THE CONDITIONAL DYNAMIC CUMULATIVE RESIDUAL ENTROPY

G. Rajesh*, E. I. Abdul-Sathar**, K. V. Reshmi**
and K. R. Muraleedharan Nair***

The Cumulative Residual Entropy (CRE), introduced by Rao et al. (2004), is viewed as a dynamic measure of uncertainty. Recently Asadi and Zohrevand (2007) proposed a dynamic form for the CRE, namely Dynamic Cumulative Residual Entropy (DCRE), and has discussed some of its properties. In this paper, we look into the problem of extending this concept to the conditionally specified models and study various properties of the new measures. We also propose nonparametric estimation for the new measures defined and performance of the estimators are compared using a simulation study.

Key words and phrases: Characterization, Cumulative residual entropy, Mean residual life function, Residual life.

1. Introduction

Let $X$ be a non-negative random variable having absolutely continuous cumulative distribution function $F(x)$ with probability density function $f(x)$. Shannon (1948) introduced a measure of uncertainty associated with the distribution function $F$, given by

$$H(f) = -\int_{0}^{\infty} f(x) \log f(x)dx.$$  

(1.1)

If we consider $X$ as the lifetime of a new unit, $H(f)$ can be viewed as a useful tool for measuring the associated uncertainty. Observing that if a unit has survived up to an age $t$, $H(f)$ is not capable of measuring the uncertainty about remaining life. Thus Ebrahimi and Pellerey (1995) and Ebrahimi (1996) proposed the concept of residual entropy. For a non-negative random variable $X$, representing the lifetime of a component, the residual entropy function is the Shannon’s entropy associated with the random variable $X$ given $X > t$, and is defined as

$$H(f, t) = -\int_{t}^{\infty} \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)}dx,$$

(1.2)

where $\overline{F}(t) = P(X > t)$ is the survival or reliability function of $X$.

Let $X$ and $Y$ be absolutely continuous non-negative random variables that describe the lifetimes of two items. We denote by $f(t), F(t)$ and $\overline{F}(t) = 1 - F(t)$,
the probability density function, distribution function and survival function of $X$, respectively and $g(t)$, $G(t)$ and $G(t)$ the corresponding functions of $Y$.

As an information distance between $F$ and $G$, Kullback and Leibler (1951) proposed the following discrimination measure, also known as relative entropy of $X$ and $Y$,

$$K(f, g) = \int_0^\infty f(x) \log \frac{f(x)}{g(x)} dx$$

More recently, Ebrahimi and Kirmani (1996) have defined the Kullback-Leibler distribution information of $X$ and $Y$ at time $t$ by

$$K(f, g; t) = \int_t^\infty \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} \frac{g(x)}{G(t)} dx, \quad t > 0$$

(1.3) is useful to compare the residual lifetimes of two items that have both survived up to time $t$.

Rao et al. (2004) introduced a new measure of information namely Cumulative Residual Entropy (CRE) that extends the Shannon entropy to continuous random variables and is defined as

$$\mathcal{E}(X) = -\int_0^\infty F(x) \log F(x) dx.$$  (1.4)

The basic idea in this definition is to replace the density function with the cumulative distribution in Shannon’s definition of entropy. The differential entropy doesn’t have the following important properties of CRE (see, Rao et al. (2004)). 1) CRE has consistent definitions in both the continuous and discrete domains; 2) CRE is always non-negative; 3) CRE can be easily computed from sample data and these computations asymptotically converge to the true values. 4) The conditional CRE of $X_1$ given $X_2$ is zero, if and only if $X_1$ is a function of $X_2$.

The distribution function is more regular than the density function, because the density is computed as the derivative of the distribution. Moreover, in practice what is of interest and/or measurable is the distribution function. For example, if the random variable 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 $t$.

Clearly $\mathcal{E}(X)$ measures the uncertainty contained in the cdf of $X$. Applications of $\mathcal{E}(X)$, and cross cumulative residual entropy related measure to image alignment and to measurements of similarity between images can be found in Wang and Vemuri (2007) and Wang et al. (2003). Further, Asadi and Zohrevand (2007) introduced the CRE for the residual life time distribution, called the dynamic cumulative residual entropy (DCRE), defined as

$$\mathcal{E}(X; t) = -\int_t^\infty \frac{F(x)}{F(t)} \log \frac{F(x)}{F(t)} dx.$$  (1.5)
It is clear that $\mathcal{E}(X;0) = \mathcal{E}(X)$. (1.5) can be rewritten as

$$(1.6) \quad \mathcal{E}(X; t) = - \int_{t}^{\infty} \frac{F(x)}{F(t)} \log F(x) dx + r(t) \log F(t),$$

where $r(t) = E(X - t \mid X \geq t)$ is the mean residual life function (MRLF) of the random variable $X$. Asadi and Zohrevand (2007) proved that

$$\mathcal{E}(X; t) = E(r(t) \mid X > t).$$

In studying the reliability aspects of multi-component system with each component having a lifetime depending on the lifetimes of the other, multivariate life distributions are employed. Reliability characteristics in the univariate case extends to the corresponding multivariate version. Even though a lot of interest has been evoked on the CRE of residual entropy in the univariate case, only few works seem to have been done in higher dimensions. For details see Rajesh and Nair (2000), Nadarajah and Zografos (2005), Nanda and Paul (2005), Ebrahimi et al. (2007), Sathar et al. (2009, 2010) and Rajesh et al. (2009, 2014).

