Amit Ghosh; Chanchal Kundu

On generalized conditional cumulative past inaccuracy measure

Applications of Mathematics, Vol. 63 (2018), No. 2, 167–193

Persistent URL: `http://dml.cz/dmlcz/147188`

ON GENERALIZED CONDITIONAL CUMULATIVE PAST INACCURACY MEASURE

Amit Ghosh, Chanchal Kundu, Jais

Received June 26, 2017. First published April 9, 2018.

Abstract. The notion of cumulative past inaccuracy (CPI) measure has recently been proposed in the literature as a generalization of cumulative past entropy (CPE) in univariate as well as bivariate setup. In this paper, we introduce the notion of CPI of order α and study the proposed measure for conditionally specified models of two components failed at different time instants, called generalized conditional CPI (GCCPI). Several properties, including the effect of monotone transformation and bounds of GCCPI are discussed. Furthermore, we characterize some bivariate distributions under the assumption of conditional proportional reversed hazard rate model. Finally, the role of GCCPI in reliability modeling has also been investigated for a real-life problem.

Keywords: cumulative past inaccuracy; marginal and conditional past lifetimes; conditional proportional reversed hazard rate model; usual stochastic order

MSC 2010: 62B10, 94A17, 62N05, 62H05

1. Introduction

Information theory, a unifying theory with profound intersections with probability, statistics, physics, economics, statistical mechanics, computer science, and many other fields, continues to set the stage for the development of communications, data storage and processing, and other information technologies. Information theory has its roots in Shannon's [41] pioneering work on communication which provides a mathematical definition of information dubbed as Shannon entropy. Since its inception, numerous entropy and information indices have been developed in the literature in both the parametric and nonparametric points of view and used extensively in

The financial support (Ref. No. 2/48(4)/2015/NBHM(R.P.)/R&D II/14130) rendered by the NBHM, Department of Atomic Energy, Government of India is acknowledged with thanks by C. Kundu for carrying out this research work.

DOI: 10.21136/AM.2018.0170-17

various disciplines. One important development in this direction is Kerridge [21] inaccuracy measure which can be thought of as a non-parametric generalization of Shannon entropy. It is worth mentioning that the continuous version of the inaccuracy is closely related to Fraser information [12] which was extensively used by Kent [19], [20], Ebrahimi et al. [11] and several others in terms of the "information gain" about a parameter.

Let X and Y be two absolutely continuous nonnegative random variables with cumulative distribution functions (*cdf*'s) F, G and probability distribution functions (*pdf*'s) f, g, respectively. If F is the actual distribution corresponding to the observations and G is the distribution assigned by the experimenter, then the inaccuracy measure of X and Y (also known as cross entropy of Y on X or relative distance between X and Y) is given by

$$K_{X,Y} = -\int_0^\infty f(x) \ln g(x)\,\mathrm{d}x, \tag{1.1}$$

which has wide range of applications in statistical inference, estimation and coding theory. In statistical inference, (1.1) is enormously used to measure the deviation of a distribution of interest from a reference distribution. The inaccuracy measure/cross entropy is proportional to the Kullback-Leibler divergence [23] and it is useful in comparing different probabilistic models when we do not know the actual probability distribution that generated some data. It allows the experimenter to assign a distribution which is an approximation to the true distribution. The assigned distribution will be more accurate to the true distribution as much as closer (1.1) will be to the Shannon entropy. Thus, the difference between the inaccuracy and the entropy is a measure of how accurate a model is.

There have been various attempts for the parametric generalizations of these information measures which make them sensitive to different shapes of probability distributions. It was Renyi [37] who embodied the idea of parametric generalization by introducing entropy of order α, called Renyi's entropy, which has a paramount importance in different areas such as physics, electronics, engineering, ecology, and statistics as a measure of uncertainty and diversity. In this direction, Nath [30] defined inaccuracy measure of order α as

$$K^{\alpha}_{X,Y} = \frac{1}{1-\alpha} \ln \int_0^\infty f(x)(g(x))^{\alpha-1}\,\mathrm{d}x, \tag{1.2}$$

where $0 < \alpha \neq 1$. As $\alpha \to 1$, (1.2) reduces to inaccuracy measure (1.1). Due to presence of an extra parameter, the proposed measure is more flexible than (1.1) and with help of this parameter α, one can make it more or less sensitive to the shape of probability distribution.

It is well known that the distribution function is more regular than the density function. Moreover, in practice what is of interest and/or measurable is the distribution function. To this aim, many researchers have shown a great interest to discover a new/novel measure of information based on the probability distribution function of a random variable rather than its density function. Rao et al. [36] introduced an alternating measure of uncertainty, called cumulative residual entropy (CRE), which is mathematically more rigorous and also overcomes several drawbacks of Shannon entropy. It is reasonable to presume that in many realistic situations uncertainty is not necessarily related to the future but can also refer to the past. Based on this idea Di Crescenzo and Longobardi [8] have studied a dual concept of CRE called cumulative past entropy (CPE) defined as

$$\overline{\varepsilon}(X) = -\int_0^\infty F(x) \ln F(x)\,\mathrm{d}x. \tag{1.3}$$

For more properties, applications and recent developments of CRE and CPE, one may refer to Wang et al. [45], Cahill et al. [6], Di Crescenzo and Longobardi [8]; [9], Navarro et al. [31], Sunoj and Linu [43], Shi et al. [42], Park and Kim [33], Kundu and Nanda [27], Psarrakos and Toomaj [35], and Toomaj et al. [44].

Motivated by the wide applicability of CPE, Kundu et al. [24] introduced the concept of cumulative past inaccuracy (CPI) which is defined as

$$\overline{\mathcal{K}}_{X,Y} = -\int_0^\infty F(x) \ln G(x)\,\mathrm{d}x. \tag{1.4}$$

The basic idea is to replace the density function by the distribution function in Kerridge inaccuracy measure defined in (1.1). Thus, (1.4) can be viewed as a suitable extension of the Kerridge inaccuracy measure to the cumulative distribution function. In analogy with cumulative residual inaccuracy (cf. Kundu et al., [24]), CPI can also be considered as a measure of inaccuracy for past lifetimes. We also recall that the cumulative Kullback-Leibler (CKL) information (cf. Park et al., [34]) of X and Y is defined as

$$\overline{\mathcal{I}}_{X,Y} = \int_0^\infty F(x) \ln \frac{F(x)}{G(x)}\,\mathrm{d}x + E(X) - E(Y), \tag{1.5}$$

where $E(X)$ and $E(Y)$ are the expected values of X and Y, respectively. Note that $\overline{\mathcal{I}}_{X,Y} \geqslant 0$ and the equality holds if and only if $F(x) = G(x)$ almost everywhere. For more details about CKL information, we refer to Di Crescenzo and Longobardi [10]. The measure given in (1.5) for past lifetimes can be considered as an analog of the cumulative residual Kullback-Leibler (CRKL) divergence (cf. Baratpour and Rad [4]). The best model is determined by minimizing the CRKL/CKL between the true distribution and the approximate model. In view of (1.3) and (1.4), (1.5) can alternatively be rewritten as

$$\overline{\mathcal{I}}_{X,Y} = \overline{\mathcal{K}}_{X,Y} - \overline{\varepsilon}(X) + E(X) - E(Y), \tag{1.6}$$

where $\overline{\varepsilon}(X)$ is the CPE of the true distribution. Thus, minimizing the cumulative Kullback-Leibler information to select the best model is equivalent to minimizing the CPI. If $G(x) = F(x)$, the CPI is said to be at a minimum and $\overline{\mathcal{K}}_{X,Y} = \overline{\varepsilon}(X)$. Indeed, closer the value of CPI is to the CPE, the better Y is an approximation of X. CPI can therefore be used to compare approximate models. From two models, the more accurate model will be the one with the lower CPI. Consider the following example to manifest the important role of CPI for measuring quality in models. For a detailed discussion on the role of inaccuracy measure (cross entropy) for comparing models one may refer to Burnham and Anderson [5] and Choe [7].

Example 1.1. Let X be the true distribution that generated some data and Y, Z be two power distributions assigned by the experimenter in order to approximate X. Write the *pdf*'s of X, Y and Z as $f(x) = 1$, $g(x) = 2x$, and $h(x) = 3x^2$ for all $x \in (0,1)$, respectively. Then after simple calculation one can see that the entropy of X is zero whereas the inaccuracy measures $K_{X,Y} = 0.3068$ and $K_{X,Z} = 0.9014$. Thus, Y has a lower inaccuracy measure on X. Therefore, Y is better than Z at approximating the true distribution X. Now, we calculate $\overline{\varepsilon}(X) = 0.25$, $\overline{\varepsilon}(Y) = 0.22$, $\overline{\varepsilon}(Z) = 0.1875$, $\overline{\mathcal{K}}_{X,Y} = 0.5$, and $\overline{\mathcal{K}}_{X,Z} = 0.75$. Thus, the CPI between X and Y is less than the CPI between X and Z. Therefore, on using the notion of CPI it is also verified that the distribution of Y gives better approximation to X than that of Z.

Next, we recall a connection between the CPI and expected inactivity time (EIT), a well-known reliability measure having application in many areas such as reliability theory, survival analysis and actuarial studies. The EIT of a random variable X is defined as $\overline{m}_X(t) = E(t - X \mid X < t)$. In the following proposition we show how CPI is related to the EIT.

Proposition 1.1. *Let X and Y be two absolutely continuous nonnegative random variables with distribution functions $F(\cdot)$ and $G(\cdot)$, respectively, and let $\overline{m}_X(t)$ be the EIT corresponding to the random variable X. If $\overline{\mathcal{K}}_{X,Y} < \infty$, then*

$$\overline{\mathcal{K}}_{X,Y} = E\Big[\frac{\overline{m}_X(Y)F(Y)}{G(Y)}\Big].$$

For further applications and perspectives of CPI along the same line, one may refer to Ghosh and Kundu [13].

In analogy with (1.2), the notion of CPI of order α (CPI(α)) is defined as

$$\overline{\mathcal{K}}^{\alpha}_{X,Y} = \frac{1}{1-\alpha} \ln \int_0^{\infty} F(x)(G(x))^{\alpha-1}\,\mathrm{d}x, \quad 0 < \alpha \neq 1. \tag{1.7}$$

Note that the well known failure entropy of order α (cf. Abbasnejad [1]), also denominated as CPE of order α, can be obtained from (1.7) by taking $G(x) = F(x)$.

