X) = -\sqrt{\pi/32\alpha}$.

In the following, to learn more about the characteristics of the CREX, we will plot it in some considered distributions in Example 2.2.

Figure 1 gives the graphs of $\xi J(X)$ for finite range with $b = 0.5$ and power distributions. Note that $\xi J(X)$ for finite range distribution is increasing and for power distribution is decreasing.

Figure 2 provides the graphs of $\xi J(X)$ for Rayleigh and Pareto distributions with $\alpha = 2$. Note that $\xi J(X)$ for Rayleigh distribution is increasing and for Pareto distribution is decreasing.

REMARK 2.3: *CREX can be used to compare the uncertainties of lifetimes of two systems. Let X_1 and X_2 be two such outcomes from two independently conducted experiments under identical conditions. Then $X_1 - X_2$ measures the uncertainty in rv X with pdf f and sf $\bar{F}$.*

https://doi.org/10.1017/S0269964819000196 Published online by Cambridge University Press

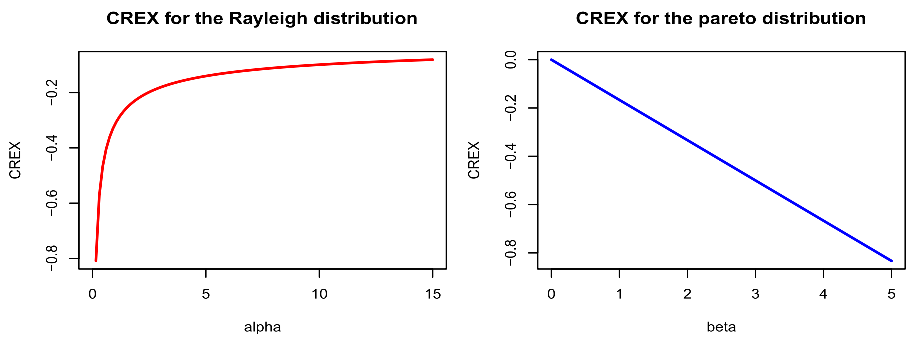

FIGURE 2. Graph of $\xi J(X)$ for $0 \leq \alpha \leq 15$ (left panel) and $0 \leq \beta \leq 5$, $\alpha = 2$ (right panel).

Since X_1 and X_2 are independent rvs, the sf of $X_1 - X_2$ is given by

$$\bar{W}(x) = \int_{-\infty}^{+\infty} f(x)\bar{F}(u+x)dx.$$

This implies the probability of $X_1 = X_2$ approximately equals to

$$\bar{W}(0) = \frac{1}{2} = -\frac{\xi J(X)}{\int_0^{+\infty} \bar{F}(x)dx}.$$

Now, let two rvs X and Y be lifetimes of two systems. If the CREX of rv X is less than Y, that is, $\xi J(X) \leq \xi J(Y)$, we can say that, X has less uncertainty than Y. This compares the uncertainties of two rvs also proposed for the extropy (see Qiu et al. [16]).

The following theorem gives the sufficient condition for CREX to be finite.

THEOREM 2.4: *Let X be a non-negative rv. If for some $p > (1/2)$, $E(X^p) < +\infty$ then $\xi J(X) \in (-\infty.0]$.*

PROOF: Using (2) it is enough to show $-\infty < -\int_0^{+\infty} \bar{F}^2(x)dx$, that is equivalent to show $\int_0^{+\infty} \bar{F}^2(x)dx < +\infty$. We can obtain

$$\begin{aligned}\int_0^{+\infty} \bar{F}^2(x)dx &= \int_0^1 \bar{F}^2(x)dx + \int_1^{+\infty} \bar{F}^2(x)dx \\ &\leq 1 + \int_1^{+\infty} \bar{F}^2(x)dx \\ &\leq 1 + \int_1^{+\infty} \left[\frac{E(X^p)}{x^p}\right]^2 dx \\ &= 1 + E^2(X^p)\int_1^{+\infty} \frac{1}{x^{2p}}dx,\end{aligned}$$

where in the third relation, we use Markov's inequality. The last integral is finite if $p > (1/2)$. Thus, the result follows. ∎

Remark 2.5: *We know that the variance of rv X is defined as*

$$\sigma^2 = Var(X) = E(X^2) - E^2(X),$$

where σ is the standard deviation. If σ^2 exists ($\sigma^2 < +\infty$), then, $E(X^2) < +\infty$ and from Theorem 2.4, $\xi J(X) \in (-\infty, 0]$. Therefore, existence of variance is the sufficient condition for the convergence of CREX.

In the following proposition, we discuss the effect of linear transformations on the CREX.

Proposition 2.6: *Let X be a non-negative rv. If $Y = aX + b$, with $a > 0$ and $b \geq 0$, then $\xi J(Y) = a\xi J(X)$.*

Proof: The result follows by noting that $\bar{F}_{aX+b}(x) = \bar{F}_X((x-b)/a), x \geq 0$ and using (2). ■

Theorem 2.7: *(Weak convergence) Let X_n be a sequence of N-dimensional random vectors converging in distribution to a random vector X. If all the X_n are bounded in L^p for some $p > (N/2)$, then*

$$\lim_{n\to+\infty} \xi J(X_n) = \xi J(X).$$

Proof: Because X_n converges to X in distribution, we get

$$\lim_{n\to+\infty} \bar{F}^2_{|X_n|}(x) = \bar{F}^2_{|X|}(x), \quad x \in R^N_+.$$

On the other hand, we can obtain

$$\begin{aligned} \bar{F}^2_{|X_n|}(x) &\leq \prod_{i=1}^{N} \bar{F}^{2/N}_{|X_i|}(x_i) \\ &\leq \prod_{i=1}^{N} \left[I_{[0,1]}(x_i) + \frac{1}{x_i^p} I_{[1,\infty]}(x_i) E(|X_{n_i}|^p)\right]^{2/N}, \end{aligned}$$

where first and second relations are obtained by Holder's inequality and Eq. (19) of Rao et al. [20], respectively. Therefore, for $(2p/N) > 1$, $\bar{F}^2_{|X_n|}(x)$ is bounded by an integrable function. Thus, dominated convergence theorem completes the proof. ■

Now, we shall focus on upper and lower bounds for CREX. In the following theorem, we show that the CREX of the sum of two independent rvs is not larger than that of each one unlike the extropy (see Qiu et al., 2018).