The present paper addresses the question of extending the definition of DCRE to bivariate setup and study its properties. The dynamic cumulative entropy of conditional distributions has been recently introduced by Rajesh et al. (2014) and studied various properties.

In order to introduce conditional, joint and cross dynamic cumulative entropies, we make use of the definition of dynamic cumulative entropy of conditional distributions in this paper.

The rest of the paper is organized as follows. In Section 2, we extend DCRE for the conditionally specified models. In Section 3, we discussed the properties of the new measures defined. In Section 4, we discussed nonparametric estimates of the new measure introduced. In Section 5, we studied the performance of the estimators using a simulation study.

2. Bivariate dynamic cumulative residual entropy

Identifying the joint distribution of $(X_1, X_2)$ when conditional distribution of $(X_1 \mid X_2 = x_2)$ and $(X_2 \mid X_1 = x_1)$ are known has been an important problem dealt with by many researchers in the past. This approach of identifying a bivariate density using the conditionals is called the conditional specification of the joint distribution. These conditional models are often useful in many two component reliability systems when the operational status of one component is known.

Ebrahimi et al. (2007) developed measures of information for multivariate distributions when their supports are truncated. The key objectives of these study are to introduce measures for assessing whether a distribution becomes more or less informative about the prediction of remaining life times using entropies of joint, marginal and conditional residual life distributions. Also, the mutual information of the residual life distributions is used to assess the dependency of the components of a system as they age.
The present paper provides another way in which the information about the predictability of the remaining life time and dependency can be studied using joint, marginal and conditional cumulative residual entropies. These measures also find applications in solving the uni and multimodal image alignment problem for parameterized (rigid, affine and projective) transformation. The key strength of the cross cumulative residual entropy over using the mutual information based on Shannon’s entropy are that the former has significantly large noise immunity and a much larger convergence range over the field of parameterized transformation (see, Wang et al. (2003)).

For a random variable vector $X = (X_1, X_2)$ represent the lifetime of the components in a two component system with joint survival function $F(x_1, x_2)$, marginal survival function $F_i(x_i)$, $i = 1, 2$. It is of special interest to consider the DCRE of conditional distributions namely, the conditional distribution of $X_i \mid X_j > t_j$ and $X_i \mid X_j = t_j$, $i, j = 1, 2$, $i \neq j$.

The dynamic cumulative residual entropy for the random variable $X_i$ given $X_j > t_j$, $i \neq j$ is defined as

$$
\mathcal{E}(X_i; t_1, t_2) = -\int_{t_i}^{\infty} \frac{F(x_i \mid X_j > t_j)}{F(t_1 \mid X_j > t_j)} \log \frac{F(x_i \mid X_j > t_j)}{F(t_1 \mid X_j > t_j)} dx_i
$$

where $F(x_i \mid X_j > t_j) = P(X_i > x_i \mid X_j > t_j)$.

Differentiating (2.1) with respect to $t_i$, $i = 1, 2$, we get

$$
\frac{\partial}{\partial t_i} \mathcal{E}(X_i; t_1, t_2) = h_i(t_1, t_2)[\mathcal{E}(X_i; t_1, t_2) - r_i(t_1, t_2)],
$$

where $h_i(t_1, t_2) = -\frac{\partial}{\partial t_i} \log F(t_1, t_2)$ is the $i$th component of the vector valued failure rate and

$$
\frac{\partial}{\partial t_i} \mathcal{E}(X_i; t_1, t_2) = \int_{t_i}^{\infty} \frac{F(x_i \mid t_j)}{F(t_1, t_2)} dx_i
$$

is the $i$th component of the vector valued mean residual life.

Analogous to the conditional residual entropy, defined by Ebrahimi et al. (2007), comparing $\mathcal{E}(X_i; t_1, t_2)$ with the marginal cumulative residual entropy $\mathcal{E}(X_i; t_1)$ gives an idea about the residual life of the component $i$ when we considering $t_1$ and $t_2$ against $t_i$, $i = 1, 2$. That is, when we consider both the ages, that may cause a reduction or excess in the uncertainty in the residual life of the components. If $X_i$ is the right tail decreasing dependence in $X_j$, $i \neq j$ and if $F(x_i)$, $i = 1, 2$ have decreasing mean residual life function, then $\mathcal{E}(X_i; t_1, t_2) \leq \mathcal{E}(X_i; t_i)$. Hence, right tail decreasing dependence and marginal decreasing mean residual life are sufficient condition for the uncertainty reduction about the residual life of one component due to the consideration of the ages of the both components.
The dynamic cumulative residual entropy for the random variable $X_i$ given $X_j = x_j$ defined as

$$
(2.4) \quad \varepsilon(X_i \mid x_j; t_1, t_2) = -\int_{t_1}^{\infty} \frac{\overline{F}(x_i \mid X_j = x_j)}{\overline{F}(t_1 \mid X_j = x_j)} \log \frac{\overline{F}(x_i \mid X_j = x_j)}{\overline{F}(t_1 \mid X_j = x_j)} dx_i,
$$

where $\overline{F}(x_i \mid X_j = x_j) = P(X_i > x_i \mid X_j = x_j)$.