Unlike the univariate case, the study of information theoretic measures based on multivariate lifetimes have attracted increasing attention in the recent years. In many practical problems, multivariate lifetime data arise in a variety of observational and experimental studies such as medicine, biology, public health, epidemiology, engineering, economic and demography. In these situations it is important to consider different multivariate models that could be used to model such multivariate lifetime data. For example, the analysis of survival times for twins, siblings or other related individuals. For an encyclopedia on various multivariate models, their properties and applications, one may refer to the book by Kotz et al. [22]. Let $X = (X_1, X_2)$ be a bivariate random vector with support $(0, l) \times (0, l)$ for $l \geqslant 0$. Assume that X_1 and X_2 describe the failure times of two components. Consider the conditional random variables $\widehat{X}_i = (X_i \mid X_1 < t_1, X_2 < t_2)$ and $\widehat{X}^*_{i|j} = (X_i \mid X_i < t_i, X_j = t_j)$ for $i = 1, 2$, $j = 3 - i$, where t_1 and t_2 may not necessarily be identical. Note that in view of the joint past lifetime $[(X_1, X_2) \mid X_1 < t_1, X_2 < t_2]$ when both components failed before inspection, $\widehat{X}_1$ and $\widehat{X}_2$ are realized as marginal past lifetimes while $\widehat{X}^*_{i|j}$, $i = 1, 2$, $j = 3 - i$ as conditional past lifetimes. Here $\widehat{X}_i$ represents the conditional distributions of X_i subject to the condition that failure of the first component had occurred in $(0, t_1)$ and the second failed before t_2. Recently, Ghosh and Kundu [13] and Kundu and Kundu [26] have considered the bivariate extension of (1.4) and CPE of order α, respectively, for marginal and conditional past lifetimes. Motivated by this, in this article we extend the concept of CPI(α) in bivariate setup with focus on marginal and conditional past lifetimes and study their properties useful in reliability modeling. It is worthwhile to mention that the concepts in past time are more appropriate than those truncated from below when the observations are predominantly from left tail. For some recent work on conditionally specified models, we refer to Navarro et al. [32], Ahmadi et al. [2], Ghosh and Kundu [14], [13], [15], Kayal and Sunoj [18], Kundu and Kundu [25], [26], and the references therein.

The outline of this paper is as follows: In Section 2, we provide one parametric generalization of conditional CPI for the marginal past lifetimes $\widehat{X}_i$, $i = 1, 2$, called generalized conditional CPI (GCCPI). We investigate several properties of the new notion, such as the effect of monotone transformations, and obtain some bounds in context of usual stochastic order. In Section 3, we provide some characterization results based on GCCPI under the assumption of conditional proportional reversed hazard rate (CPRHR) model. Note that the results corresponding to the conditional past lifetimes $\widehat{X}^*_{i|j}$, $i=1,2$, $j=3-i$ are analogous to that of Section 2 and hence not presented here. Finally, in Section 4, we conclude the present study with a real-life application in reliability modeling.

2. Definition and properties of GCCPI for $(X_i \mid X_j < t_j)$

Let (X_1,X_2), (Y_1,Y_2), and (Z_1,Z_2) be three absolutely continuous bivariate random vectors with common support $(0,l)\times(0,l)$ for $l \geqslant 0$. Note that l can be equal to ∞. The joint *pdf* and *cdf* of (X_1,X_2) are denoted by f and F and those of (Y_1,Y_2) by g and G and of (Z_1,Z_2) by h and H, respectively. Consider the marginal past lifetimes $(X_i \mid X_1 < t_1, X_2 < t_2)$, $(Y_i \mid Y_1 < t_1, Y_2 < t_2)$ and $(Z_i \mid Z_1 < t_1, Z_2 < t_2)$ for $i=1,2$. For simplicity we denote them by $\widehat{X}_i$, $\widehat{Y}_i$ and $\widehat{Z}_i$, $i=1,2$, respectively. We denote the *pdf*'s and *cdf*'s of these random variables by $f_{\widehat{X}_i}$, $F_{\widehat{X}_i}$; $g_{\widehat{Y}_i}$, $G_{\widehat{Y}_i}$ and $h_{\widehat{Z}_i}$, $H_{\widehat{Z}_i}$, respectively. Now, for $i=1,2$, we define $f_{\widehat{X}_1}(x_1)=f_1(x_1,t_2)/F(t_1,t_2)$, $f_{\widehat{X}_2}(x_2)=f_2(t_1,x_2)/F(t_1,t_2)$, $F_{\widehat{X}_1}=F(x_1,t_2)/F(t_1,t_2)$, and $F_{\widehat{X}_2}=F(t_1,x_2)/F(t_1,t_2)$, where $f_1(x_1,t_2)=(\partial/\partial x_1)F(x_1,t_2)$, $f_2(t_1,x_2)=(\partial/\partial x_2)F(t_1,x_2)$, $0\leqslant x_1\leqslant t_1$, $0\leqslant x_2\leqslant t_2$, with a similar definition for $\widehat{Y}_i$ and $\widehat{Z}_i$. Then the CPI for $\widehat{X}_i$ and $\widehat{Y}_i$, called conditional CPI (CCPI), are defined as

$$\mathcal{C}\overline{K}_{X_1,Y_1}(t_1,t_2) = -\int_0^{t_1} \frac{F(x_1,t_2)}{F(t_1,t_2)} \ln\Big(\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big)\,\mathrm{d}x_1 \tag{2.1}$$

and

$$\mathcal{C}\overline{K}_{X_2,Y_2}(t_1,t_2) = -\int_0^{t_2} \frac{F(t_1,x_2)}{F(t_1,t_2)} \ln\Big(\frac{G(t_1,x_2)}{G(t_1,t_2)}\Big)\,\mathrm{d}x_2, \tag{2.2}$$

where $t_1,t_2\geqslant 0$. Several aspects of (2.1)–(2.2) have recently been discussed in Ghosh and Kundu [13]. Along a similar line, CPI(α) for $\widehat{X}_i$ and $\widehat{Y}_i$, $i=1,2$, called generalized conditional CPI (GCCPI) measure, can be defined as

$$\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2) = \frac{1}{1-\alpha}\ln\int_0^{t_1} \frac{F(x_1,t_2)}{F(t_1,t_2)} \Big(\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big)^{\alpha-1}\,\mathrm{d}x_1 \tag{2.3}$$

and

$$\mathcal{C}\overline{K}^{\alpha}_{X_2,Y_2}(t_1,t_2) = \frac{1}{1-\alpha}\ln\int_0^{t_2} \frac{F(t_1,x_2)}{F(t_1,t_2)} \Big(\frac{G(t_1,x_2)}{G(t_1,t_2)}\Big)^{\alpha-1}\,\mathrm{d}x_2, \tag{2.4}$$

where $t_1, t_2 \geqslant 0$ and $0 < \alpha \neq 1$. Note that (2.3) and (2.4) can be thought of as a generalization of conditional CPE of order α (CCPE(α)) studied by Kundu and Kundu [26] which are given below. The CCPE(α) of the bivariate random vector $X = (X_1, X_2)$, takes the form

$$\overline{\varepsilon}^*_{1,\alpha}(X; t_1, t_2) = \frac{1}{1-\alpha} \ln \int_0^{t_1} \Big(\frac{F(x_1, t_2)}{F(t_1, t_2)}\Big)^{\alpha} \, \mathrm{d}x_1 \tag{2.5}$$

and

$$\overline{\varepsilon}^*_{2,\alpha}(X; t_1, t_2) = \frac{1}{1-\alpha} \ln \int_0^{t_2} \Big(\frac{F(t_1, x_2)}{F(t_1, t_2)}\Big)^{\alpha} \, \mathrm{d}x_2, \tag{2.6}$$

where $t_1, t_2 \geqslant 0$ and $0 < \alpha \neq 1$. Now we consider the following example.

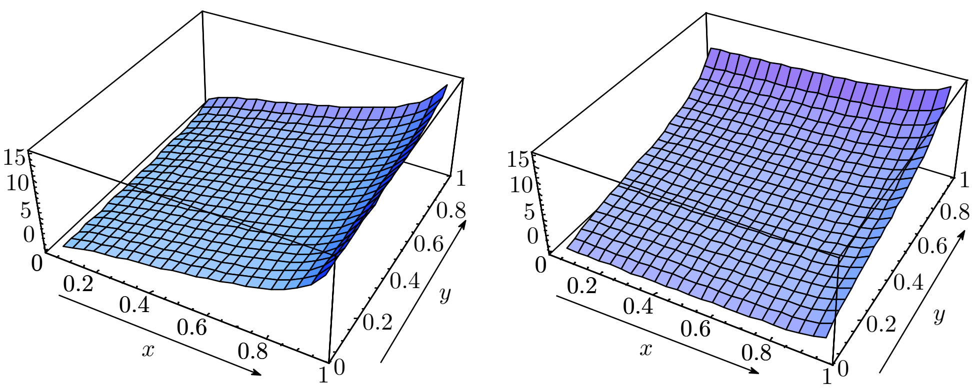
Example 2.1. Let $X = (X_1, X_2)$ follow the standard bivariate logistic distribution (cf. Gumbel [16]) with joint *cdf*

$$F(t_1, t_2) = (1 + \mathrm{e}^{-t_1} + \mathrm{e}^{-t_2})^{-1}, \quad t_1, t_2 > 0,$$

and let $Y = (Y_1, Y_2)$ follow the bivariate inverse exponential distribution

$$G(t_1, t_2) = \exp\Big(-\frac{1}{t_1} - \frac{1}{t_2} - \frac{\theta}{t_1 t_2}\Big), \quad 0 \leqslant \theta \leqslant 1, \ t_1, t_2 > 0.$$

Then Figure 1 gives $\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2)$ for $\alpha = 1.5$ and $\theta = 0.75$. It is to be mentioned here that while plotting curves, the substitutions $t_1 = -\ln x$ and $t_2 = -\ln y$ have been used so that $\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1, t_2) = \mathcal{C}\overline{K}^{\alpha}_1(x, y)$ and $\mathcal{C}\overline{K}^{\alpha}_{X_2,Y_2}(t_1, t_2) = \mathcal{C}\overline{K}^{\alpha}_2(x, y)$, (say).



(i) Plot of $\mathcal{C}\overline{K}^{\alpha}_1(x, y)$. (ii) Plot of $\mathcal{C}\overline{K}^{\alpha}_2(x, y)$.

Figure 1. Graphical representation of $\mathcal{C}\overline{K}^{\alpha}_1(x, y)$ and $\mathcal{C}\overline{K}^{\alpha}_2(x, y)$ (Example 2.1).