Theorem 2.8: *Let X and Y be two non-negative and independent rvs with sfs $\bar{F}$ and $\bar{G}$, respectively. Then*

$$\xi J(X+Y) \geq \max\left\{\xi J(X) - \frac{E(Y)}{2}, \xi J(Y) - \frac{E(X)}{2}\right\}.$$

PROOF: Since X and Y are independent rvs, the sf of $X+Y$ is given by

$$\bar{H}(x)=\int_0^{+\infty}\bar{F}(x-t)dG(t).$$

Using Jensen's inequality, we have

$$\begin{aligned}\bar{H}^2(x)&=\left[\int_0^{+\infty}\bar{F}(x-t)dG(t)\right]^2\\&\leq\int_0^{+\infty}\bar{F}^2(x-t)dG(t).\end{aligned}$$

Integrating both sides of the above inequality with respect to x from 0 to $+\infty$, we obtain

$$\begin{aligned}\xi J(X+Y)&\geq-\frac{1}{2}\int_0^{+\infty}\left(\int_0^{+\infty}\bar{F}^2(x-t)dG(t)\right)dx\\&=-\frac{1}{2}\int_0^{+\infty}dG(t)\left(\int_0^{t}\bar{F}^2(x-t)dx+\int_t^{+\infty}\bar{F}^2(x-t)dx\right)\\&=-\frac{E(Y)}{2}+\xi J(X).\end{aligned}$$

Similarly, we can obtain $\xi J(X+Y)\geq-(E(X)/2)+\xi J(Y)$. This completes the proof. ■

REMARK 2.9: *From (2), the CREX of a rv is always non-positive. Thus, Theorem 2.8 holds that*

$$\xi J(X+Y)\geq\xi J(X)+\xi J(Y)-\left(\frac{E(X)+E(Y)}{2}\right).$$

THEOREM 2.10: *Let X be a non-negative rv with CRE, $\xi(X)$ and CREX, $\xi J(X)$. Then,*
(i)

$$\xi J(X)\geq-\frac{E(X)}{2}.\tag{3}$$

(ii)

$$\xi J(X)\leq\frac{1}{2}\big[\xi(X)-E(X)\big].\tag{4}$$

PROOF: (*i*) Since $\bar{F}^2(x)\leq\bar{F}(x)$, the result follows by recalling definition of CREX. The proof of (*ii*) follows from inequality $-\log x\geq(1-x)$, for $x>0$. ■

REMARK 2.11: (*i*) *For a non-negative rv X, Rao et al. [20] have shown that $\xi(X)\leq(E(X^2))/(2E(X))$. Thus, (3) and (4) rewritten as*

$$\xi J(X)\geq-\frac{E(X^2)}{4\xi(X)},\tag{5}$$

and

$$\xi J(X)\leq\frac{1}{2}\left[\frac{E(X^2)}{2E(X)}-E(X)\right].\tag{6}$$

https://doi.org/10.1017/S0269964819000196 Published online by Cambridge University Press

It should be noted here that lower and upper bounds in (3) and (6) are sharper than the lower bound and the upper bound in (5) and (4), respectively. Since CREX of an exponential variable with mean λ is $-(\lambda/4)$, inequality (6) shows that the exponentially distributed rv with mean

$$-2\left[\frac{E(X^2)}{2E(X)}-E(X)\right]$$

have maximum CREX.

(ii) For the random lifetime X of a system, the residual life $[X-t|X>t]$ is the time elapsing between the failure time X and an inspection time t given that at time t the system has been found working. For all $x \geq 0$ such that $F(x)>0$ the mean residual life is given by

$$m_F(t)=E(X-t|X>t)=\int_t^{+\infty}\frac{\bar{F}(x)}{\bar{F}(t)}dx.$$

Asadi and Zohrevand [3] expressed that $\xi(X)=E(m_F(X))$. Thus, (4) is equivalent to

$$\xi J(X)\leq\frac{1}{2}\big[E(m_F(X))-E(X)\big].$$

To illustrate the above results let us consider the following example.

Example 2.12: *Let X be distributed by uniformly on $(0,a)$, $a>0$. For CREX in this case, Theorem 2.10 gives $\xi J(X)\geq-(a/4)$ and $\xi J(X)\leq-(a/8)$. Based on the results (5) and (6) lower and upper bounds are $\xi J(X)\geq-(a/3)$ and $\xi J(X)\leq-(a/12)$. This means that the lower bound and the upper bound in (3) and (6) are sharper than the lower bound and the upper bound in (5) and (4), respectively.*

Let $X_1,X_2,\ldots,X_n$ be n iid non-negative rvs having sfs $\bar{F}$. If $X_{i:n}$ denotes the ith order statistics in this sample of size n, then the lifetime of a series system is determined by $X_{1:n}$ and the lifetime of a parallel system is determined by $X_{n:n}$ with sfs $\bar{F}_{1:n}$ and $\bar{F}_{n:n}$, respectively. The following proposition provides lower bounds for CREX of series and parallel systems based on the mean lifetime of their components.

Proposition 2.13: *Let $X_1,X_2,\ldots,X_n$ be iid non-negative continuous rvs with common sfs $\bar{F}$. Then,*

(i) $\xi J(X_{n:n})\geq-\frac{n^2E(X)}{2}$,

(ii) $\xi J(X_{n:n})\geq n^2\xi J(X)$,

(iii) $\xi J(X_{1:n})\geq-\frac{E(X)}{2}$,

(iv) $\xi J(X_{1:n})\geq\xi J(X)$,

where $X_{1:n}=min\{X_1,X_2,\ldots,X_n\}$ *and* $X_{n:n}=max\{X_1,X_2,\ldots,X_n\}$.

https://doi.org/10.1017/S0269964819000196 Published online by Cambridge University Press

Proof: From (2), we have

$$\begin{aligned}\xi J(X_{n:n}) &= -\frac{1}{2}\int_0^{+\infty}\left[1-(1-\bar{F}(x))^n\right]^2 dx\\ &\geq -\frac{1}{2}\int_0^{+\infty}\left[1-(1-n\bar{F}(x))\right]^2 dx\\ &= -\frac{n^2}{2}\int_0^{+\infty}\bar{F}^2(x)dx\\ &\geq -\frac{n^2}{2}\int_0^{+\infty}\bar{F}(x)dx = -\frac{n^2 E(X)}{2},\end{aligned}$$

where the first inequality is obtained by using Bernoulli's inequality. Also, the third relation gives, $\xi J(X_{n:n}) \geq n^2 \xi J(X)$.

We know that, $\bar{F}^2_{1:n}(x) = \bar{F}^{2n}(x) \leq \bar{F}^2(x) \leq \bar{F}(x)$. Hence, from (2) the proof for both (iii) and (iv) is quite straightforward. ■

Example 2.14: *Let $X_1, X_2, \ldots, X_n$ be iid non-negative continuous rvs uniformly distributed on $(0,1)$. Here, $E(X) = (1/2)$ and*

$$\xi J(X_{n:n}) = -\frac{n^2}{(2n+1)(n+1)} \geq -\frac{n^2}{4},$$

for $n \geq 1$. Thus, (i) holds. From Example 2.2, $\xi J(X) = -(1/6)$, which confirms (ii). It is clear that also (iii) holds,

$$\xi J(X_{1:n}) = -\frac{1}{4n+2} \geq -\frac{1}{4},\ n \geq 1.$$

On the other hand, $\xi J(X_{1:n}) \geq -(1/6)$, for $n \geq 1$, which confirms (iv).