Differentiating (2.4) with respect to $t_i$, $i = 1, 2$, we get

$$
(2.5) \quad \frac{\partial}{\partial t_i} \varepsilon(X_i \mid x_j; t_1, t_2) = h_i(t_i \mid x_j)[\varepsilon(X_i \mid x_j; t_1, t_2) - r_i(t_i \mid x_j)],
$$

where $h_i(t_i \mid x_j) = -\frac{\partial}{\partial t_i} \log \overline{F}(t_i \mid x_j)$ and $r_i(t_i \mid x_j) = \int_{t_i}^{\infty} \frac{\overline{F}(x_i \mid x_j)}{\overline{F}(t_i \mid x_j)} dx_i$ are the conditional failure rate and mean residual life function defined for the random variable $X_i \mid X_j > t_i$, $X_j = x_j$.

The conditional dynamic cumulative entropy function is defined as

$$
(2.6) \quad \varepsilon(X_i \mid X_j; t_1, t_2) = \int_{t_1}^{\infty} \varepsilon(X_i \mid x_j; t_1, t_2) f_j(x_j; t_1, t_2) dx_j,
$$

where $f_j(x_j; t_1, t_2)$ is the marginal residual density for the random variable $X_j \mid X_1 > t_1, X_2 > t_2$. For $i = 1, j = 2$,

$$
(2.7) \quad \varepsilon(X_1 \mid X_2; t_1, t_2) = \int_{t_2}^{\infty} \int_{t_1}^{\infty} \frac{\partial}{\partial x_2} \frac{F(x_1, x_2)}{F(t_1, t_2)} \log \left( \frac{\partial}{\partial x_2} \frac{F(x_1, x_2)}{F(t_1, t_2)} \right) dx_1 dx_2.
$$

The joint cumulative residual entropy can be decomposed as

$$
(2.8) \quad \varepsilon(X_1, X_2; t_1, t_2) = \varepsilon(X_1; t_1, t_2) + \varepsilon(X_2 \mid X_1; t_1, t_2).
$$

It is to be noted that

$$
(2.9) \quad \varepsilon(X_1, X_2; t_1, t_2) \neq \varepsilon(X_2, X_1; t_1, t_2).
$$

Now the cumulative entropy difference, is defined as

$$
(2.10) \quad \tau(X_i \mid x_j; t_1, t_2) = \varepsilon(X_i; t_1, t_2) - \varepsilon(X_i \mid x_j; t_1, t_2).
$$

(2.10) can be viewed as the Kullback-Leibler discrimination function, denoted as $K[\overline{F}_{i,j}(x_i \mid x_j; t_1, t_2); \overline{F}_i(x_i; t_1, t_2)]$ where $\overline{F}_{i,j}(x_i \mid x_j; t_1, t_2)$ and $\overline{F}_i(x_i; t_1, t_2)$ are the conditional survival functions of the random variables $(X_i \mid X_i > x_i, X_j = x_j)$ and $(X_i \mid X_i > x_i, X_j > x_j)$ respectively. (2.10) can be negative or positive depending upon whether $\overline{F}_{i,j}(x_i \mid x_j; t_1, t_2)$ or $\overline{F}_i(x_i; t_1, t_2)$ is more informative.

Now we define the cross cumulative residual entropy (CCRE) (analogous to residual mutual information (Ebrahimi et al. (2007))), is defined as

$$
(2.11) \quad CCRE(X_i, X_j; t_1, t_2)
$$
where \( K[F_{ij}(x_i \mid x_j; t_1, t_2) : F_i(x_i; t_1, t_2)] \) is the Kullback measure defined in (1.3).

The Kullback-Leibler and entropy representation of two residual lifetimes, given by

\[
\begin{align*}
CCRE(X_i, X_j; t_1, t_2) &= K[P(X_i > x_i, X_j = x_j \mid X_1 > t_1, X_2 > t_2) : \\
&\quad f_j(x_j; t_1, t_2)F_i(x_i; t_1, t_2)] \\
&= \mathcal{E}(X_i; t_1, t_2) + \mathcal{E}(X_j; t_1, t_2) - \mathcal{E}(X_i, X_j; t_1, t_2) \\
&= \mathcal{E}(X_i; t_1, t_2) - \mathcal{E}(X_i \mid X_j; t_1, t_2).
\end{align*}
\]

By (2.12), \( CCRE(X_i, X_j; t_1, t_2) \geq 0 \) and it can be verified that

\[
CCRE(X_1, X_2; 0, 0) = CCRE(X_1, X_2)
\]

where \( CCRE(X_1, X_2) = \mathcal{E}(X_1) - E[\mathcal{E}(X_1 \mid X_2)] \) is the cross cumulative residual entropy of \((X_1, X_2)\).

For any bivariate distributions, the residual mutual information is symmetric with respect to its coordinates. But CCRE is not symmetric. That is \( CCRE(X_1, X_2; t_1, t_2) \neq CCRE(X_2, X_1; t_1, t_2) \). In order to derive a symmetric version of CCRE we define

\[
CCRE(X_1, X_2; t_1, t_2) = \frac{1}{2} [CCRE(X_1, X_2; t_1, t_2) + CCRE(X_2, X_1; t_1, t_2)].
\]

It is clear that \( SCCRE(X_1, X_2; t_1, t_2) = SCCRE(X_2, X_1; t_1, t_2) \). A summarization of formulas of these measures are provided in Table 1. From (2.13) and (2.14), we have

\[
\mathcal{E}(X_i, X_j; t_1, t_2) \leq \mathcal{E}(X_i; t_1, t_2) + \mathcal{E}(X_j; t_1, t_2), \quad i \neq j
\]

and

\[
\mathcal{E}(X_i \mid X_j; t_1, t_2) \leq \mathcal{E}(X_i; t_1, t_2), \quad i \neq j.
\]

The equality holds iff \( X_1 \) and \( X_2 \) are independent. Clearly SCCRE is non-negative. We will however use the nonsymmetric CCRE in our study, as it is sufficient to yield the desired results. In the following examples we compute the various dynamic CRE measures defined above, using some bivariate distributions.