In order to pinpoint a probabilistic meaning of GCCPI, let us introduce the following functions for $0 \leqslant x_i < t_i$, $i = 1, 2$,

$$\eta_1^{(2)}(x_1, t_1; t_2) = \int_{x_1}^{t_1} (G(v_1, t_2))^{(\alpha-1)} \, \mathrm{d}v_1$$

and

$$\eta_2^{(2)}(x_2, t_2; t_1) = \int_{x_2}^{t_2} (G(t_1, v_2))^{(\alpha-1)} \, \mathrm{d}v_2.$$

It is worth mentioning that $(\partial/\partial t_1)\eta_1^{(2)}(x_1, t_1; t_2) = (G(t_1, t_2))^{(\alpha-1)}$. Thus, the significance of $\eta_1^{(2)}(x_1, t_1; t_2)$ is that its partial derivative is closely related to the distribution function of Y. The interpretation for $\eta_2^{(2)}(x_2, t_2; t_1)$ is similar. The following theorem provides a relation between GCCPI and $\eta_i^{(2)}(x_i, t_i; t_j)$, $i = 1, 2$, $j = 3 - i$, which also generalizes Theorem 3.1 of Kundu and Kundu [26].

Theorem 2.1. *Let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ be two absolutely continuous nonnegative bivariate random variables with joint cdf's $F(t_1, t_2)$ and $G(t_1, t_2)$, respectively. Then, for all $t_1, t_2 > 0$ and $i = 1, 2$, $j = 3 - i$,*

$$E[\eta_i^{(2)}(X_i, t_i; t_j) \mid X_1 < t_1, X_2 < t_2] = (G(t_1, t_2))^{(\alpha-1)} \mathrm{e}^{(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_i, Y_i}(t_1, t_2)},$$

where $0 < \alpha \neq 1$.

Proof. Using Fubini's theorem for $i = 1$ and $t_1, t_2 > 0$, we have

$$\begin{aligned}
E[\eta_1^{(2)}&(X_1, t_1; t_2) \mid X_1 < t_1, X_2 < t_2] \\
&= \int_0^{t_1} \Big(\int_{x_1}^{t_1} (G(v_1, t_2))^{(\alpha-1)} \, \mathrm{d}v_1 \Big) \frac{f_1(x_1, t_2)}{F(t_1, t_2)} \, \mathrm{d}x_1 \\
&= \int_0^{t_1} \Big(\int_0^{v_1} \frac{f_1(x_1, t_2)}{F(t_1, t_2)} \, \mathrm{d}x_1 \Big) (G(v_1, t_2))^{(\alpha-1)} \, \mathrm{d}v_1 \\
&= \int_0^{t_1} \frac{F(v_1, t_2)}{F(t_1, t_2)} \Big(\frac{G(v_1, t_2)}{G(t_1, t_2)} \Big)^{(\alpha-1)} (G(t_1, t_2))^{(\alpha-1)} \, \mathrm{d}x_1,
\end{aligned}$$

which gives the stated result for $i = 1$. The proof for $i = 2$ follows in the same line. □

Let us now define bivariate reversed hazard rate (BRHR) and bivariate expected inactivity time (BEIT). For more properties of these two functions one may refer to Roy [38] and Nair and Asha [29], respectively.

Definition 2.1. For a random vector $X = (X_1, X_2)$ with distribution function $F(t_1, t_2)$,

(i) the BRHR is defined as a vector $\phi^X(t_1, t_2) = (\phi_1^X(t_1, t_2), \phi_2^X(t_1, t_2))$, where $\phi_i^X(t_1, t_2) = (\partial/\partial t_i) \ln F(t_1, t_2)$, $i = 1, 2$ are the components of BRHR;

(ii) the BEIT, is defined as a vector $\overline{m}^X(t_1, t_2) = (\overline{m}_1^X(t_1, t_2), \overline{m}_2^X(t_1, t_2))$, where

$$\overline{m}_i^X(t_1, t_2) = E(t_i - X_i \mid X_1 < t_1, X_2 < t_2), \quad i = 1, 2, \ t_i > 0.$$

A fundamental relationship between BRHR and BEIT is

$$\phi_1^X(t_1, t_2)\overline{m}_1^X(t_1, t_2) = 1 - \frac{\partial}{\partial t_1}\overline{m}_1^X(t_1, t_2). \tag{2.7}$$

Hereafter, we study some properties and obtain few bounds of GCCPI in terms of CCPE(α) and BEIT. It is worthwhile to mention that for most of the well-known bivariate models, the proposed measures have no simple closed form to perform the analytic treatment and this necessitates to construct such bounds to get some idea of the corresponding measures very easily. Also, these bounds are quite useful in studying comparative behaviour of the proposed measures in reliability theory and for applications in other disciplines.

Theorem 2.2. *Let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ be two nonnegative bivariate random variables. Then for $0 < \alpha \neq 1$, we have*

$$\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2) \geqslant \frac{1}{1-\alpha} \ln[\overline{m}_i^X(t_1, t_2)],$$

where $\overline{m}_i^X(t_1, t_2)$, $i = 1, 2$, are the components of BEIT of X.

Proof. It is known that for $i = 1$, $G(x_1, t_2)$ is nondecreasing in x_1 for fixed t_2. As a consequence, for $x_1 < t_1$ we have

$$\Big(\frac{G(x_1, t_2)}{G(t_1, t_2)}\Big)^{\alpha-1} \begin{cases} \geqslant 1 \\ \leqslant 1 \end{cases} \text{according as} \begin{cases} 0 < \alpha < 1, \\ \alpha > 1. \end{cases}$$

Hence the stated result, for $i = 1$, follows from (2.3). The proof for $i = 2$ is analogous and hence omitted. □

We give the following example to support the above theorem.

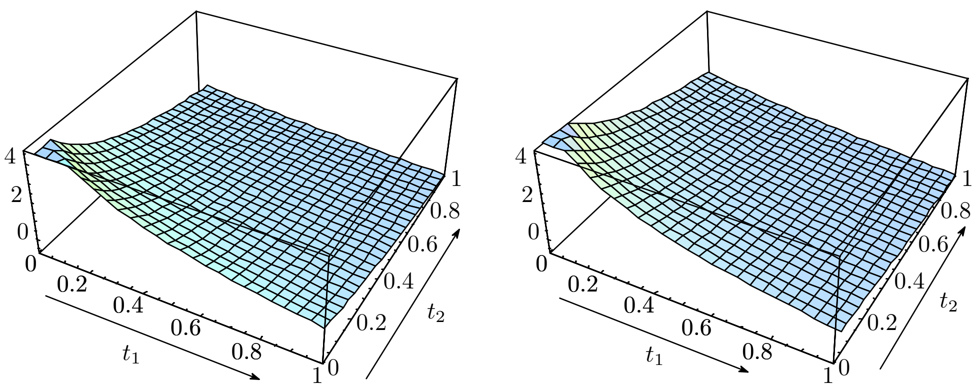
E x a m p l e 2.2. Let X and Y be two nonnegative bivariate random variables with *cdf*'s

$$F(t_1,t_2)=t_1t_2,\quad 0<t_1,\ t_2<1$$

and

$$G(t_1,t_2)=\frac{t_1t_2(t_1+t_2)}{2},\quad 0<t_1,\ t_2<1,$$

respectively. Then Fig. 2 shows that $[\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2)-1/(1-\alpha)^{-1}\ln(\overline{m}^X_i(t_1,t_2))]=\xi^{\alpha}_i(t_1,t_2)$, (say), are always positive for $\alpha=0.75$ and $i=1,2$, satisfying Theorem 2.2.



(i) Plot of $\xi^{\alpha}_1(t_1,t_2)$. (ii) Plot of $\xi^{\alpha}_2(t_1,t_2)$.

Figure 2. Graphical representation of ξ^{α}_1 and ξ^{α}_2 (Example 2.2).

Recall that for two univariate nonnegative random variables X and Y with *cdf*'s F and G, respectively, X is said to be less than Y in the convex order, written as $X\leqslant_{\rm cx}Y$, if for all convex functions $\phi\colon\ \mathbb{R}\to\mathbb{R}$,

$$E[\phi(X)]\leqslant E[\phi(Y)] \tag{2.8}$$

provided the expectations exist (see Shaked and Shanthikumar [40]). One can see that (2.8) is equivalent to

$$\int_0^x F(u)\,{\rm d}u\leqslant\int_0^x G(u)\,{\rm d}u\quad\forall\,x.$$

Theorem 2.3. *Let $X=(X_1,X_2)$ and $Y=(Y_1,Y_2)$ be two nonnegative bivariate random variables. For all $t_1,t_2>0$*

(i) *if $\widehat{X}_i\geqslant_{\rm cx}\widehat{Y}_i$ and $0<\alpha<1$, then*

$$\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2)\geqslant\frac{1}{1-\alpha}\ln[\overline{m}^Y_i(t_1,t_2)],$$

(ii) *if* $\widehat{X}_i \leqslant_{\rm cx} \widehat{Y}_i$ *and* $\alpha > 1$, *then*

$$\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2) \geqslant \frac{1}{1-\alpha}\ln[\overline{m}^{Y}_{i}(t_1,t_2)],$$

where $\overline{m}^{Y}_{i}(t_1,t_2)$, $i=1,2$, *are the components of BEIT of* Y.

P r o o f. Let us assume for $i=1$ that $\widehat{X}_1 \geqslant_{\rm cx} \widehat{Y}_1$, which leads to

$$\int_0^{t_1}\frac{F(x_1,t_2)}{F(t_1,t_2)}\,{\rm d}x_1 \geqslant \int_0^{t_1}\frac{G(x_1,t_2)}{G(t_1,t_2)}\,{\rm d}x_1. \tag{2.9}$$

Now, using the nondecreasing property of $G(x_1,t_2)$ with respect to x_1 for fixed t_2, we have for $0<\alpha<1$

$$\int_0^{t_1}\frac{F(x_1,t_2)}{F(t_1,t_2)}\Big(\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big)^{\alpha-1}{\rm d}x_1 \geqslant \int_0^{t_1}\frac{F(x_1,t_2)}{F(t_1,t_2)}\,{\rm d}x_1.$$

Using (2.9), we get

$$\ln\int_0^{t_1}\frac{F(x_1,t_2)}{F(t_1,t_2)}\Big(\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big)^{\alpha-1}{\rm d}x_1 \geqslant \ln\int_0^{t_1}\frac{G(x_1,t_2)}{G(t_1,t_2)}\,{\rm d}x_1.$$

Multiplying both sides by $1/(1-\alpha)$, we obtain the required result for $i=1$. The proof of the other cases is similar and hence omitted. $\square$

Theorem 2.4. *Let* X *and* Y *be two nonnegative bivariate random variables. Then for* $\alpha>1$ *we have*

$$\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2) \geqslant \frac{2-\alpha}{1-\alpha}\overline{\varepsilon}^{*}_{i,\alpha-1}(Y;t_1,t_2), \quad i=1,2,$$

where $\overline{\varepsilon}^{*}_{i,\alpha-1}(Y;t_1,t_2)$ *is the conditional CPE of order* $(\alpha-1)$ *for the random vector* Y.