Following theorem provided an alternative expression of the CREX, which is expressed in terms of stop-loss transforms. The stop-loss transform $H_F(t)$ of a rv X is defined as

$$H_F(t) = E(max\{X-t, 0\}) = \int_t^{+\infty}\bar{F}(x)dx.$$

Theorem 2.15: *Let X be a non-negative absolutely continuous rv with CREX, $\xi J(X)$. Then, we have*

$$\xi J(X) = -\frac{1}{2}[E(X) - E(H_F(X))], \tag{7}$$

Proof: From (2) and by Fubini's theorem, we have

$$\begin{aligned}\xi J(X) &= -\frac{1}{2}\int_0^{+\infty}\bar{F}^2(x)dx = -\frac{1}{2}\int_0^{+\infty}\bar{F}(x)\left(\int_x^{+\infty}f(t)dt\right)dx\\ &= -\frac{1}{2}\int_0^{+\infty}f(t)\left(\int_0^t\bar{F}(x)dx\right)dt.\end{aligned} \tag{8}$$

On the other hand,

$$\int_0^t\bar{F}(x)dx = \int_0^{+\infty}\bar{F}(x)dx - \int_t^{+\infty}\bar{F}(x)dx = E(X) - H_F(t), \tag{9}$$

where last equation is obtained from (2). The proof of (7) then follows from the substitution of (9) in (8). ■

REMARK 2.16: *Using the relation* $m_F(t) = (H_F(t))/(\bar{F}(t))$*, we can obtain another alternative expression of the CREX, which is expressed in terms of mean residual life function:*

$$\xi J(X) = -\frac{1}{2}\left[E(X) - E(\bar{F}(X)m_F(X))\right]. \tag{10}$$

EXAMPLE 2.17: *Let* X *have an uniform distribution in* $[0, b]$*. We have* $H_F(x) = ((b-x)^2)/(2b)$ *and* $m_F(x) = ((b-x)/2)$*. Hence,*

$$\xi J(X) = -\frac{1}{2}\left[E(X) - E\left(\frac{(b-X)^2}{2b}\right)\right] = -\frac{b}{6}.$$

In the following, we give a relationship between CREX and extropy.

Let X be a rv with pdf f and let $w(x)$ be a non-negative real function such that $0 < E(w(X)) < \infty$. Then a rv Y is said to have the weighted distribution associated to X and $w(x)$ if its sf is given by

$$\bar{F}^w(t) = \frac{E[w(X)|X > t]}{E[w(X)]}\bar{F}(t).$$

Rao [18,19] presented a unified theory of weighted distributions, identifying various situations which can be modeled by using them. The equilibrium rv Y corresponding to a renewal process with lifetime X is a rv with sf given by

$$\bar{F}^e(t) = \frac{1}{E(X)}\int_t^{+\infty} \bar{F}(x)dx,$$

which is a weighted rv obtained from X and $w(x) = (\bar{F}(x))/(f(x))$. On the other hand,

$$E(H_F(X)) = E(X)E(\bar{F}^e(X)).$$

Hence, by replacing above equation in (7), we have

$$\xi J(X) = -\frac{E(X)}{2}\left[1 - E(\bar{F}^e(X))\right].$$

For the equilibrium rv Y corresponding to rv X we obtain the following result. The proof is omitted.

PROPOSITION 2.18: *Let* X *be a non-negative continuous rv and let* Y *be the equilibrium rv corresponding to* X*. Then,*

$$J(Y) = \frac{\xi J(X)}{E^2(X)}.$$

3. PROPORTIONAL HAZARD MODEL AND GINI INDEX

The well-known proportional hazard model is described by the following relation between survival functions of random life times as

$$\bar{F}_{\theta^*}(x) = \left[\bar{F}(x)\right]^{\theta}, \quad x \in R, \quad \theta > 0, \tag{11}$$

where $\bar{F}_{\theta^*}$ is sf of non-negative rv X_{θ^*}. The mean value of X_{θ^*}, $\theta \in (0, 1]$, arises in the constructions of actuarial risk measures (see Subsection 6.1). Gupta and Gupta [10] studied this model from a reliability point of view and discussed the monotonicity of failure rates. It is easy to see that

$$\xi J(X) = -\frac{E(X_{\theta^*})}{2},$$

$\theta = 2$. Now, consider the following result to compare $\xi J(X)$, $\xi J(X_{\theta^*})$ and $\xi J(\theta X)$.

PROPOSITION 3.1: *There holds*

$$\xi J(X_{\theta^*}) \geq \xi J(X) \geq \xi J(\theta X) \quad \text{if} \quad \theta \geq 1,$$

the inequality being reversed if $0 < \theta \leq 1$.

PROOF: From Definition 2.1 and (11), for $\theta > 1$ $(0 < \theta \leq 1)$ we obtain

$$\begin{aligned} \xi J(X_{\theta^*}) &= -\frac{1}{2}\int_0^{+\infty} \bar{F}^{2\theta}(x)dx \\ &\geq (\leq) -\frac{1}{2}\int_0^{+\infty} \bar{F}^2(x)dx = \xi J(X). \end{aligned} \tag{12}$$

On the other hand, we have

$$\xi J(\theta X) = -\frac{1}{2}\int_0^{+\infty} \bar{F}^2\left(\frac{x}{\theta}\right)dx = \theta \xi J(X) \leq (\geq) \xi J(X).$$

■

The following result immediately follows from Proposition 3.1 and by recalling that the right-hand side of (11) is sf of the minimum iid rvs $(X_{1:n})$ when the power is integer.

COROLLARY 3.2: *Let* $X_1, X_2, \ldots, X_n$ *be iid rvs, with* n *a positive integer. Then,*

$$\xi J(nX_1) \leq \xi J(X_{1:n}),$$

where $X_{1:n} = min\{X_1, X_2, \ldots, X_n\}$.

The Gini index or Gini coefficient is a statistical measure of distribution. It is often used as a gauge of economic inequality, measuring income distribution or, less commonly, wealth distribution among a population. The Gini index is given by

$$gini(X) = \frac{E(|X - Y|)}{E(X + Y)} = \frac{E(|X - Y|)}{2E(X)} = 1 - \frac{\int_0^{+\infty} \bar{F}^2(x)dx}{E(X)},$$

where X and Y are independent rvs and have the same distribution as X. See Wang [26] for more details. Using above equation, the CREX can be expressed as Gini index:

$$\xi J(X) = \frac{E(X)}{2}\left[gini(X) - 1\right]. \tag{13}$$

EXAMPLE 3.3:

(I) For the Gamma distribution with mean $\alpha\beta$, $\alpha > 0, \beta > 0$ *we have*

$$gini(X) = \frac{\Gamma((2\alpha+1)/2)}{\alpha\Gamma(\alpha)\sqrt{\pi}}.$$

Thus, (13) *gives*

$$\xi J(X) = \frac{\beta}{2}\left[\frac{\Gamma((2\alpha+1)/2)}{\Gamma(\alpha)\sqrt{\pi}} - 1\right].$$

(II) For the log-normal distribution with mean $\exp\{\mu + (\sigma^2/2)\}$, $\mu \in R, \sigma > 0$ *we have* $gini(X) = erf(\sigma/2)$ *where* erf *is the error function. Therefore, the CREX for this case is*

$$\xi J(X) = \frac{\exp\{\mu + (\sigma^2/2)\}}{2}[erf(\sigma/2) - 1].$$

From (13) and Proposition 3.1, we have the following result. The proof is omitted.