**Example 1.** Consider the following bivariate distribution with the survival function given by

\[
\mathcal{F}(x_1, x_2) = (1 + x_1 + x_2)^{-\alpha}, \quad x_1, x_2 \geq 0, \quad \alpha > 0
\]
Table 1. Measures of information.

| Sl. No. | Measures                  | Formula                                                                                                                                 |
|---------|---------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| 1       | Shannon’s entropy         | $H(f) = -\int_{0}^{\infty} f(x) \log f(x) dx$                                                                                           |
| 2       | Residual entropy          | $H(f, t) = -\int_{t}^{\infty} \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} dx$                                                            |
| 3       | CRE                      | $\mathcal{E}(X) = -\int_{0}^{\infty} \hat{F}(x) \log \hat{F}(x) dx$                                                                    |
| 4       | DCRE                     | $\mathcal{E}(X; t) = -\int_{t}^{\infty} \frac{\hat{F}(x)}{F(t)} \log \frac{\hat{F}(x)}{F(t)} dx$                                  |
| 5       | Marginal CRE’s           | $\mathcal{E}(X_i; t_1, t_2) = -\int_{t_1}^{\infty} \frac{\hat{F}(x_i|X_j=x_j)}{F(t_1,t_2)} \log \frac{\hat{F}(x_i|X_j=x_j)}{F(t_1,t_2)} dx_i$ and $\mathcal{E}(X_i|X_j=x_j; t_1, t_2) = -\int_{t_1}^{\infty} \frac{\hat{F}(x_i|X_j=x_j)}{F(t_1|X_j=x_j)} \log \frac{\hat{F}(x_i|X_j=x_j)}{F(t_1|X_j=x_j)} dx_i$ |
| 6       | Conditional CRE’s        | $\mathcal{E}(X_i | X_j; t_1, t_2) = \int_{t_1}^{\infty} \mathcal{E}(X_i | x_j; t_1, t_2) f_j(x_j; t_1, t_2) dx_j$                             |
| 7       | Joint CRE’s              | $\mathcal{E}(X_1, X_2; t_1, t_2) = \mathcal{E}(X_1; t_1, t_2) + \mathcal{E}(X_2 | X_1; t_1, t_2)$                                       |
| 8       | Kullback function        | $K(f, g) = \int_{0}^{\infty} f(x) \log \frac{f(x)}{g(x)} dx$                                                                           |
| 9       | CCRE                     | $CCRE(X_i, X_j; t_1, t_2) = \int_{t_1}^{\infty} K[\hat{F}_i(x_i | x_j; t_1, t_2) : \hat{F}_j(x_j; t_1, t_2)] f_j(x_j; t_1, t_2) dx_i$ |
and the corresponding density function is

\[ f(x_1, x_2) = \alpha(\alpha + 1)(1 + x_1 + x_2)^{-(\alpha + 2)}, \quad x_1, x_2 \geq 0, \quad \alpha > 0. \]  

(2.19)

Using (2.1), (2.4), (2.6), (2.8) and (2.14), we get

\[ \mathcal{E}(X_i; t_1, t_2) = \frac{\alpha(1 + t_1 + t_2)}{\alpha - 1}, \quad i = 1, 2 \]  

(2.20)

\[ \mathcal{E}(X_i \mid t; t_1, t_2) = \frac{\alpha + 1}{\alpha^2}(1 + t_1 + t_2), \quad i \neq j = 1, 2 \]  

(2.21)

\[ \mathcal{E}(X_i \mid X_j; t_1, t_2) = \frac{\alpha + 1}{\alpha(\alpha - 1)}(1 + t_1 + t_2), \quad i \neq j = 1, 2 \]  

(2.22)

\[ \mathcal{E}(X_1, X_2; t_1, t_2) = \frac{(2\alpha^2 - 1)}{\alpha(\alpha - 1)^2}(1 + t_1 + t_2) \]

\[ = \mathcal{E}(X_2, X_1; t_1, t_2), \quad i \neq j = 1, 2 \]  

(2.23)

\[ CRE(X_i, X_j; t_1, t_2) = \frac{1}{\alpha(\alpha - 1)^2}(1 + t_1 + t_2)(1 - \alpha - \alpha^2) \]  

(2.24)

It can be shown that all cumulative residual entropies are increasing in \( t_1 \) and \( t_2 \).

\[ Example \ 2. \] Consider the bivariate distribution with the following density as the unit square

\[ f(x_1, x_2) = x_1 + x_2, \quad 0 \leq x_1, x_2 \leq 1 \]

The survival function is

\[ F(t_1, t_2) = \frac{1}{2}(1 - t_1)(1 - t_2)(2 + t_1 + t_2). \]  

(2.25)

The marginal CRE for the random variable \( X_1 \) given \( X_2 > t_2 \) for (2.25) simplifies to

\[ \mathcal{E}(X_1; t_1, t_2) = \frac{1}{18(t_1 - 1)(2 + t_1 + t_2)} \times [ -41 + 60 \log(3 + t_2) - 4t_1^3 - 6t_1^2(1 + t_2) \]

\[ - 3(2 + t_2^2)(5 + t_2) \log(2 + t_1 + t_2) + 3(7 + 3t_2) \log(1 - t_1) + 3t_1(17 + t_2(10 + t_2)) \]

\[ + 3t_2 - 8 + 24 \log(3 + t_2) + t_2(-1 + 9 \log(3 + t_2) + t_2 \log(3 + t_2)]]. \]