P r o o f. For $i=1$, using the nondecreasing property of $F(x_1,t_2)$ in x_1 for fixed t_2, we get

$$\frac{F(x_1,t_2)}{F(t_1,t_2)} \leqslant 1,$$

for $x_1 \leqslant t_1$. Multiplying both sides by $\Big(\dfrac{G(x_1,t_2)}{G(t_1,t_2)}\Big)^{\alpha-1}$, we obtain

$$\ln\int_0^{t_1}\frac{F(x_1,t_2)}{F(t_1,t_2)}\Big(\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big)^{\alpha-1}{\rm d}x_1 \leqslant \ln\int_0^{t_1}\Big(\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big)^{\alpha-1}{\rm d}x_1.$$

Hence, the stated result is obtained for $i=1$ by multiplying both sides by $1/(1-\alpha)$. The proof for $i=2$ being analogous is omitted. $\square$

The following example illustrates the above theorem.

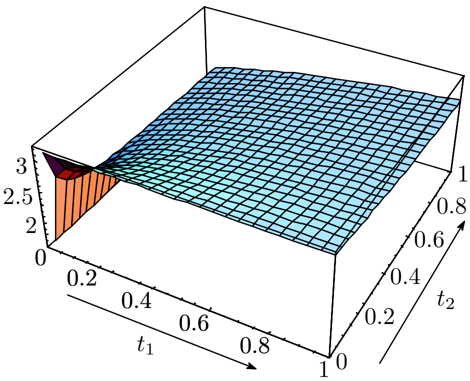
Example 2.3. Let X and Y be two nonnegative bivariate random variables with *cdf*'s

$$F(t_1,t_2)=\frac{t_1t_2(t_1+t_2)}{2},\quad 0<t_1,\ t_2<1,$$

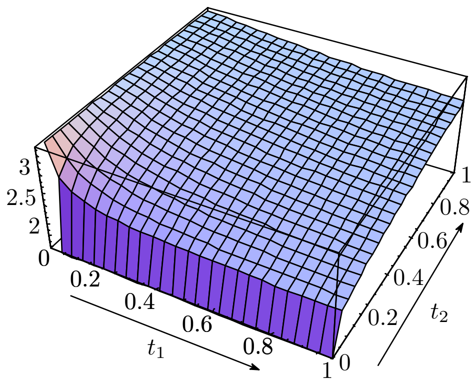
and

$$G(t_1,t_2)=\frac{1}{1/t_1+1/t_2-1},\quad 0<t_1,\ t_2<1,$$

respectively. Then Figure 3 shows that $[\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2)-(\frac{2-\alpha}{1-\alpha})\overline{\varepsilon}^{*}_{i,\alpha-1}(Y;t_1,t_2)]=\psi_i^{\alpha}(t_1,t_2)$, $i=1,2$, (say), are always positive for $\alpha=1.25$ satisfying Theorem 2.4.



(i) Plot of $\psi_1^{\alpha}(t_1,t_2)$.



(ii) Plot of $\psi_2^{\alpha}(t_1,t_2)$.

Figure 3. Graphical representation of ψ_1^{α} and ψ_2^{α} (Example 2.3).

Theorem 2.5. *Let X and Y be two nonnegative bivariate random vectors. Then, for $i=1,2$ and $t_1,t_2>0$, we have*

$$[1-\mathrm{e}^{-(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2)}]\leqslant(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2)\leqslant[\mathrm{e}^{(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2)}-1].$$

Proof. For $i=1$, using the fact that $\ln x\leqslant x-1$, we obtain

$$\ln\int_0^{t_1}\frac{F(x_1,t_2)}{F(t_1,t_2)}\Big(\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big)^{\alpha-1}\,\mathrm{d}x_1\leqslant\int_0^{t_1}\frac{F(x_1,t_2)}{F(t_1,t_2)}\Big(\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big)^{\alpha-1}\,\mathrm{d}x_1-1.$$

By some algebraic manipulation and with help of (2.3), we have

$$(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2)\leqslant[\mathrm{e}^{(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2)}-1]\quad\text{for } 0<\alpha\neq1. \tag{2.10}$$

Again, using $\ln x \geqslant (x-1)/x$, along a similar line, we get

$$(2.11)\qquad (1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2) \geqslant \frac{\mathrm{e}^{(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2)}-1}{\mathrm{e}^{(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2)}},$$

where $0 < \alpha \neq 1$. Thus the desired result for $i = 1$ follows by combining (2.10) and (2.11). The proof for $i = 2$ follows in the same line. $\square$

In the next theorem, we analyze the effect of monotone transformation on GCCPI to obtain their bounds. Heuristically, the results enable one to examine the information properties of lifetime models that can be obtained by transformation of simpler models.

Theorem 2.6. *Let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ be two absolutely continuous nonnegative bivariate random variables. Suppose $\varphi(x)$ is a strictly increasing, continuous and differentiable function. If $a \leqslant \varphi'(x) \leqslant b, a, b > 0$, then for all $t_1, t_2 > 0$ and $i = 1, 2$,*

$$\begin{aligned} a\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(\varphi^{-1}(t_1),\varphi^{-1}(t_2)) &\leqslant \mathcal{C}\overline{K}^{\alpha}_{\varphi(X_i),\varphi(Y_i)}(t_1,t_2)\\ &\leqslant b\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(\varphi^{-1}(t_1),\varphi^{-1}(t_2)), \quad \text{for } 0<\alpha<1 \end{aligned}$$

and

$$\begin{aligned} b\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(\varphi^{-1}(t_1),\varphi^{-1}(t_2)) &\leqslant \mathcal{C}\overline{K}^{\alpha}_{\varphi(X_i),\varphi(Y_i)}(t_1,t_2)\\ &\leqslant a\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(\varphi^{-1}(t_1),\varphi^{-1}(t_2)), \quad \text{for } \alpha>1. \end{aligned}$$

P r o o f. Let us consider the conditional random variables

$$\varphi(\widehat{X}_i) = (\varphi(X_i) \mid \varphi(X_1) < t_1,\ \varphi(X_2) < t_2), \quad i = 1,2,$$

and similar expression for $\varphi(\widehat{Y}_i)$. First we prove the result for $i = 1$. From (2.3), we have

$$\begin{aligned} &\mathcal{C}\overline{K}^{\alpha}_{\varphi(X_1),\varphi(Y_1)}(t_1,t_2)\\ &= \frac{1}{1-\alpha}\ln\int_0^{t_1}\frac{F(\varphi^{-1}(x_1),\varphi^{-1}(t_2))}{F(\varphi^{-1}(t_1),\varphi^{-1}(t_2))}\Big(\frac{G(\varphi^{-1}(x_1),\varphi^{-1}(t_2))}{G(\varphi^{-1}(t_1),\varphi^{-1}(t_2))}\Big)^{\alpha-1}\,\mathrm{d}x_1\\ &= \frac{1}{1-\alpha}\ln\int_0^{\varphi^{-1}(t_1)}\frac{F(y_1,\varphi^{-1}(t_2))}{F(\varphi^{-1}(t_1),\varphi^{-1}(t_2))}\Big(\frac{G(y_1,\varphi^{-1}(t_2))}{G(\varphi^{-1}(t_1),\varphi^{-1}(t_2))}\Big)^{\alpha-1}\varphi'(y_1)\,\mathrm{d}y_1. \end{aligned}$$

Thus, the stated result is obtained by using the fact that $a \leqslant \varphi'(x) \leqslant b$ and $0 < \alpha < 1$ or $\alpha > 1$. A similar result is obtained for $i = 2$. $\square$

In the sequel we obtain some bounds for GCCPI based on the usual stochastic order. We recall that for two univariate random variables X and Y with *cdf*'s F and G, respectively, X is said to be less than Y in the usual stochastic order, written as $X \leqslant_{\rm st} Y$, if $F(x) \geqslant G(x)$.

Theorem 2.7. *Let X and Y be two absolutely continuous nonnegative random vectors. For $i = 1, 2$ and $t_1, t_2 > 0$*

(i) *if $\widehat{X}_i \leqslant_{\rm st} \widehat{Y}_i$ then $\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2) \geqslant \max\{\overline{\varepsilon}^*_{i,\alpha}(X;t_1,t_2), \overline{\varepsilon}^*_{i,\alpha}(Y;t_1,t_2)\}$ when $0 < \alpha < 1$, and $\overline{\varepsilon}^*_{i,\alpha}(X;t_1,t_2) \leqslant \mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2) \leqslant \overline{\varepsilon}^*_{i,\alpha}(Y;t_1,t_2)$ when $\alpha > 1$;*

(ii) *if $\widehat{X}_i \geqslant_{\rm st} \widehat{Y}_i$ then $\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2) \leqslant \min\{\overline{\varepsilon}^*_{i,\alpha}(X;t_1,t_2), \overline{\varepsilon}^*_{i,\alpha}(Y;t_1,t_2)\}$ for $0 < \alpha < 1$ and $\overline{\varepsilon}^*_{i,\alpha}(Y;t_1,t_2) \leqslant \mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2) \leqslant \overline{\varepsilon}^*_{i,\alpha}(X;t_1,t_2)$ for $\alpha > 1$.*

P r o o f. Let us suppose, for $i = 1$, that $\widehat{X}_1 \leqslant_{\rm st} \widehat{Y}_1$ holds. Then

$$F_{\widehat{X}_1}(x_1,t_2) \geqslant G_{\widehat{Y}_1}(x_1,t_2) \tag{2.12}$$

for $x_1 \leqslant t_1$. Taking algebraic power $(\alpha - 1)$ on both sides of (2.12) and multiplying by $F_{\widehat{X}_1}(x_1,t_2)$, we have

$$\ln \int_0^{t_1} F_{\widehat{X}_1}(x_1,t_2)(F_{\widehat{X}_1}(x_1,t_2))^{\alpha-1}\,{\rm d}x_1 \begin{Bmatrix} \leqslant \\ \geqslant \end{Bmatrix} \ln \int_0^{t_1} F_{\widehat{X}_1}(x_1,t_2)(G_{\widehat{Y}_1}(x_1,t_2))^{\alpha-1}\,{\rm d}x_1,$$

according as $0 < \alpha < 1$ or $\alpha > 1$. Multiplying both sides by $1/(1-\alpha)$ and simplifying, we obtain

$$\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2) \geqslant \overline{\varepsilon}^*_{1,\alpha}(X;t_1,t_2) \quad \text{for } 0 < \alpha \neq 1. \tag{2.13}$$

Again, multiplying both sides of (2.12) by $(G_{\widehat{Y}_1}(x_1,t_2))^{\alpha-1}$ and integrating, we get

$$\ln \int_0^{t_1} F_{\widehat{X}_1}(x_1,t_2)(G_{\widehat{Y}_1}(x_1,t_2))^{\alpha-1}\,{\rm d}x_1 \geqslant \ln \int_0^{t_1} (G_{\widehat{Y}_1}(x_1,t_2))^{\alpha}\,{\rm d}x_1.$$

Multiplying both sides by $1/(1-\alpha)$, we have

$$\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2) \begin{Bmatrix} \leqslant \\ \geqslant \end{Bmatrix} \overline{\varepsilon}^*_{1,\alpha}(Y;t_1,t_2), \text{ according as } 0 < \alpha < 1 \text{ or } \alpha > 1. \tag{2.14}$$

Combining (2.13) and (2.14), one can easily conclude that

$$\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2) \geqslant \max\{\overline{\varepsilon}^*_{1,\alpha}(X;t_1,t_2), \overline{\varepsilon}^*_{1,\alpha}(Y;t_1,t_2)\} \quad \text{when } 0 < \alpha < 1$$

and

$$\overline{\varepsilon}^*_{1,\alpha}(X;t_1,t_2) \leqslant \mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1,t_2) \leqslant \overline{\varepsilon}^*_{1,\alpha}(Y;t_1,t_2) \quad \text{when } \alpha > 1.$$

Hence (i) holds for $i = 1$. The proof of the other cases being analogous is omitted. $\square$

The proof of the following theorem is similar to that of Theorem 2.7 and hence omitted.