PROPOSITION 3.4: *There holds*

$$gini(X_{\theta^*}) \geq gini(X) \geq gini(\theta X) \quad if \quad \theta \geq 1,$$

the inequality being reversed if $0 < \theta \leq 1$.

REMARK 3.5: *Schezhtman and Yitzhaki [22] introduced a measure of association between* X *and* Y *based on Gini's mean difference (GMD). The GMD associated to rv* X *is defined as*

$$GMD = E(|X - Y|) = 2\left[E(X) - \int_0^{+\infty} \bar{F}^2(x)dx\right], \tag{14}$$

where X *and* Y *are independent rvs distributed as* X. *It is argued by Yitzhaki [29] that the GMD, as a measure of variability, shares many properties of the variance of* X *and is more informative than the variance for the distributions that are far from normality. An interesting observation regarding the variance of* X *and the GMD is that both of them can be written as special cases of covariance. In fact, one can write* $Var(X) = Cov(X, X)$, *while* $GMD = 4Cov(X, F(X))$ *(see Yitzhaki and Schechtman, [30]). By Definition 2.1 and* (14), *the CREX can be expressed as GMD:*

$$\xi J(X) = \frac{GMD - 2E(X)}{4} = Cov(X, F(X)) - \frac{E(X)}{2}.$$

4. STOCHASTIC ORDERS

In this section, we provide some results on the CREX ordering of rvs. We need the following definition in which X and Y denote random variables with cdfs F and G, pdfs f and g, and sfs $\bar{F}$ and $\bar{G}$, respectively.

DEFINITION 4.1 ([23]): *X is said to be smaller than Y*

1) in the likelihood ratio order, denoted by $X \leq^{lr} Y$, *if* $(f(x)/g(x))$ *is decreasing in* x,

https://doi.org/10.1017/S0269964819000196 Published online by Cambridge University Press

2) *in the hazard rate order, denoted by* $X \leq^{hr} Y$*, if* $\lambda_F(x) \geq \lambda_G(x)$ *for all* x*, where* $\lambda_F(x) = (f(x)/\bar{F}(x))$ *is the hazard rate function,*
3) *in the usual stochastic order, denoted by* $X \leq^{st} Y$*, if* $\bar{F}(x) \leq \bar{G}(x)$ *for all* x*,*
4) *in the dispersive order, denoted by* $X \leq^{disp} Y$*, if* $G^{-1}(F(x)) - x$ *is increasing in* $x \geq 0$*,*
5) *in the increasing concave (convex) order, denoted by* $X \leq^{icv} Y (\leq^{icx})$*, if* $E[\phi(X)] \leq E[\phi(Y)]$ *for all*

increasing concave (convex) functions ϕ *such that the expectations exist of X.*

The connections between the earlier mentioned stochastic orders are described in the following diagram (see Shaked and shanthikumar, 2007)

$$X \leq^{lr} Y \Rightarrow X \leq^{hr} Y \Rightarrow X \leq^{st} Y \Rightarrow X \leq^{icx} Y,$$
$$X \leq^{disp} Y \Rightarrow X \leq^{st} Y \ (\geq^{st}) \ \text{ if } \ l_X = l_Y > -\infty \ (u_X = u_Y < +\infty),$$

where l_X, u_X, l_Y, and u_Y are endpoints of supports of X and Y, respectively. The following theorem describes the relationship between CREX and increasing concave ordering.

THEOREM 4.2: *Let* X *and* Y *be two non-negative absolutely continuos rvs with sfs* $\bar{F}$ *and* $\bar{G}$*, respectively. If* $X \leq^{icv} Y$ *then* $\xi J(X) \geq \xi J(Y)$*.*

PROOF: We know that $\int_0^t \bar{F}(x)dx$ is an increasing concave function. Hence, the proof follows from (8) and recalling the definition of increasing concave order. ■

REMARK 4.3: *The increasing concave ordering is the corresponding ordering for returns instead of losses. This is also known as second order stochastic dominance (SSD), especially in the economic literature. If* $X \leq^{icv} Y$ *holds, where* X *and* Y *are risky returns, then any risk averse decision maker prefers* Y *to* X *(see Rothschild and Stiglitz [21]). Note that* $X \leq^{icv} Y$ *holds, if and only if* $-X \geq^{icx} -Y$*. The increasing convex ordering is also known as stop-loss order in actuarial sciences. The reason is that* $X \leq^{icx} Y$ *holds if and only if* $H_F(t) \leq H_G(t)$*.*

The following theorem describes the relationship between CREX and usual stochastic ordering. The proof is omitted.

THEOREM 4.4: *Let* X *and* Y *be two non-negative absolutely continuos rvs with sfs* $\bar{F}$ *and* $\bar{G}$*, respectively. If* $X \leq^{st} Y$ *then* $\xi J(X) \geq \xi J(Y)$*.*

EXAMPLE 4.5: *Let* X *and* Y *have cdfs* $F(x) = (x/b)$*,* $0 \leq x \leq b$ *and* $G(x) = (x/c)$*,* $0 \leq x \leq c$*, respectively. It is easy to see that* $X \leq^{st} Y$*, for* $c \geq b$*. But, using Example 2.2, we obtain*

$$\xi J(X) = -\frac{b}{6} \geq -\frac{c}{6} = \xi J(Y).$$

REMARK 4.6: *For some families of distributions such as exponential and Parto,* $\xi J(X)$ *can be easily computed in closed form and thus, the ordering can be directly obtained. But the ordering for other distributions can be obtained by application of the Theorem 4.4 (ordering parametric families) and using dispersion, likelihood ratio, and usual stochastic ordering. For example, it can be easily shown that, if* X *has a gamma distribution with shape parameter* θ*,* *then for $\theta_0 < \theta_1$, $X_{\theta_0} \le^{lr} X_{\theta_1}$ and thus $X_{\theta_0} \le^{st} X_{\theta_1}$. So we have $\xi J(X_{\theta_0}) \ge \xi J(X_{\theta_1})$. If X has a Weibull distribution with shape parameter γ, then for $\gamma_0 < \gamma_1$, $X_{\gamma_0} \ge^{disp} X_{\gamma_1}$, and thus $X_{\gamma_0} \ge^{st} X_{\gamma_1}$. Therefore, $\xi J(X_{\gamma_0}) \le \xi J(X_{\gamma_1})$.*

Theorem 4.4 can be used in order statistics and record values as following corollary. For comprehensive discussion on various concepts of order statistics and record values see Arnold et al. [2] and David and Nagaraja [7].

COROLLARY 4.7: (i) *Suppose that $X_{i:n}$ and $Y_{i:n}$ denote the ith order statistic in samples of size n, $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ respectively. If $X \le^{st} Y$, then $X_{i:n} \le^{st} Y_{i:n}$ and we have $\xi J(X_{i:n}) \ge \xi J(Y_{i:n})$, $i = 1, 2, \ldots, n$.*

(ii) *Suppose that U_n and V_n denote the nth record of two sequences of random variables $\{X_n, n \ge 1\}$ and $\{Y_n, n \ge 1\}$, respectively. If $X \le^{st} Y$, then $U_n \le^{st} V_n$ and we have $\xi J(U_n) \ge \xi J(V_n)$.*

In the following theorem, we show that CREX can be a superadditive functional.