(2.26)

The expressions of \( \mathcal{E}(X_i \mid X_j; t_1, t_2) \), \( \mathcal{E}(X_i, X_j; t_1, t_2) \) and \( CRE(X_i, X_j; t_1, t_2) \) computed using (2.25) are more complicated. But it can be shown that all cumulative residual entropies are increasing in \( t_1 \) and \( t_2 \) and the rate of increase is highest on the diagonal \((t, t)\).
Figure 1 shows the graphs of three cumulative residual entropies. The joint cumulative residual entropy $\mathcal{E}(X_1, X_2; t, t)$ is shown by the solid curve and the marginal cumulative residual entropy $\mathcal{E}(X_1; t, t)$ by the dashed curve and the conditional cumulative residual entropy $\mathcal{E}(X_1 \mid X_2; t, t)$ by the dotted curve.

The CCRE decreases in $t_1$ and $t_2$. Thus by (2.13) the joint CRE increases with a faster rate than the total CRE’s and by (2.14), the conditional CRE decreases with a faster rate than the marginal residual entropy. Figure 2 shows the graph of $CCRE(X_1, X_2; t, t)$. 
3. Properties

In this section we study some properties of the measures defined in Section 2. We first look into the situation where the dynamic cumulative residual entropy (DCRE) coincides with the conditional dynamic residual entropy.

**Theorem 3.1.** *The conditional cumulative residual entropy*

\[
\varepsilon(X_i \mid X_j; t_1, t_2) = \varepsilon(X_i; t_1, t_2), \quad i, j = 1, 2, \ i \neq j
\]

or

\[
CCRE(X_i, X_j; t_1, t_2) = 0, \quad i, j = 1, 2, \ i \neq j
\]

holds if and only if \(X_1\) and \(X_2\) are independent.

**Proof.** When (3.1) holds, we have for \(i = 1, j = 2\)

\[
\int_{t_2}^{\infty} \int_{t_1}^{\infty} \frac{\partial}{\partial x_2} \frac{F(x_1, x_2)}{F(t_1, t_2)} \log \left( \frac{\partial}{\partial x_2} \frac{F(x_1, x_2)}{F(t_1, t_2)} \right) dx_1 dx_2
\]

\[
= - \int_{t_1}^{\infty} \frac{F(x_1, t_2)}{F(t_1, t_2)} \log \frac{F(x_1, t_2)}{F(t_1, t_2)} dx_1.
\]

Differentiating (3.2) both sides with respect to \(t_2\), we get

\[
\frac{\partial}{\partial t_2} \log \frac{F(x_1, t_2)}{F(t_1, t_2)} \int_{t_1}^{\infty} \frac{\partial}{\partial t_2} \frac{F(x_1, t_2)}{F(t_1, t_2)} dx_1
\]

\[
+ \int_{t_1}^{\infty} \frac{\partial}{\partial t_2} F(x_1, t_2) \log \frac{\partial}{\partial t_2} F(t_1, t_2) dx_1
\]

\[
= - \int_{t_1}^{\infty} \left[ \frac{F(x_1, t_2)}{F(t_1, t_2)} \frac{\partial}{\partial t_2} \log F(x_1, t_2) + \frac{\partial}{\partial t_2} F(x_1, t_2) \log F(x_1, t_2) \right] dx_1
\]

\[
+ \log F(t_1, t_2) \int_{t_1}^{\infty} \frac{\partial}{\partial t_2} F(x_1, t_2) dx_1
\]

\[
+ \frac{\partial}{\partial t_2} \log F(t_1, t_2) \int_{t_1}^{\infty} F(x_1, t_2) dx_1.
\]

Differentiating (3.3), with respect to \(t_1\) on both sides, and simplifying, we get,

\[
\frac{\partial}{\partial t_1} \log \frac{\partial}{\partial t_2} F(t_1, t_2) \int_{t_1}^{\infty} \frac{\partial}{\partial t_2} F(x_1, t_2) dx_1
\]

\[
= \frac{\partial}{\partial t_1} \log F(t_1, t_2) \int_{t_1}^{\infty} \frac{\partial}{\partial t_2} F(x_1, t_2) dx_1
\]

\[
+ \frac{\partial^2}{\partial t_1 \partial t_2} \log F(t_1, t_2) \int_{t_1}^{\infty} F(x_1, t_2) dx_1.
\]
which is equivalent to

$$\frac{\partial}{\partial t_2} \log \left[ \int_{t_1}^{\infty} \frac{F(x_1, t_2) dx_1}{F(t_1, t_2)} \right] = 0$$

or

$$\frac{\partial}{\partial t_2} \log r_1(t_1, t_2) = 0$$

or

$$\frac{1}{r_1(t_1, t_2)} \frac{\partial}{\partial t_2} r_1(t_1, t_2) = 0.$$ 

That is

(3.4) $$\frac{\partial}{\partial t_2} r_1(t_1, t_2) = 0.$$ 

This means $r_1(t_1, t_2)$ is independent of $t_2$ or

(3.5) $$r_1(t_1, t_2) = r_1(t_1, 0).$$

Differentiating (3.5) with respect to $t_1$ on both sides, we get

$$\frac{\partial}{\partial t_1} r_1(t_1, t_2) = \frac{\partial}{\partial t_1} r_1(t_1, 0).$$