Theorem 2.8. *For two nonnegative bivariate random vectors* $X = (X_1, X_2)$ *and* $Y = (Y_1, Y_2)$,

(i) *if* $\widehat{X}_i \leqslant_{\rm st} \widehat{Y}_i$, $i = 1, 2$, *then*

$$\mathcal{C}\overline{K}^{\alpha}_{Y_i,X_i}(t_1, t_2) \leqslant \overline{\varepsilon}^*_{i,\alpha}(X; t_1, t_2) \leqslant \mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2) \quad \text{when } 0 < \alpha < 1$$

and

$$\overline{\varepsilon}^*_{i,\alpha}(X; t_1, t_2) \leqslant \min\{\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2), \mathcal{C}\overline{K}^{\alpha}_{Y_i,X_i}(t_1, t_2)\} \quad \text{when } \alpha > 1;$$

(ii) *Again, if* $\widehat{X}_i \geqslant_{\rm st} \widehat{Y}_i$, $i = 1, 2$, *then*

$$\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2) \leqslant \overline{\varepsilon}^*_{i,\alpha}(X; t_1, t_2) \leqslant \mathcal{C}\overline{K}^{\alpha}_{Y_i,X_i}(t_1, t_2) \quad \text{when } 0 < \alpha < 1$$

and

$$\overline{\varepsilon}^*_{i,\alpha}(X; t_1, t_2) \geqslant \max\{\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2), \mathcal{C}\overline{K}^{\alpha}_{Y_i,X_i}(t_1, t_2)\} \quad \text{when } \alpha > 1,$$

where $t_1, t_2 > 0$.

Theorem 2.9. *Let* $X = (X_1, X_2)$, $Y = (Y_1, Y_2)$ *and* $Z = (Z_1, Z_2)$ *be three nonnegative bivariate random vectors. If* $\widehat{Y}_i \{ {\leqslant \atop \geqslant} \} \widehat{Z}_i$, $i = 1, 2$, *then for* $t_1, t_2 > 0$

$$\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2) \left\{ {\leqslant \atop \geqslant} \right\} \mathcal{C}\overline{K}^{\alpha}_{X_i,Z_i}(t_1, t_2),$$

where $0 < \alpha \neq 1$.

P r o o f. Let $\widehat{Y}_i \leqslant_{\rm st} \widehat{Z}_i$ hold for $i = 1$. As an immediate consequence we have

$$\frac{G(x_1, t_2)/G(t_1, t_2)}{H(x_1, t_2)/H(t_1, t_2)} \geqslant 1. \tag{2.15}$$

Again, (2.3) can alternatively be rewritten as

$$\begin{aligned}(1 - \alpha)\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1, t_2) \\ = \ln \int_0^{t_1} \frac{F(x_1, t_2)}{F(t_1, t_2)} \Big(\frac{H(x_1, t_2)}{H(t_1, t_2)}\Big)^{\alpha-1} \Big(\frac{G(x_1, t_2)/G(t_1, t_2)}{H(x_1, t_2)/H(t_1, t_2)}\Big)^{\alpha-1} \, {\rm d}x_1.\end{aligned} \tag{2.16}$$

Hence, the stated result for $i = 1$ follows from (2.16) by making use of (2.15). The proof of the other cases follows in the same line. This completes the proof. $\square$

Theorem 2.10. *Let X, Y and Z be three nonnegative bivariate random vectors. If $\widehat{X}_i \leqslant_{\rm st} \widehat{Y}_i$, then for $t_1, t_2 > 0$ and $i = 1, 2$ we have*

$$\mathcal{C}\overline{K}^{\alpha}_{X_i,Z_i}(t_1,t_2) \begin{Bmatrix} \geqslant \\ \leqslant \end{Bmatrix} \mathcal{C}\overline{K}^{\alpha}_{Y_i,Z_i}(t_1,t_2)$$

for $0 < \alpha < 1$ or $\alpha > 1$.

P r o o f. Let us assume that $\widehat{X}_i \leqslant_{\rm st} \widehat{Y}_i$, $i = 1, 2$. Then for $i = 1$, $\widehat{X}_1 \leqslant_{\rm st} \widehat{Y}_1$ leads to

$$\frac{F(x_1,t_2)/F(t_1,t_2)}{G(x_1,t_2)/G(t_1,t_2)} \geqslant 1. \tag{2.17}$$

By applying (2.3), $\mathcal{C}\overline{K}^{\alpha}_{X_1,Z_1}(t_1,t_2)$ can be rewritten as

$$\begin{aligned}(1-\alpha)&\mathcal{C}\overline{K}^{\alpha}_{X_1,Z_1}(t_1,t_2) \\ &= \ln \int_0^{t_1} \frac{G(x_1,t_2)}{G(t_1,t_2)} \Big(\frac{H(x_1,t_2)}{H(t_1,t_2)}\Big)^{\alpha-1} \frac{F(x_1,t_2)/F(t_1,t_2)}{G(x_1,t_2)/G(t_1,t_2)} \,\mathrm{d}x_1.\end{aligned} \tag{2.18}$$

Therefore, the desired result for $i = 1$ is obtained by using (2.17) and (2.18). The proof of the other cases is similar and therefore omitted. □

As a continuation of the above result we have the following theorem.

Theorem 2.11. *Let X, Y and Z be three nonnegative bivariate random vectors. If for $i = 1, 2$, $\widehat{X}_i \leqslant_{\rm st} \widehat{Z}_i \leqslant_{\rm st} \widehat{Y}_i$, then*

$$\mathcal{C}\overline{K}^{\alpha}_{Y_i,X_i}(t_1,t_2) \leqslant \min\{\mathcal{C}\overline{K}^{\alpha}_{Y_i,Z_i}(t_1,t_2), \mathcal{C}\overline{K}^{\alpha}_{Z_i,X_i}(t_1,t_2)\}, \quad 0 < \alpha < 1,$$

and

$$\mathcal{C}\overline{K}^{\alpha}_{Z_i,X_i}(t_1,t_2) \leqslant \mathcal{C}\overline{K}^{\alpha}_{Y_i,X_i}(t_1,t_2) \leqslant \mathcal{C}\overline{K}^{\alpha}_{Y_i,Z_i}(t_1,t_2), \quad \alpha > 1,$$

where $t_1, t_2 > 0$.

P r o o f. With help of (2.3), for $i = 1$, $\mathcal{C}\overline{K}^{\alpha}_{Y_1,X_1}(t_1,t_2)$ can be written as

$$\begin{aligned}&\mathcal{C}\overline{K}^{\alpha}_{Y_1,X_1}(t_1,t_2) \\ &= \frac{1}{1-\alpha} \ln \int_0^{t_1} \frac{G(x_1,t_2)}{G(t_1,t_2)} \Big(\frac{H(x_1,t_2)}{H(t_1,t_2)}\Big)^{\alpha-1} \Big(\frac{F(x_1,t_2)/F(t_1,t_2)}{H(x_1,t_2)/H(t_1,t_2)}\Big)^{\alpha-1} \,\mathrm{d}x_1.\end{aligned} \tag{2.19}$$

Alternatively (2.19) can also be represented as

$$\begin{aligned}&\mathcal{C}\overline{K}^{\alpha}_{Y_1,X_1}(t_1,t_2) \\ &= \frac{1}{1-\alpha} \ln \int_0^{t_1} \frac{H(x_1,t_2)}{H(t_1,t_2)} \Big(\frac{F(x_1,t_2)}{F(t_1,t_2)}\Big)^{\alpha-1} \frac{G(x_1,t_2)/G(t_1,t_2)}{H(x_1,t_2)/H(t_1,t_2)} \,\mathrm{d}x_1.\end{aligned} \tag{2.20}$$

Now, let us assume for $i = 1$ that $\widehat{X}_1 \leqslant_{\mathrm{st}} \widehat{Z}_1$ holds. As a consequence, we have

$$\frac{F(x_1,t_2)/F(t_1,t_2)}{H(x_1,t_2)/H(t_1,t_2)} \geqslant 1.$$

After some algebraic manipulations, we obtain

$$\begin{aligned}\ln\int_0^{t_1}\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big(\frac{H(x_1,t_2)}{H(t_1,t_2)}\Big)^{\alpha-1}\Big(\frac{F(x_1,t_2)/F(t_1,t_2)}{H(x_1,t_2)/H(t_1,t_2)}\Big)^{\alpha-1}\,\mathrm{d}x_1\\ \begin{Bmatrix}\leqslant\\ \geqslant\end{Bmatrix}\ln\int_0^{t_1}\frac{G(x_1,t_2)}{G(t_1,t_2)}\Big(\frac{H(x_1,t_2)}{H(t_1,t_2)}\Big)^{\alpha-1}\,\mathrm{d}x_1,\end{aligned}$$

according as $0 < \alpha < 1$ or $\alpha > 1$. Multiplying both sides by $1/(1-\alpha)$ and using (2.19), we obtain

$$\mathcal{C}\overline{K}^{\alpha}_{Y_1,X_1}(t_1,t_2) \leqslant \mathcal{C}\overline{K}^{\alpha}_{Y_1,Z_1}(t_1,t_2) \quad \text{for } 0 < \alpha \neq 1. \tag{2.21}$$