THEOREM 4.8: *Let X and Y be two independent non-negative rvs with right-end support points $u_X = u_Y < +\infty$. If X and Y have log-concave densities, then*

(i) $\xi J(X + Y) \ge \max\{\xi J(X), \xi J(Y)\}$.

(ii) $\xi J(X + Y) \ge \xi J(X) + \xi J(Y)$.

PROOF: Let X have a log-concave density. From Theorem 3.B.7 of Shaked and Shanthikumar [23], one can conclude that $X \le^{disp} X + Y$ for any rv Y independent of X. Since $u_X = u_Y < +\infty$, we have, $X \ge^{st} X + Y$. Hence, Theorem 4.4 implies that $\xi J(X + Y) \ge \xi J(X)$. Similar result also holds when Y has a log-concave density i.e. $\xi J(X + Y) \ge \xi J(Y)$. This completes the proof of part (i). Also, we can prove part (ii), noting that the CREX of a rv is always non-positive. ■

5. ESTIMATION

In this section we consider the problem of estimating the CREX by means of the empirical CREX. In this regards, we use two different empirical estimators of the cdf to estimate CREX.

DEFINITION 5.1: *Let $X_1, \ldots, X_n$ be a random sample drawn from a population having cdf F. From (1) we define the empirical CREX as*

$$\xi J(F_n) = -\frac{1}{2} \int_0^{\infty} (1 - F_n(x))^2 dx.$$

https://doi.org/10.1017/S0269964819000196 Published online by Cambridge University Press

where F_n is an empirical estimator of F. Hence, denoting by $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ the order statistic of random sample, we get

$$\xi J(F_n) = -\frac{1}{2} \sum_{j=1}^{n-1} \int_{X_{(j)}}^{X_{(j+1)}} (1 - F_n(x))^2 dx, \tag{15}$$

which is the empirical CREX and converges to CREX of X:

$$\xi J(F_n) \to \xi J(X) \quad a.s. \quad as \quad n \to \infty.$$

The first estimator ($\xi J_1(F_n)$) can be obtained by replacing empirical cdf in (15) as

$$\xi J_1(F_n) = -\frac{1}{2} \sum_{j=1}^{n-1} \left(X_{(j+1)} - X_{(j)}\right) \left(1 - \frac{j}{n}\right)^2, \tag{16}$$

where for $X_{(j)} \leq x < X_{(j+1)}$

$$F_n(x) = \frac{j}{n}, \quad j = 1, 2, \ldots, n-1.$$

Since smoothed estimators have a better performance compared to non-smoothed estimators, we make another estimator based on kernel-smoothed estimator of the cdf.

In a similar way, the second estimator ($\xi J_2(F_h)$) can be achieved by replacing empirical kernel-smoothed estimator in (15) as

$$\begin{aligned} \xi J_2(F_h) &= -\frac{1}{2} \sum_{j=1}^{n-1} \int_{X_{(j)}}^{X_{(j+1)}} (1 - F_h(x_j))^2 dx, \\ &= -\frac{1}{2} \sum_{j=1}^{n-1} \left(X_{(j+1)} - X_{(j)}\right) (1 - F_h(x_j))^2, \end{aligned} \tag{17}$$

where $F_h(\cdot)$ is defined by Nadaraya (1964) as

$$F_h(x) = \frac{1}{n} \sum_{i=1}^{n} L\left(\frac{x - X_i}{h}\right),$$

where L is the cdf of a positive kernel K, i.e $L(x) = \int_{-\infty}^{x} K(t)dt$ and h is a bandwidth parameter. Meanwhile, we use the normal kernel which is optimal in a mean square error sense, though the loss of efficiency is small and due to its convenient mathematical properties is often used, which means $K(x) = \phi(x)$, where ϕ is the standard normal density function.

It is important to point out, to estimate the bandwidth (h) we use Sarda (1993) method which is considered the following selection method that minimizes the leave-one-out criterion

$$CV(h) = \frac{1}{n} \sum_{i=1}^{n} \{F_{h,-i}(X_i) - F_n(X_i)\}^2,$$

where $F_{h,-i}(X_i)$ is leave-one-out version of the kernel-smoothed estimator of cdf which is defined as

$$F_{h,-i}(x) = \frac{1}{n-1} \sum_{j \neq i} L\left(\frac{x - X_j}{h}\right).$$

TABLE 1. Mean and variance of the empirical CREX

n	$E[\xi J_1(F_n)]$	$Var[\xi J_1(F_n)]$	$E[\xi J_2(F_h)]$	$Var[\xi J_2(F_h)]$
5	−0.200	0.012	−0.289	0.022
10	−0.225	0.007	−0.291	0.009
20	−0.237	0.004	−0.296	0.004
30	−0.241	0.003	−0.299	0.003
40	−0.242	0.002	−0.308	0.002
50	−0.245	0.002	−0.310	0.002
100	−0.248	0.001	−0.313	0.001

EXAMPLE 5.2: *Let $X_1, X_2, \ldots, X_n$ be a random sample of exponentially distributed rvs with parameter θ. Recalling Pyke (1965), the sample spacings are independent, with $X_{(j)} - X_{(j-1)}$ exponentially distributed with parameter $\theta(n-j)$. Hence, from (16) we have:*

$$E[\xi J_1(F_n)] = -\frac{1}{2n\theta}\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right), \quad Var[\xi J_1(F_n)] = \frac{1}{4n^2\theta^2}\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)^2,$$

similarly,

$$E[\xi J_2(F_h)] = -\frac{1}{2\theta}\sum_{j=1}^{n-1}\frac{(1-F_h(x_j))^2}{n-j}, \quad Var[\xi J_2(F_h)] = \frac{1}{4\theta^2}\sum_{j=1}^{n-1}\frac{(1-F_h(x_j))^4}{(n-j)^2}.$$

Table 1 shows mean and variance of the empirical CREX for random samples from exponential distribution with mean 1 and some choices of n. Based on the results of Table 1, by increasing sample size the values of mean and variance of the proposed estimators are decreased which results from dependence of the mean and variance of the empirical estimators to the sample size.

The following are two examples that clarify the effectiveness of the empirical CREX measure to perform some estimations.

EXAMPLE 5.3: *This data set includes an active repair time (in hours) for an airborne communication transceiver reported by [4], which was originally given by Chhikara and Folks [6]. The actual observations are listed below.*

0.2, 0.3, 0.5, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.7, 0.8, 0.8, 1.0, 1.0, 1.0, 1.0, 1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2.0, 2.0, 2.2, 2.5, 2.7, 3.0, 3.0, 3.3, 3.3, 4.0, 4.0, 4.5, 4.7, 5.0, 5.4, 5.4, 7.0, 7.5, 8.8, 9.0, 10.3, 22.0, 24.5.

To check the validity of using exponential distribution for fitting to repair time data, Kolmogorov–Smirnov (K-S) test is applied. The K-S statistic of the distance between the fitted and the empirical distribution functions (based on the parameter $\theta = 0.2773$) is 0.1597 and the corresponding p-value is 0.1914. Therefore, it is reasonable to use the exponential distribution to fit the data which can be seen in left panel of Figure 3 too.