Using the relationship between vector valued hazard rate and mean residual life function, we get

$$r_1(t_1, t_2) h_1(t_1, t_2) = r_1(t_1, 0) h_1(t_1, 0).$$

Therefore

$$h_1(t_1, t_2) = h_1(t_1, 0).$$

or

$$-\frac{\partial}{\partial t_1} \log F(t_1, t_2) = -\frac{\partial}{\partial t_1} \log F_1(t_1, 0),$$

where $\log \frac{F(t_1, t_2)}{F_1(t_1, 0)}$ does not depend $t_1$. Therefore

$$\log \frac{F(t_1, t_2)}{F_1(t_1, 0)} = \log \frac{F_2(0, t_2)}{F(0, 0)} = \log F_2(0, t_2).$$

or

$$\bar{F}(t_1, t_2) = \bar{F}_1(t_1, 0) \bar{F}_2(0, t_2) \quad \forall t_1, t_2.$$ 

Hence $X_1$ and $X_2$ are independent. □

The following theorem characterizes conditional cumulative residual entropy in terms of conditional mean residual life function.
**Theorem 3.2.** The conditional cumulative residual entropies

\( \mathcal{E}(X_1 \mid X_2; t_1, t_2) = \mathcal{E}(X_2 \mid X_1; t_2, t_1) \)

if and only if

\[ r_1(t_1 \mid t_2) = r(t_2 \mid t_1), \]

where

\[ r_1(t_1 \mid t_2) = E(X_1 - t_1 \mid X_1 > t_1, X_2 = t_2) \]

and

\[ r(t_2 \mid t_1) = E(X_2 - t_2 \mid X_1 = t_1, X_2 > t_2). \]

**Proof.** When (3.6) holds, we have

\[
\int_{t_1}^{\infty} \int_{t_2}^{\infty} \frac{\partial}{\partial x_2} \frac{F(x_1, x_2)}{F(t_1, t_2)} \log \left( \frac{\partial}{\partial x_2} \frac{F(x_1, x_2)}{F(t_1, t_2)} \right) dx_1 dx_2
\]

\[
= \int_{t_2}^{\infty} \int_{t_1}^{\infty} \frac{\partial}{\partial x_1} \frac{F(x_1, x_2)}{F(t_1, t_2)} \log \left( \frac{\partial}{\partial x_1} \frac{F(x_1, x_2)}{F(t_1, t_2)} \right) dx_1 dx_2.
\]

Differentiating with respect to \( t_2 \) on both sides, we get

\[
\frac{\partial}{\partial t_1} \log \left( \frac{\partial}{\partial t_2} F(t_1, t_2) \right) \int_{t_1}^{\infty} \frac{\partial}{\partial t_2} F(x_1, t_2) dx_1
\]

\[
= \frac{\partial}{\partial t_2} \log \left( \frac{\partial}{\partial t_1} F(t_1, t_2) \right) \int_{t_2}^{\infty} \frac{\partial}{\partial t_1} F(t_1, x_2) dx_2.
\]

Differentiating with respect to \( t_1 \) on both sides, we get

\[
\frac{\partial}{\partial t_1} F(t_1, t_2) = \int_{t_1}^{\infty} \frac{\partial}{\partial t_1} F(x_1, t_2) dx_2
\]

\[
\frac{\partial}{\partial t_2} F(t_1, t_2) = \int_{t_2}^{\infty} \frac{\partial}{\partial t_2} F(x_1, t_2) dx_1.
\]

The above equation can be written as,

\[
\frac{1}{F(t_1 \mid t_2)} \int_{t_1}^{\infty} F(x_1 \mid t_2) dx_1 = \frac{1}{F(t_2 \mid t_1)} \int_{t_2}^{\infty} F(x_2 \mid t_1) dx_2,
\]

where

\[ F(t_1 \mid t_2) = P(X_1 > t_1 \mid X_2 = t_2). \]
This gives

\[ r_1(t_1 \mid t_2) = r(t_2 \mid t_1). \]

Like other reliability measures as failure rate, mean residual life function, residual entropy function, CRE is said to be memoryless if it is free from the ages of the components. In the univariate case, memoryless CRE characterizes the exponential distribution. In the following theorem, bivariate independent exponential distribution is characterized by memoryless conditional CRE.

**Theorem 3.3.** The conditional cumulative residual entropies \( \mathcal{E}(X_i \mid X_j; t_1, t_2), i, j = 1, 2, i \neq j \) are constants free from \( t_1, t_2 \) if and only if \( X_1 \) and \( X_2 \) are independently exponentially distributed. That is,

\[
\overline{F}(x_1, x_2) = e^{-\lambda_1 x_1 - \lambda_2 x_2}, \quad x_i > 0, \; \lambda_i > 0, \; i = 1, 2.
\]

**Proof.** To prove sufficiency part, by (3.6) for two independent random variables \( X_1 \) and \( X_2 \),

\[
\mathcal{E}(X_i \mid X_j; t_1, t_2) = \mathcal{E}(X_i; t_1, t_2).
\]

Using the result of Ebrahimi (1996), dynamic cumulative entropy of an exponentially distributed random variable being constant independent of time.