Again, assuming for $i = 1$ that $\widehat{Z}_1 \leqslant_{\mathrm{st}} \widehat{Y}_1$ holds and then proceeding in the same line, we get

$$\begin{aligned}\ln\int_0^{t_1}\frac{H(x_1,t_2)}{H(t_1,t_2)}\Big(\frac{F(x_1,t_2)}{F(t_1,t_2)}\Big)^{\alpha-1}\frac{G(x_1,t_2)/G(t_1,t_2)}{H(x_1,t_2)/H(t_1,t_2)}\,\mathrm{d}x_1\\ \leqslant \ln\int_0^{t_1}\frac{H(x_1,t_2)}{H(t_1,t_2)}\Big(\frac{F(x_1,t_2)}{F(t_1,t_2)}\Big)^{\alpha-1}\,\mathrm{d}x_1,\end{aligned} \tag{2.22}$$

for $0 < \alpha \neq 1$. Multiplying by $1/(1-\alpha)$ and then with the aid of (2.20), we have

$$\mathcal{C}\overline{K}^{\alpha}_{Y_1,X_1}(t_1,t_2)\begin{Bmatrix}\leqslant\\ \geqslant\end{Bmatrix}\mathcal{C}\overline{K}^{\alpha}_{Z_1,X_1}(t_1,t_2), \text{ according as } 0 < \alpha < 1 \text{ or } \alpha > 1. \tag{2.23}$$

Thus, from (2.21) and (2.22), we conclude that the given results hold for $i = 1$. The proof for $i = 2$ being analogous is omitted. □

3. Characterization

In the last three decades, several attempts have been made to characterize probability distributions in both the univariate and multivariate setup. Characterizations of multivariate distributions have become a topic of great interest in the literature of applied statistics, reliability and information theory. In this section we discuss characterization theorems associated with some bivariate models based on the functional form of GCCPI, BRHR and BEIT. Although the uniqueness of a characterization result is highly questionable, nonetheless under some conditions GCCPI uniquely determines the parent distribution. Recall that the random vectors $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ are said to satisfy the conditional proportional reversed hazard rate (CPRHR) model when the corresponding conditional reversed hazard rate functions of $\widehat{X}_i$ and $\widehat{Y}_i$ satisfy

$$\phi_i^Y(t_1, t_2) = \theta_i(t_j)\phi_i^X(t_1, t_2) \tag{3.1}$$

for $i = 1, 2$; $j = 3 - i$, and $t_1, t_2 \geqslant 0$. Here $\theta_1(t_2)$ and $\theta_2(t_1)$ are positive functions of t_2 and t_1, respectively. In the following theorem we show that GCCPI determines the distribution function uniquely under CPRHR model assumption.

Theorem 3.1. *Let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ be two absolutely continuous nonnegative random vectors with cdf's $F(t_1, t_2)$, $G(t_1, t_2)$, respectively, and satisfy the CPRHR model given in* (3.1). *If $(\partial/\partial t_i)F(t_1, t_2)$ is continuous in t_i, $i = 1, 2$, then for each α, $\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2)$ uniquely determine $F(t_1, t_2)$.*

Proof. Let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ satisfy the CPRHR model with the constant of proportion $\theta_i(t_j)$, $i = 1, 2$; $j = 3 - i$. Assume that $Z = (Z_1, Z_2)$ and $W = (W_1, W_2)$ are other two absolutely continuous random variables with the same support as that of X and Y and also satisfy the CPRHR model with the same constant of proportion, i.e. $\theta_i(t_j)$. Now for $i = 1$, $\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1, t_2) = \mathcal{C}\overline{K}^{\alpha}_{Z_1,W_1}(t_1, t_2)$ implies that

$$(1-\alpha)\frac{\partial}{\partial t_1}\mathcal{C}\overline{K}^{\alpha}_{X_1,Y_1}(t_1, t_2) = (1-\alpha)\frac{\partial}{\partial t_1}\mathcal{C}\overline{K}^{\alpha}_{Z_1,W_1}(t_1, t_2).$$

The above equation simplifies to

$$(\alpha-1)\phi_1^Y(t_1, t_2) + \phi_1^X(t_1, t_2) = (\alpha-1)\phi_1^W(t_1, t_2) + \phi_1^Z(t_1, t_2).$$

Under the CPRHR model assumption, we have

$$\phi_1^X(t_1, t_2) = \phi_1^Z(t_1, t_2).$$

By retracing the above steps, we obtain

$$\phi_2^X(t_1, t_2) = \phi_2^Z(t_1, t_2).$$

Thus, in general

$$\phi_i^X(t_1, t_2) = \phi_i^Z(t_1, t_2), \quad i = 1, 2.$$

Hence, the result follows by using the fact that the vector valued reversed hazard rate uniquely determines the bivariate distribution function (cf. Roy [38]). □

In the sequel, we provide some characterization results with dependent marginals.

Theorem 3.2. *Let* $X = (X_1, X_2)$ *and* $Y = (Y_1, Y_2)$ *be two random vectors satisfying the CPRHR model given in* (3.1) *for* $i = 1, 2$, $j = 3 - i$. *If* $(\partial/\partial t_i)F(t_1, t_2)$ *is continuous in* t_i, *then for* $0 < t_1, t_2 < 1$, $i = 1, 2$; $j = 3 - i$,

$$(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2) = \ln[\omega_i(t_j)\overline{m}^X_i(t_1,t_2)] \tag{3.2}$$

if and only if X *follows the distribution*

$$F(t_1,t_2) = t_1^{1+\mu\ln t_2}t_2, \quad 0 < t_1, t_2 < 1,\ \mu \leqslant 0, \tag{3.3}$$

where

$$\omega_i(t_j) = \frac{2+\mu\ln t_j}{1+[\theta_i(t_j)(\alpha-1)+1](1+\mu\ln t_j)}.$$

In particular, for $\mu = 0$, (3.3) *is bivariate uniform.*

P r o o f. The if part is trivial when noting that if X follows the distribution (3.3) then for $i = 1, 2$; $j = 3 - i$,

$$\overline{m}^X_i(t_1,t_2) = \frac{t_i}{2+\mu\ln t_j}$$

and

$$\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2) = \frac{1}{1-\alpha}\ln\frac{t_i}{1+[\theta_i(t_j)(\alpha-1)+1](1+\mu\ln t_j)}.$$

To prove the converse let us assume that (3.2) holds. Under the CPRHR model assumptions, for $i = 1$, differentiating both sides of (3.2) with respect to t_1, we get

$$\phi^X_1(t_1,t_2)\overline{m}^X_1(t_1,t_2) = \frac{1+\mu\ln t_2}{2+\mu\ln t_2}.$$

Using (2.7), we have

$$\frac{\partial}{\partial t_1}\overline{m}^X_1(t_1,t_2) = \frac{1}{2+\mu\ln t_2},$$

which by integration gives

$$\overline{m}^X_1(t_1,t_2) = \frac{1}{2+\mu\ln t_2}t_1 + l_1(t_2),$$

where $l_1(t_2)$ is a constant of integration. As $t_1 \to 0$, $\overline{m}^X_1(t_1,t_2) \to 0$ implies $l_1(t_2) = 0$, which in turn gives the bivariate EIT of (3.3). Hence, the result follows for $i = 1$ by virtue of the fact that bivariate EIT determines the distribution uniquely. The proof of the other case is similar and hence omitted. □

The next theorem gives a characterization of the bivariate power distribution.

Theorem 3.3. *Let* (X_1, X_2) *and* (Y_1, Y_2) *be two absolutely continuous nonnegative random vectors with common support* $(0, b_1) \times (0, b_2)$ *satisfying the CPRHR model given in* (3.1) *for* $i = 1, 2$, $j = 3 - i$. *If* $(\partial/\partial t_i)F(t_1, t_2)$ *is continuous in* t_i, $i = 1, 2$, *then*

$$(1-\alpha)\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1,t_2) = \ln[d_i(t_j)\overline{m}^X_i(t_1,t_2)], \tag{3.4}$$

where

$$0 < d_i(t_j) < \frac{1}{1+(\alpha-1)\theta_i(t_j)}, \quad 0 < \theta_i(t_j) < \frac{1}{1-\alpha}$$

and

$$d_i(t_j) > \frac{1}{1+(\alpha-1)\theta_i(t_j)}, \quad \theta_i(t_j) > 0$$

according as $0 < \alpha < 1$ *and* $\alpha > 1$, *respectively, characterizes the bivariate power distribution*

$$F(t_1,t_2) = \Big(\frac{t_1}{b_1}\Big)^{c_1}\Big(\frac{t_2}{b_2}\Big)^{c_2+\mu\ln(t_1/b_1)}, \quad \mu \leqslant 0, \tag{3.5}$$

where

$$c_i = \frac{1-d_i(b_j)}{[(\alpha-1)\theta_i(b_j)+1]d_i(b_j)-1}.$$

P r o o f. The if part is straightforward. To prove the reverse implication, let us assume that (3.4) holds. Under the CPRHR model assumptions, for $i = 1$, (3.4) can alternatively be rewritten as

$$\int_0^{t_1}(F(x_1,t_2))^{\theta_1(t_2)(\alpha-1)+1}\,\mathrm{d}x_1 = d_1(t_2)(F(t_1,t_2))^{\theta_1(t_2)(\alpha-1)}\int_0^{t_1}F(x_1,t_2)\,\mathrm{d}x_1.$$

Differentiating both sides with respect to t_1 and using (2.7), we get after some algebraic manipulation

$$\frac{\partial}{\partial t_1}\overline{m}^X_1(t_1,t_2) = \frac{[(\alpha-1)\theta_1(t_2)+1]d_1(t_2)-1}{d_1(t_2)\theta_1(t_2)(\alpha-1)},$$

the integration of which gives

$$\overline{m}^X_1(t_1,t_2) = \frac{[(\alpha-1)\theta_1(t_2)+1]d_1(t_2)-1}{d_1(t_2)\theta_1(t_2)(\alpha-1)}t_1 + n_1(t_2),$$

where $n_1(t_2)$ is a constant of integration. As $t_1 \to 0$, $\overline{m}^X_1(t_1,t_2) \to 0$ implies $n_1(t_2) = 0$. By retracing the above steps, we obtain in general

$$\overline{m}^X_i(t_1,t_2) = \frac{[(\alpha-1)\theta_i(t_j)+1]d_i(t_j)-1}{d_i(t_j)\theta_i(t_j)(\alpha-1)}t_i, \quad i=1,2;\ j=3-i.$$