The value of CREX based on the exponential distribution is equal to −0.9016 but empirical estimator ($\xi J_1(F_n)$) and kernel-smoothed estimator ($\xi J_2(F_h)$) are −0.6685 and −0.7441 respectively which $\xi J_2(F_h)$ is more closer than $\xi J_1(F_n)$ to the theoretical value.

The right panel in Figure 3 shows that by increasing sample size, the empirical estimators become closer to the theoretical value which is due to dependence of them on n. Also, it

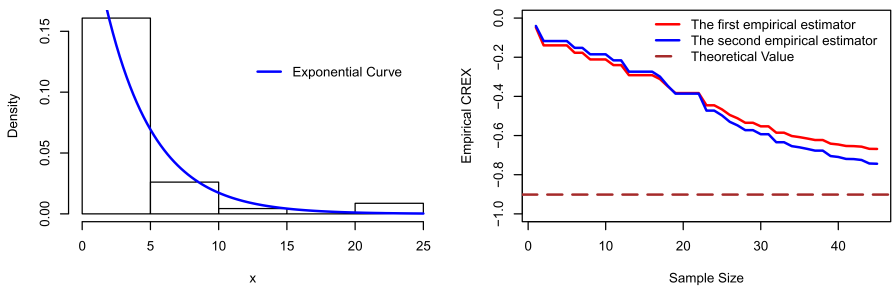

FIGURE **3.** Histogram of the active repair times data (left panel) and empirical estimators (right panel).

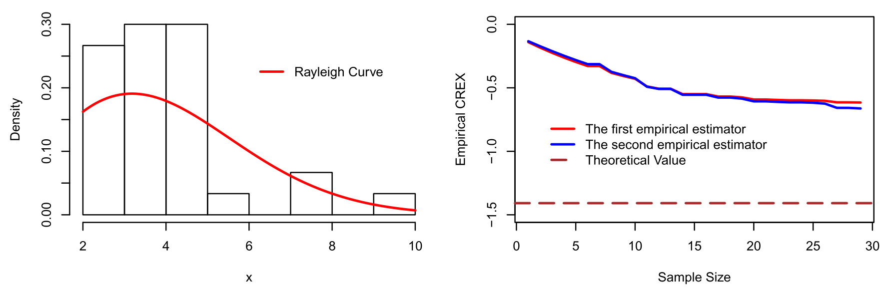

FIGURE **4.** Histogram of the average wind speed data (left panel) and empirical estimators (right panel).

can be seen that the second empirical estimator approaches the amount of the theory value more quickly than the first empirical estimator. Therefore, we conclude that the proposed kernel-smoothed estimator is more accurate than empirical estimator.

EXAMPLE 5.4: *In this example, we consider one average wind speed data analysis reported in Best et al. [5]. The following data represent 30 average daily wind speeds (in km/h) for the month of November 2007 recorded at Elanora Heights, a northeastern suburb of Sydney, Australia:*

2.7, 3.2, 2.1, 4.8, 7.6, 4.7, 4.2, 4.0, 2.9, 2.9, 4.6, 4.8, 4.3, 4.6, 3.7, 2.4, 4.9, 4.0, 7.7, 10.0, 5.2, 2.6, 4.2, 3.6, 2.5, 3.3, 3.1, 3.7, 2.8, 4.0.

The wind speed data were analyzed initially by Best *et al.* [5] and Alizadeh *et al.* [1], who fitted the Rayleigh distribution successfully. As we checked, since the probability value of the K-S test is equal to 0.1262 so that the Rayleigh distribution has a good fit to these data which can be seen in the left panel of Figure 4 too.

The value of CREX based on the Rayleigh distribution is equal to -1.408 but empirical estimator $(\xi J_1(F_n))$ and kernel-smoothed estimator $(\xi J_2(F_h))$ are -0.6158 and -0.6627 respectively which $\xi J_2(F_h)$ is more closer than $\xi J_1(F_n)$ to the theoretical value.

It can be seen in the right panel of Figure 4 that by increasing sample size the empirical estimators become closer to the theoretical value. Also, as can be seen that the second empirical estimator approaches the amount of the theory value more quickly than the first empirical estimator. Therefore, we conclude once again that the $\xi J_2(F_h)$ estimator is more accurate than the $\xi J_1(F_n)$ estimator.

In the following, we perform a simulation study to compare the performance of the proposed empirical estimators based on simulated data from active repair time and average

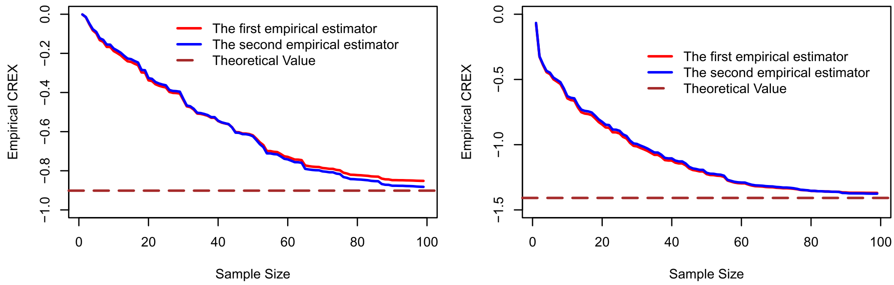

FIGURE 5. Empirical estimators for exponential distribution (left panel) and the Rayleigh distribution (right panel).

wind speed data. Figure 5 provides a convenient visual summary of comparing proposed estimators which one can easily infer that the proposed estimators are affected by sample size and generally the second empirical estimator is more accurate than the first estimator. In addition, it is evident that for sample size more than 80 the proposed estimators have good accuracy to estimate the theoretical values.

6. APPLICATIONS

In this section, we derived two applications of the CREX to risk measure and independence.

6.1. Risk Measures

Yang [27] studied CRE as an alternative risk measure for heavy-tailed distribution when variance does not exist. Also, Ramsay [17] showed that standard deviation $\sigma(X)$ is not an appropriate tool to measure large insurance risks with large tailed skewed distributions. Hence, Wang [26] proposed a measure of the right-tail risk, namely, the right-tail deviation given by

$$D(X) = \int_0^{+\infty} \sqrt{\bar{F}(x)}dx - E(X).$$

$|\xi J(X)|$ preserves some basic properties of $D(X)$. From Proposition 2.6, we obtain $|\xi J(aX + b)|$
$= a|\xi J(X)|$, $a > 0, b > 0$. Furthermore, we get a monotonicity property for $|\xi J(X)|$ under the hypothesis of usual stochastic order (see Section 4). Also, from Theorem 4.8, subadditivity property of $|\xi J(X)|$ can be hold. Therefore, we can consider $|\xi J(X)|$ as a risk measure. It is to be noted that when $\sigma(X)$ of heavy-tailed distribution does not exist, you can use CREX to measure the risk, such as Pareto distribution for $1 < \alpha < 2$ (see Example 2.2). To compare $|\xi J(X)|$, $D(X)$ and $\sigma(X)$ we present the following examples.