The proof of the necessary part follows from the fact that

\[
\mathcal{E}(X_1 \mid X_2; t_1, t_2) = \int_{t_1}^{\infty} \mathcal{E}(X_1 \mid x_2; t_1, t_2) f_2(x_2; t_1, t_2) dx_2
\]

\[
= \int_{t_2}^{\infty} \int_{t_1}^{\infty} \frac{\partial}{\partial x_2} \overline{F}(x_1, x_2) \log \left[ \frac{\partial}{\partial x_2} \overline{F}(x_1, x_2) \right] dx_1 dx_2.
\]

Differentiating with respect to \( t_2 \) on both sides we get

\[
\frac{\partial}{\partial x_2} \mathcal{E}(X_1 \mid X_2; t_1, t_2) = \frac{\partial}{\partial t_2} \overline{F}(t_1, t_2) \left[ \mathcal{E}(X_1 \mid x_2; t_1, t_2) - \mathcal{E}(X_1 \mid X_2; t_1, t_2) \right] = 0
\]

Hence \( \mathcal{E}(X_1 \mid X_2; t_1, t_2) = \mathcal{E}(X_1 \mid x_2; t_1, t_2) = k_1 \), a constant free from \( t_1, t_2 \geq 0 \). It follows from Ebrahimi (1996) that for some \( \lambda_1 > 0 \),

\[
\overline{F}_{1|2}(x_1 \mid x_2; t_1, t_2) = e^{-\lambda_1 x_1}, \quad x_1 \geq 0, \; t_1, t_2 \geq 0.
\]

Hence

\[
\overline{F}(x_1, x_2) = e^{-\lambda_1 x_1 - \lambda_2 x_2}, \quad \lambda_1 > 0, \; \lambda_2 > 0, \; x_1, x_2 \geq 0. \]
A bivariate distribution possesses the bivariate lack of memory property if

\[ \overline{F}(s_1 + t, s_2 + t) = \overline{F}(s_1, s_2) \overline{F}(t, t). \]

(3.7)

When \( s_2 = 0 \), (3.7) becomes

\[ \overline{F}(s_1 + t, t) = \overline{F}(s_1) \overline{F}(t, t). \]

(3.8)

Taking proper steps in (3.7) and (3.8), leads to

\[ \varepsilon(X_i; t_1, t_2) = \varepsilon(X_i; t_i), \quad i = 1, 2 \]

and

\[ \varepsilon(X_1, X_2; t_1, t_2) = \varepsilon(X_1, X_2). \]

In the following theorem, we address the question of cumulative residual entropy of lifetimes of parallel and series systems of two components.

**Theorem 3.4.** Suppose that \( X_1 \) and \( X_2 \) are independent and identically distributed random variables.

(a) If \( \varepsilon(X_1, X_2; t, t) \) is increasing in \( t \), then \( \varepsilon(\min(X_1, X_2); t, t) \) is increasing in \( t \).

(b) If \( \varepsilon(X_1, X_2; t, t) \) is decreasing in \( t \), then \( \varepsilon(\max(X_1, X_2); t, t) \) is decreasing in \( t \).

**Proof.** (a) We have

\[ \varepsilon(\min(X_1, X_2); t, t) = \frac{-2}{(\overline{F}(t))^2} \int_t^\infty (\overline{F}(x))^2 \log \frac{\overline{F}(x)}{F(t)} dx \]

(3.9)

\[ = 2\varepsilon(X_j; t) - \frac{2}{(\overline{F}(t))^2} \int_t^\infty f(x) \overline{F}(x) \left( \varepsilon(X_j; x) - r(x) \log \frac{\overline{F}(x)}{F(t)} \right) dx \]

\[ \geq \varepsilon(X_j; t) \geq r(t). \]

Differentiating (3.9) with respect to \( t \) on both sides we get

\[ \frac{\partial}{\partial t} \varepsilon(\min(X_1, X_2); t, t) \]

\[ = 2h(t) \left[ \varepsilon(\min(X_1, X_2); t, t) - \frac{1}{(\overline{F}(t))^2} \int_t^\infty (\overline{F}(x))^2 dx \right] \]

\[ = 2h(t) \left[ \varepsilon(\min(X_1, X_2); t, t) - r(t) + \frac{1}{(\overline{F}(t))^2} \int_t^\infty f(x) \overline{F}(x) r(x) dx \right] \]

\[ \geq 2h(t)[\varepsilon(\min(X_1, X_2); t, t) - r(t)] \geq 0. \]
Differentiating (3.10) with respect to \( t \), we get

\[
\frac{\partial}{\partial t} \varepsilon(\max(X_1, X_2); t, t)
= \int_t^\infty \frac{1 - F^2(x)}{1 - F^2(t)} \log \frac{1 - F^2(x)}{1 - F^2(t)} \, dx
\]

\[
= \frac{\varepsilon(X_j; t)}{1 + F(t)} - \frac{r(t)}{1 + F(t)} \log F(t)
- \frac{1}{1 - F^2(t)} \int_t^\infty (1 + F(x))(1 - F(x)) \log(1 + F(x)) \, dx
- \frac{1}{1 - F^2(t)} \int_t^\infty F(x)F(x) \log F(x) \, dx
+ \frac{\log(1 - F^2(t))}{1 - F^2(t)} \int_t^\infty (1 + F(x))F(x) \, dx
\]

\[
\leq \varepsilon(X_j; t) \leq r(t), \quad \text{if } \varepsilon(X_j; t) \text{ is decreasing.} \quad \Box
\]

In the following theorem we give a sufficient condition for \( \varepsilon(X_i, X_j; t_1, t_2) \), \( i \neq j \), to be monotone in terms of the mean residual life function.