The rest of the proof follows from Theorem 2.1 of Nair and Asha [29]. $\square$

Theorem 3.4. *Let (X_1, X_2) and (Y_1, Y_2) be two absolutely continuous nonnegative random vectors with common support $(0, b_1) \times (0, b_2)$ satisfying the CPRHR model given in (3.1) for $i = 1, 2$; $j = 3 - i$. If $(\partial/\partial t_i)F(t_1, t_2)$ is continuous in t_i, $i = 1, 2$, then*

$$\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2) = C_i(t_j) - \frac{1}{1-\alpha}\ln[\phi_i^X(t_1, t_2)], \tag{3.6}$$

where

$$C_i(t_j) < \frac{1}{1-\alpha}\ln\frac{1}{1+(\alpha-1)\theta_i(t_j)}, \quad 0 < \theta_i(t_j) < \frac{1}{1-\alpha}$$

and

$$C_i(t_j) > \frac{1}{\alpha-1}\ln[1+(\alpha-1)\theta_i(t_j)], \quad \theta_i(t_j) > 0$$

according as $0 < \alpha < 1$ and $\alpha > 1$, respectively, characterizes the bivariate power distribution given in (3.5) with

$$c_i = \frac{\mathrm{e}^{(1-\alpha)C_i(b_j)}}{1 - [(\alpha-1)\theta_i(b_j)+1]\mathrm{e}^{(1-\alpha)C_i(b_j)}}.$$

Proof. The if part is straightforward. To prove the converse part, let us assume that (3.6) holds. Under the CPRHR model assumptions, for $i = 1$, some algebraic manipulations from (3.6), yield

$$\int_0^{t_1}(F(x_1, t_2))^{\theta_1(t_2)(\alpha-1)+1}\,\mathrm{d}x_1 = \frac{A_1(t_2)}{\phi_1^X(t_1, t_2)}(F(t_1, t_2))^{\theta_1(t_2)(\alpha-1)+1},$$

where $A_1(t_2) = \mathrm{e}^{(1-\alpha)C_1(t_2)}$. Differentiating both sides with respect to t_1 and simplifying, we obtain

$$\frac{\partial}{\partial t_1}\frac{1}{\phi_1^X(t_1, t_2)} = \frac{1 - [1+(\alpha-1)\theta_1(t_2)]A_1(t_2)}{A_1(t_2)},$$

which with help of (2.7) leads to

$$\phi_1^X(t_1, t_2)\overline{m}_1^X(t_1, t_2) = \frac{A_1(t_2)}{1-(\alpha-1)\theta_1(t_2)A_1(t_2)}.$$

Now, the rest of the proof follows from Theorem 3.1 of Nair and Asha [29]. A similar result is obtained for $i = 2$. This completes the proof. □

Now we provide a result where the support of the components is not restricted to positive real half line. Here we assume that $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ have a common support $\mathcal{D} = (-\infty, b_1) \times (-\infty, b_2)$ with $b_i < \infty$. Consequently, the lower limit of integrations in (2.3)–(2.4) will be taken with $-\infty$ instead of zero.

Theorem 3.5. *Let (X_1, X_2) and (Y_1, Y_2) be two absolutely continuous nonnegative random vectors with common support $\mathcal{D}$ and satisfy the CPRHR model given in* (3.1). *If $(\partial/\partial t_i)F(t_1, t_2)$ is continuous in t_i, $i = 1, 2$, then the following conditions are equivalent:*

(i) $\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2)$ *only depends on t_j for $i = 1, 2$; $j = 3 - i$,*

(ii) (X_1, X_2) *has the joint cdf's of the form*

$$F(t_1, t_2) = \exp[c_1(t_1 - b_1) + c_2(t_2 - b_2) + c_3(t_1 - b_1)(t_2 - b_2)],$$

where $c_i \geqslant 0$.

Proof. To prove (i) implies (ii), let us assume that $\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2)$ only depends on t_j for $i = 1, 2$; $j = 3 - i$, i.e. $\mathcal{C}\overline{K}^{\alpha}_{X_i,Y_i}(t_1, t_2) = c_i(t_j)$. Then for $i = 1$ we have from (2.3)

$$\int_{-\infty}^{t_1} (F(x_1, t_2))^{\theta_1(t_2)(\alpha-1)+1} \, \mathrm{d}x_1 = A_1(t_2)(F(t_1, t_2))^{\theta_1(t_2)(\alpha-1)+1},$$

where $A_1(t_2) = \mathrm{e}^{(1-\alpha)c_1(t_2)} > 0$. Now differentiating both sides partially with respect to t_1, we have

$$\phi_1^X(t_1, t_2) = \frac{1}{A_1(t_2)[\theta_1(t_2)(\alpha - 1) + 1]} = \frac{1}{B_1(t_2)}, \text{ (say)}.$$

Thus, $\phi_1^X(t_1, t_2)$ is independent of t_1 and only depends on t_2. Similarly, it can be shown that $\phi_2^X(t_1, t_2)$ only depends on t_1. Let us assume that $\phi_1^X(t_1, t_2) = \gamma_1(t_2)$ and $\phi_2^X(t_1, t_2) = \gamma_2(t_1)$. Using (2.7), we obtain after some algebraic manipulation $\overline{m}_i^X(t_1, t_2) = 1/\gamma_i(t_j) = \alpha_i(t_j)$, (say), for $i = 1, 2$; $j = 3 - i$. Now, the rest of the proof follows from Theorem 2.3 of Nair and Asha [29]. The converse part is straightforward.

4. Application to electrical appliance failure data

In this section, we illustrate the contribution of our proposed measure (i.e. GCCPI) in a real-life problem. Note that our approach is parametric which involves specific families of distributions.

Information measures appear to be suitable performance criteria in estimation and model selection. The use of information-theoretic approaches provides a new paradigm for model selection in the analysis of empirical data. Model selection based on information theory represents a quite different approach in the statistical sciences, and the resulting selected model may differ substantially from model selection based on some form of statistical null hypothesis testing. Though the information-theoretic methods may not always be the very best for a particular situation, they do represent a unified and rigorous theory, an extension of the likelihood theory, an important application of the information theory, and they are objective and practical to employ across a very wide class of empirical problems. Model selection, under the information-theoretic approach, attempts to identify the (likely) best model from the candidate models available and orders the models from best to worst. According to the information-theoretic approach for model selection due to Burnham and Anderson [5], for a given data set, the best fitted model is the one which has minimum Kullback-Leibler (K-L) information or distance. The K-L distance between models is a fundamental quantity in science and information theory (see Akaike [3]) and is the logical basis for model selection in conjunction with likelihood inference. A good model contains the information in the data, leaving only noise. Of course, we seek an approximating model that looses as little information as possible; this is equivalent to minimizing K-L information. Again, in view of (1.6), minimizing the K-L distance is equivalent to minimizing the (cumulative) inaccuracy measure (cross entropy). Note that the most well-known model selection criterion, the Akaike information criterion (AIC), is an asymptotically unbiased estimator of the inaccuracy measure from a parametric distribution to the true distribution of data. Minimizing the AIC can be interpreted as minimizing an asymptotically unbiased estimator of the inaccuracy/cross entropy. For some flavour of fascinating growth of cross entropy for comparing models, one may refer to Burnham and Anderson [5], Jurafsky and Martin [17] and Choe [7]. With this motivation, various generalized inaccuracy measures have been proposed in the literature which may contain minimum inaccuracy (but could not be less than the information given by the Shannon entropy).

The use of CPI ($\overline{\mathcal{K}}_{X,Y}$) for comparing the true distribution to the used distribution in statistical modeling has been discussed by Kundu et al. [24] in the univariate case. To see the effectiveness of GCCPI in reliability modeling we consider the electrical appliance failure data which are from an experiment in which new models of a small electrical appliance were being tested (see Lawless [28]). The data set consists of failure times or censoring times for 36 appliances subjected to an automated life test. Failures are mainly classified into 18 different modes, though among 33 observed failures only 7 modes are present and only modes 6 and 9 appear more than once. We are mainly interested in the failure mode 9. The data consist of two causes of failures, $\delta = 1$ (failure mode 9), $\delta = 2$ (all other failure modes), and $\delta = 0$ indicates that the data are censored at that time point. The data are given below:

Data Set: $(11, 2)$, $(35, 2)$, $(49, 2)$, $(170, 2)$, $(329, 2)$, $(381, 2)$, $(708, 2)$, $(958, 2)$, $(1062, 2)$, $(1167, 1)$, $(1594, 2)$, $(1925, 1)$, $(1990, 1)$, $(2223, 1)$, $(2327, 2)$, $(2400, 1)$, $(2451, 2)$, $(2471, 1)$, $(2551, 1)$, $(2565, 0)$, $(2568, 1)$, $(2694, 1)$, $(2702, 2)$, $(2761, 2)$, $(2831, 2)$, $(3034, 1)$, $(3059, 2)$, $(3112, 1)$, $(3214, 1)$, $(3478, 1)$, $(3504, 1)$, $(4329, 1)$, $(6367, 0)$, $(6976, 1)$, $(7846, 1)$, $(13403, 0)$.

The joint distribution of failure times and failure modes is of special interest. This can be used to help plan further development and testing of the appliance. The failure time distribution will change as the appliance is developed, and product improvements effectively remove certain causes of failure. In the final stages, the failure time distribution model can be used to predict the implications of a warranty plan for the appliance.

We now turn to an examination of statistical models for this data set. Sankaran and Kundu [39] used Lindley-Singpurwalla bivariate Pareto (LSBP) distribution

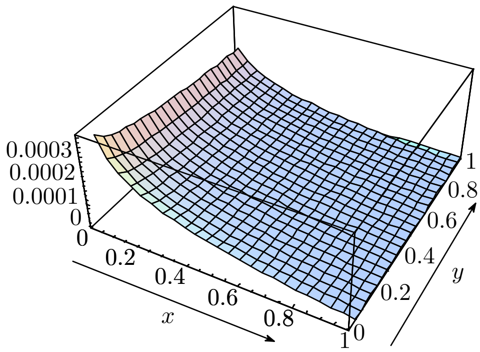
$$\overline{F}(x_1, x_2) = (1 + \beta_1 x_1 + \beta_2 x_2)^{-\theta}, \quad x_1, x_2 > 0, \tag{4.1}$$

to analyze the data and estimated $\widehat{\beta}_1 = 0.00019$, $\widehat{\beta}_2 = 0.00272$, and $\widehat{\theta} = 0.46688$. They also observed that the Sankaran-Nair bivariate Pareto (SNBP) distribution given by

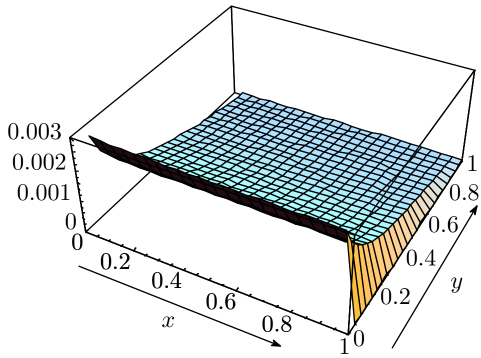
$$\overline{F}(x_1, x_2) = (1 + \gamma_1 x_1 + \gamma_2 x_2 + \gamma_0 x_1 x_2)^{-\phi}, \quad x_1, x_2 > 0, \tag{4.2}$$

can also be fitted to the data where $\widehat{\gamma}_1 = 0.000340$, $\widehat{\gamma}_2 = 0.00454$, $\widehat{\gamma}_0 = 0.000074$, and $\widehat{\phi} = 0.43103$. They have concluded that (4.1) provides a better fit than (4.2) to the given data set. Now we will use GCCPI to identify which of these two models is the closest to the distribution that generated the given data set. It has already been mentioned that between two models, the more accurate model will be the one with the lower inaccuracy measure. Let $X = (X_1, X_2)$ and $Y = (Y_1, Y_2)$ be two nonnegative bivariate random vectors which follow the distributions (4.1) and (4.2), respectively. Then Figure 4 shows that $[\mathcal{C}K^{\alpha}_{X_i,Y_i}(t_1, t_2) - \mathcal{C}K^{\alpha}_{Y_i,X_i}(t_1, t_2)] = \vartheta^{\alpha}_i(t_1, t_2)$, $i = 1, 2$, are always positive for $\alpha = 1.5$. Note that the substitutions $t_1 = -\ln x$ and $t_2 = -\ln y$ have been used while plotting curves so that $\vartheta^{\alpha}_i(t_1, t_2) = \vartheta^{\alpha}_i(x, y)$, $i = 1, 2$.