EXAMPLE 6.1: *Again consider Example 2.2. Then*

(I) For a uniform distribution we get

$$|\xi J(X)| = D(X) = \frac{b}{6} < \frac{b}{2\sqrt{3}} = \sigma(X).$$

https://doi.org/10.1017/S0269964819000196 Published online by Cambridge University Press

(II) For the exponential distribution we have

$$|\xi J(X)| = \frac{1}{4\lambda} < \frac{1}{\lambda} = D(X) = \sigma(X).$$

(III) For pareto distribution with parameters $\alpha > 2$ *and* $\beta > 0$ *we can see that*

$$|\xi J(X)| = \frac{\beta}{2(2\alpha - 1)} < \frac{\alpha\beta}{(\alpha - 1)(\alpha - 2)} = D(X).$$

Moreover, it follows that

$$|\xi J(X)| < \frac{\beta\sqrt{\alpha}}{(\alpha - 1)\sqrt{\alpha - 2}} = \sigma(X).$$

In this case for $1 < \alpha < 2$*,* $D(X)$ *and* $\sigma(X)$ *do not exist and we can use CREX to measure the risk.*

EXAMPLE 6.2: *Let* X *has a Weibull distribution with parameters* $\lambda > 0, \gamma > 0$ *and sf* $\bar{F}(x) = \exp\{-\lambda x^{\gamma}\}$*,* $x \geq 0$*. After some standard calculations, we obtain*

$$|\xi J(X)| = \frac{\Gamma(1/\gamma)}{4\gamma\lambda^{1/\gamma}} > (<) \frac{\Gamma(1/\gamma)[2^{1/\gamma} - 1]}{\gamma\lambda^{1/\gamma}}, \quad \textit{for } \gamma > 3.10628 \ (0 < \gamma < 3.10628),$$

where $\Gamma(.)$ *is the complete gamma function. Moreover,*

$$|\xi J(X)| > (<) \frac{\left[\Gamma(1 + (2/\gamma)) - \Gamma^2(1 + (1/\gamma))\right]^{1/2}}{\lambda}, \quad \textit{for } \gamma > \gamma_0 \ (0 < \gamma < \gamma_0),$$

where γ_0 *is near a point* 2.78*.*

6.2. Independence of Two rvs

In this section, a new measure of independence using CREX was derived, which can be used to measure the independence between two rvs.

Suppose we have n samples of two continuous rvs (X, Y), $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, we aim to infer from the sample data if X and Y are independent rvs. The commonly used method is to calculate the correlation between X and Y, then perform the hypothesis testing of $H_0 : X$ and Y are independent, H_1: Not H_0. However, we know that $corr(X, Y) = 0$ does not imply X, Y are independent except for the normal distribution for which the independence and uncorrelated are equivalent. Here, we derived a new measure of independence using CREX, which can measure the true independence rather than the correlation between two rvs. First, we defined the conditional CREX as

$$\xi J(X|Y) = -\frac{1}{2}\int_0^{+\infty} \bar{F}^2_{X|Y}(x|y)\, dx, \quad y > 0, \tag{18}$$

where $\bar{F}_{X|Y}(x|y) = P(X > x|Y > y)$. Note that when X and Y are independent, then from (18) it is easy to obtain

$$\xi J(X|Y) = \xi J(X).$$

Definition 6.3: *Let X, Y be two real valued rvs, define*

$$r_{\xi J}(X,Y) = 1 - \frac{\xi J(Y|X)}{\xi J(Y)}.$$

The simulation results show that CREX measure is asymptotically better than the correlation measure, especially when n is sufficiently large.

Proposition 6.4: *For any rv X and Y,*

$$r_{\xi J}(X,Y) = \begin{cases} 0 & if\, X\, and\, Y\, are\, independent, \\ 1 & if\, Y\, is\, a\, function\, of\, X. \end{cases}$$

Proof: It follows from the property of conditional CREX. ■

Corollary 6.5: *Let $\{X_n\}_{n\in N}$ be rvs converging to X in distribution, and $\{Y_n\}_{n\in N}$ be rvs converging to Y in distribution. Then*

$$\lim_{n\to\infty} r_{\xi J_n}(X_n,Y_n) = \begin{cases} 0 & if\, X_n\, and\, Y_n\, are\, independent, \\ 1 & if\, Y_n\, is\, a\, function\, of\, X_n. \end{cases}$$

Proof: It follows from the weak convergence theorem and the above proposition. ■

We perform two experiments to test the independence between two rvs using the measure $r_{\xi J_n}$ and the correlation method. Since all random numbers can be obtained from the uniform random numbers, we focus on the the uniform distributions.

Experiment 1. Generate samples from uniform distribution in $(0,1)$ as follows:

- step 1, generate 50 samples, $\{x_1^{(1)}, x_2^{(1)}, \ldots, x_{50}^{(1)}\}$ and let $X^{(1)}$ denote this array;
- step 2, generate 100 samples, $\{x_1^{(2)}, x_2^{(2)}, \ldots, x_{100}^{(2)}\}$ and let $X^{(2)}$ denote this array;
- step n, generate 50n samples, $\{x_1^{(n)}, x_2^{(n)}, \ldots, x_{50n}^{(n)}\}$ and let $X^{(n)}$ denote this array.

We generate 200 times and get a sequence of arrays $X^{(1)}, X^{(2)}, \ldots, X^{(200)}$ and the length of $X^{(n)}$ is 50n. Do the same procedure and obtain the other sequence of arrays $Y^{(1)}, Y^{(2)}, \ldots, Y^{(200)}$.

Since we generate two sets of random samples independently, we may assume $X^{(n)}$ and $Y^{(n)}$ are independent. Based on Corollary 3, $r_{\xi J_n}(X^{(n)}, Y^{(n)}) \to 0$ as $n \to \infty$.

The numerical results in Figure 6 show that as $n\to\infty$, $r_{\xi J_n}(X^{(n)}, Y^{(n)})$ and $corr(X_n, Y_n)$ decrease to zero asymptotically.

Experiment 2. Construct a new sequence $\{Z^{(1)}, Z^{(2)}, \ldots, Z^{(200)}\}$ where $Z^{(n)} = sin(100X^{(n)})$, $n = 1, 2, \ldots, 200$, a highly non-linear function of $X^{(n)}$. Recalculate $r_{\xi J_n}$ of $(X^{(n)}, Z^{(n)})$ and $corr(X^{(n)}, Z^{(n)})$. The results in Figure 7 show that $r_{\xi J_n}$ concentrated around 0.44 as n increases, which entails dependence between $X^{(n)}$ and $Z^{(n)}$. But the values of $corr(X_n, Z_n)$ reveal no information.