From Figure 4 we observe that the GCCPI of X on Y, for some specific value of the parameter, is less than when their role is inverted, which indeed shows that X provides a better approximation to the true distribution that generated the data set than Y. Hence, in agreement with Sankaran and Kundu [39], we can conclude that (4.1) gives a better fit to the given data set.



(i) Plot of $\vartheta_1^\alpha(x,y)$.



(ii) Plot of $\vartheta_2^\alpha(x,y)$.

Figure 4. Graphical representation of $\vartheta_1^\alpha(x,y)$ and $\vartheta_2^\alpha(x,y)$.

A c k n o w l e d g e m e n t s. We sincerely thank the anonymous reviewer for several helpful comments and suggestions which have considerably improved the content of the paper.

References

[1] *M. Abbasnejad*: Some characterization results based on dynamic survival and failure entropies. Commun. Stat. Appl. Methods *18* (2011), 787–798. doi

[2] *J. Ahmadi, A. Di Crescenzo, M. Longobardi*: On dynamic mutual information for bivariate lifetimes. Adv. Appl. Probab. *47* (2015), 1157–1174. zbl MR doi

[3] *H. Akaike*: Information measures and model selection. Bull. Int. Stat. Inst. *50* (1983), 277–290. zbl MR

[4] *S. Baratpour, A. H. Rad*: Testing goodness-of-fit for exponential distribution based on cumulative residual entropy. Commun. Stat., Theory Methods *41* (2012), 1387–1396. zbl MR doi

[5] *K. P. Burnham, D. R. Anderson*: Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, New York, 2002. zbl MR doi

[6] *N. D. Cahill, J. A. Schnabel, J. A. Noble, D. J. Hawkes*: Overlap invariance of cumulative residual entropy measures for multimodal image alignment. Medical Imaging 2009 (J. P. W. Pluim, B. M. Dawant, eds.). Proceedings of SPIE 7259, Society of Photo-Optical Instrumentation Engineers, Washington, 2009, Article ID 72590I. doi

[7] *Y. Choe*: Information criterion for minimum cross-entropy model selection. Available at `https://arxiv.org/abs/1704.04315` (2017), 32 pages.

[8] *A. Di Crescenzo, M. Longobardi*: On cumulative entropies. J. Stat. Plann. Inference *139* (2009), 4072–4087. zbl MR doi

[9] *A. Di Crescenzo, M. Longobardi*: Stochastic comparisons of cumulative entropies. Stochastic Orders in Reliability and Risk (H. Li, X. Li, eds.). Lecture Notes in Statistics 208, Springer, New York, 2013, pp. 167–182. zbl MR doi

[10] *A. Di Crescenzo, M. Longobardi*: Some properties and applications of cumulative Kullback-Leibler information. Appl. Stoch. Models Bus. Ind. *31* (2015), 875–891. MR doi

[11] *N. Ebrahimi, E. S. Soofi, R. Soyer*: Information measures in perspective. Int. Stat. Rev. *78* (2010), 383–412. doi

[12] *D. A. S. Fraser*: On information in statistics. Ann. Math. Stat. *36* (1965), 890–896. zbl MR doi

[13] *A. Ghosh, C. Kundu*: Bivariate extension of (dynamic) cumulative residual and past inaccuracy measures. To appear in Stat. Pap. doi

[14] *A. Ghosh, C. Kundu*: Chernoff distance for conditionally specified models. To appear in Stat. Pap. doi

[15] *A. Ghosh, C. Kundu*: On some dynamic generalized measures of information for conditionally specified models in past life. Statistics *51* (2017), 1398–1418. zbl MR doi

[16] *E. J. Gumbel*: Bivariate logistic distributions. J. Am. Stat. Assoc. *56* (1961), 335–349. zbl MR doi

[17] *D. Jurafsky, J. H. Martin*: Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice-Hall, Englewood Cliffs, 2009.

[18] *S. Kayal, S. M. Sunoj*: Generalized Kerridge's inaccuracy measure for conditionally specified models. Commun. Stat., Theory Methods *46* (2017), 8257–8268. zbl MR doi

[19] *J. T. Kent*: Robust properties of likelihood ratio tests. Biometrika *69* (1982), 19–27. zbl MR doi

[20] *J. T. Kent*: Information gain and a general measure of correlation. Biometrika *70* (1983), 163–173. zbl MR doi

[21] *D. F. Kerridge*: Inaccuracy and inference. J. R. Stat. Soc., Ser. B *23* (1961), 184–194. zbl MR

[22] *S. Kotz, N. Balakrishnan, N. L. Johnson*: Continuous Multivariate Distributions. Vol. 1: Models and Applications. Wiley, New York, 2000. zbl MR doi

[23] *S. Kullback, R. A. Leibler*: On information and sufficiency. Ann. Math. Stat. *22* (1951), 79–86. zbl MR doi

[24] *C. Kundu, A. Di Crescenzo, M. Longobardi*: On cumulative residual (past) inaccuracy for truncated random variables. Metrika *79* (2016), 335–356. zbl MR doi

[25] *A. Kundu, C. Kundu*: Bivariate extension of (dynamic) cumulative past entropy. Commun. Stat., Theory Methods *46* (2017), 4163–4180. zbl MR doi

[26] *A. Kundu, C. Kundu*: Bivariate extension of generalized cumulative past entropy. Commun. Stat., Theory Methods *47* (2018), 1962–1977. MR doi

[27] *A. Kundu, A. K. Nanda*: On study of dynamic survival and cumulative past entropies. Commun. Stat., Theory Methods *45* (2016), 104–122. zbl MR doi

[28] *J. F. Lawless*: Statistical Models and Methods for Lifetime Data. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York, 1982. zbl MR

[29] *N. U. Nair, G. Asha*: Some characterizations based on bivariate reversed mean residual life. ProbStat Forum *1* (2008), 1–14. zbl

[30] *P. Nath*: Inaccuracy and coding theory. Metrika *13* (1968), 123–135. zbl MR doi

[31] *J. Navarro, Y. del Aguila, M. Asadi*: Some new results on the cumulative residual entropy. J. Stat. Plann. Inference *140* (2010), 310–322. zbl MR doi

[32] *J. Navarro, S. M. Sunoj, N. N. Linu*: Characterizations of bivariate models using some dynamic conditional information divergence measures. Commun. Stat., Theory Methods *43* (2014), 1939–1948. zbl MR doi

[33] *S. Park, I. Kim*: On cumulative residual entropy of order statistics. Stat. Probab. Lett. *94* (2014), 170–175. zbl MR doi

[34] *S. Park, M. Rao, D. W. Shin*: On cumulative residual Kullback-Leibler information. Stat. Probab. Lett. *82* (2012), 2025–2032. zbl MR doi

[35] *G. Psarrakos, A. Toomaj*: On the generalized cumulative residual entropy with applications in actuarial science. J. Comput. Appl. Math. *309* (2017), 186–199. zbl MR doi

[36] *M. Rao, Y. Chen, B. C. Vemuri, F. Wang*: Cumulative residual entropy: a new measure of information. IEEE Trans. Inf. Theory *50* (2004), 1220–1228. zbl MR doi

[37] *A. Rényi*: On measures of entropy and information. Proc. 4th Berkeley Symp. Math. Stat. Probab. 1. Univ. California Press, Berkeley, 1961, pp. 547–561. zbl MR

[38] *D. Roy*: A characterization of model approach for generating bivariate life distributions using reversed hazard rates. J. Jpn. Stat. Soc. *32* (2002), 239–245. zbl MR doi

[39] *P. G. Sankaran, D. Kundu*: A bivariate Pareto model. Statistics *48* (2014), 241–255. zbl MR doi
[40] *M. Shaked, J. G. Shanthikumar*: Stochastic Orders. Springer Series in Statistics, Springer, New York, 2007. zbl MR doi
[41] *C. E. Shannon*: A mathematical theory of communication. Bell Syst. Tech. J. *27* (1948), 379–423, 623–656. zbl MR doi
[42] *J. Shi, Y. Cai, J. Zhu, J. Zhong, F. Wang*: SEMG-based hand motion recognition using cumulative residual entropy and extreme learning machine. Medical & Biological Engineering & Computing *51* (2013), 417–427. doi
[43] *S. M. Sunoj, M. N. Linu*: Dynamic cumulative residual Rényi's entropy. Statistics *46* (2012), 41–56. zbl MR doi
[44] *A. Toomaj, S. M. Sunoj, J. Navarro*: Some properties of the cumulative residual entropy of coherent and mixed systems. J. Appl. Probab. *54* (2017), 379–393. MR doi
[45] *F. Wang, B. C. Vemuri, M. Rao, Y. Chen*: A new & robust information theoretic measure and its application to image alignment. Information Processing in Medical Imaging (C. Taylor, J. A. Noble, eds.). Lecture Notes in Computer Science 2732, Springer, Berlin, 2003, pp. 388–400. doi

Authors' address: *Amit Ghosh*, *Chanchal Kundu* (corresponding author), Department of Mathematics, Rajiv Gandhi Institute of Petroleum Technology, Jais 229 304, Uttar Pradesh, India, e-mail: `ckundu@rgipt.ac.in`, `chanchal_kundu@yahoo.com`.