7. CONCLUSION

In this paper, an alternative measure of uncertainty of a rv was introduced which was called cumulative residual extropy. Proposed measure is based on the cumulative distribution

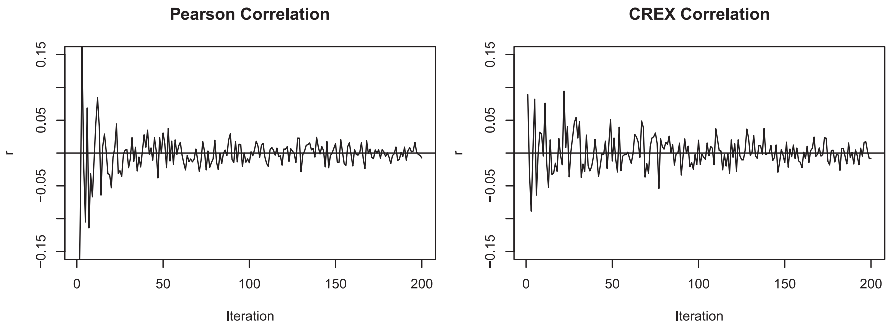

FIGURE 6. Pearson correlation plot (left panel) and CREX correlation plot (right panel).

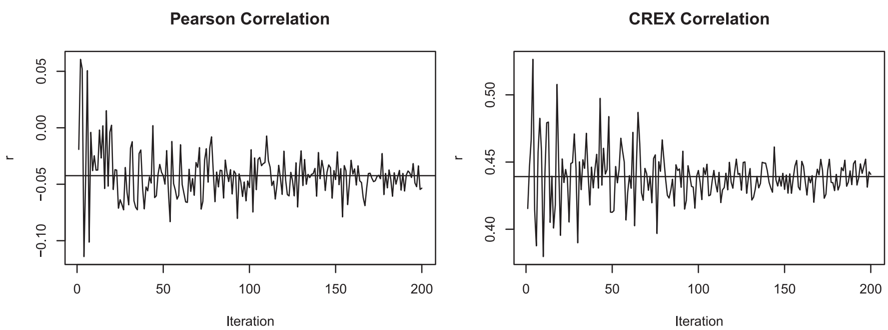

FIGURE 7. Pearson correlation plot (left panel) and CREX correlation plot (right panel).

function F. Upper and lower bounds were determined and inequalities concerning CREX and it shows that the exponentially distributed rv have maximum CREX. Proportional hazard model and Gini index were studied. Also, some results on the CREX ordering of rvs were provided, such as the relationship between CREX and increasing concave ordering and the relationship between CREX and usual stochastic ordering. Then, we considered the problem of estimating the CREX by means of the empirical cumulative extropy by proposing two different empirical estimators of cdf to estimate the extropy. We concluded that the proposed estimators are affected by sample size and generally the second empirical estimator is more accurate than the first estimator. Finally, two applications of the CREX were presented to risk measure and independence problem.

In the future, we will consider the dynamic version of CREX and study some properties of it.

Acknowledgment

The authors would like to thank the Editor and anonymous reviewers for the constructive and valuable suggestions which led to the improvement in an earlier version of this article.

References

1. Alizadeh, R., Alizadeh, H., & Ebrahimi, A. (2014). An entropy test for the rayleigh distribution and power comparison. *Journal of Statistical Computation and Simulation* 84: 151–158.
2. Arnold, B.C., Balakrishnan, B.N., & Nagaraja, H.N. (1992). *A first course in order statistics*. New York: John Wiley and Sons.

3. Asadi, M. & Zohrevand, Y. (2007). On the dynamic cumulative residual entropy. *Journal of Statistical Planning and Inference* 137: 1931–1941.
4. Balakrishanan, N., Leiva, V., Sanhuzea, A., & Cabrera, E. (2009). Mixture inverse gaussian distributions and its transformations, moments and applications. *Statistics* 43: 91–104.
5. Best, D.J., Rayner, J.C.W., & Thas, O. (2010). Easily applied tests of fit for the Rayleigh distribution. *Sankhya B* 72: 254–263.
6. Chhikara, R.S. & Folks, J.L. (1989). *The inverse gaussian distribution: Theory, methodology, and applications*, New York: Springer, Marcel Dekker.
7. David, H.A. & Nagaraja, H.N (2003). *Order statistics*. 3rd ed. New York: Wiley.
8. Furuichi, S. & Mitroi, F.C. (2012). Mathematical inequalities for some divergences. *Physica A,* Elsevier 391: 388–400.
9. Gneiting, T. & Raftery, A.E. (2007). Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association* 102: 359–378.
10. Gupta, R.C. & Gupta, R.D. (2007). Proportional reversed Hazard rate model and its applications. *Journal of Statistical Planning and Inference* 137: 3525–3536.
11. Jaynes, E.T. (2003). *Probability theory: the logic of science*. Cambridge: Cambridge Univ. Press.
12. Lad, F., Sanfilippo, G., & Agro, G. (2015). Extropy: complementary dual of entropy. *Statistical Science* 30: 40–58.
13. Qiu, G. (2017). The extropy of order statistics and record values. *Statistics and Probability Letters,* Elsevier 120: 52–60.
14. Qiu, G. & Jia, K. (2018a). The residual extropy of order statistics. *Statistics and Probability Letters,* Elsevier 133, 15–22.
15. Qiu, G. & Jia, K. (2018b). Extropy estimators with applications in testing uniformity. *Journal of Nonparametric Statistics*, 30, 182–196.
16. Qiu, G., Wang, L., & Wang, X. (2018). On extropy properties of mixed systems. *Probability in the Engineering and Informational Science* 42, 1–16.
17. Ramsay, C.M. (1993). Loading Gross Premiums for Risk Without Using Utility Theory, Transactions of the Society of Actuaries XLV, 305–349.
18. Rao, C.R. (1965). On discrete distributions arising out of methods of ascertainment. *Sankhya A* 27: 31–324.
19. Rao, C.R. (1985). *Weighted distributions arising out of methods of as certainment: what population does a sample represent?*. NewYork: Springer-Verlag, 543–569.
20. Rao, M., Chen, Y., Vemuri, B.C., & Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE Transactions on Information Theory* 50: 1220–1228.
21. Rothschild, M. & Stiglitz, J.E. (1970). Increasing risk: I a definition. *Journal of Economic Theory*, 2: 225–243.
22. Schezhtman, E. & Yitzhaki, S. (1987). A measure of association based on Gini's mean difference. *Communication in Statistics Theory and Methods* 16: 207–231.
23. Shaked, M. & Shanthikumar, J.G. (2007). *Stochastic orders*. New York: Springer Verlag.
24. Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27: 379–423.
25. Vontobel, P.O. (2013). The Bethe permanent of a nonnegative matrix. *IEEE Transactions on Information Theory* 59: 1866–1901.
26. Wang, S. (1998). An actuarial index of the right-tail risk. *North American Actuarial Journal* 2: 88–101.
27. Yang, V. (2012). Study on Cumulative Residual Entropy and Variance as Risk Measure, Fifth International Conference on Business Intelligence and Financial Engineering, published in IEEE, 4 pages.
28. Yang, J., Xia, W., & Hu, T. (2018). Bounds on extropy with variational distance constraint. *Probability in the Engineering and Informational Sciences* 33, 1–19.
29. Yitzhaki, S. (2003). Gini's mean difference: a superior measure of variability for non-normal distributions. *Metron* 61: 285–316.
30. Yitzhaki, S. & Schechtman, E. (2013). *The Gini methodology: a primer on a statistical methodology.* New York: Springer Verlag.