**Theorem 3.5.** If (a) \( r_i(x; t_1, t_2) \) is increasing (decreasing) in \( x_i \), \( i = 1, 2 \), and (b) \( \varepsilon(X_i | x_j; t_1, t_2) \), \( i \neq j \), \( j = 1, 2 \) is decreasing (increasing) in \( x_i \) for each \( j = 1, 2 \) then \( \varepsilon(X_j, X_i; t_1, t_2) \), \( i \neq j \), \( j = 1, 2 \) is decreasing (increasing) in \( t_1, t_2 \).
Proof. We have

\begin{equation}
\mathcal{E}(X_2, X_1; t_1, t_2) = \mathcal{E}(X_2; t_1, t_2) + \mathcal{E}(X_1 | X_2; t_1, t_2).
\end{equation}

Differentiating (3.11) with respect to \( t_2 \) on both sides, we get

\[
\frac{\partial}{\partial t_2} \mathcal{E}(X_2, X_1; t_1, t_2) = \frac{\partial}{\partial t_2} \mathcal{E}(X_2; t_1, t_2) + \frac{\partial}{\partial t_2} \mathcal{E}(X_1 | X_2; t_1, t_2)
\]

\[
= h_2(t_1, t_2) \left[ \mathcal{E}(X_2; t_1, t_2) - r_2(t_1, t_2) \right]
\]

\[
- \mathcal{E}(X_1 | x_2; t_1, t_2) - \mathcal{E}(X_1 | X_2; t_1, t_2) \]

Since \( r_2(t_1, t_2) \) is decreasing and \( \mathcal{E}(X_1 | X_2; t_1, t_2) \) is decreasing in \( t_2 \). That is \( \mathcal{E}(X_2, X_1; t_1, t_2) \) is decreasing in \( t_1, t_2 \). \( r_2(t_1, t_2) \) is decreasing, then \( \mathcal{E}(X_2; t_1, t_2) \leq r_2(t_1, t_2) \). By assumption (b) we have \( \mathcal{E}(X_1 | X_2; t_1, t_2) \leq \mathcal{E}(X_1 | x_2; t_1, t_2) \).

Hence we have \( \frac{\partial}{\partial t_j} \mathcal{E}(X_1, X_2; t_1, t_2) \leq 0 \) for \( j = 1, 2 \). This completes the proof for the increasing part. The proof for the decreasing part is similar. \( \square \)

4. Non parametric estimation

Let \( (X_i, Y_i), i = 1, 2, \ldots, n \) be \( n \) independent and identically distributed pairs of failure times with survival function \( \overline{F}(t_1, t_2) \). Based on this observation, \( \overline{F}(t_1, t_2) \) is, generally estimated by the empirical survival function \( \overline{F}_n(t_1, t_2) \), is defined as

\begin{equation}
\overline{F}_n(t_1, t_2) = \frac{1}{n} \sum_{k=1}^{n} I(X_{1_k} \geq t_1, X_{2_k} > t_2).
\end{equation}

The estimator for the dynamic cumulative residual entropy \( \mathcal{E}(X_i; t_1, t_2) \) for the random variable \( X_i | X_j > t_j \) is obtained as

\begin{equation}
\mathcal{E}_n(X_i; t_1, t_2)
\end{equation}

\[
= - \int_{t_i}^{\infty} \frac{\sum_{k=1}^{n} I(X_{i_k} > x_i, X_{j_k} > t_j)}{\sum_{k=1}^{n} I(X_{1_k} > t_1, X_{2_k} > t_2)} \times \log \left[ \frac{\sum_{k=1}^{n} I(X_{i_k} > x_i, X_{j_k} > t_j)}{\sum_{k=1}^{n} I(X_{1_k} > t_1, X_{2_k} > t_2)} \right] dx_i, \quad i \neq j, \ i, j = 1, 2.
\]

For estimating the conditional cumulative entropy function \( \mathcal{E}(X_i | X_j; t_1, t_2) \), we use the usual kernel estimator of conditional distributions. Let \( h_{jn} \), \( j = 1, 2 \) denote the bandwidth such that \( h_{jn} \to 0 \) as \( n \to \infty \). Kernel estimate of the conditional survival function of \( X_i \) given \( X_j = x_j, i \neq j = 1, 2 \), is given by

\[
\overline{F}_n(x_i | X_j = x_j) = \frac{\sum_{k=1}^{n} w \left( \frac{x_j - X_{j_k}}{h_{jn}} \right) I(X_{i_k} \geq x_i)}{w \left( \frac{x_j - X_{j_k}}{h_{jn}} \right)}, \quad i \neq j = 1, 2,
\]
where \( w(\cdot) \) denote the Kernel. The estimator of the dynamic cumulative entropy of \( X_i \) given \( X_j = t_j \), \( \mathcal{E}(X_i \mid x_j; t_1, t_2) \), is defined as

\[
\mathcal{E}_n(X_i \mid x_j; t_1, t_2) = -\int_t^\infty \sum_{k=1}^n w^{(x_j - X_{jk})} I(X_{ik} \geq x_i) \sum_{k=1}^n w^{(x_j - X_{jk})} I(X_{ik} \geq t_i) \times \log \left[ \sum_{k=1}^n w^{(x_j - X_{jk})} I(X_{ik} \geq x_i) \right] \; dx_i, \quad i \neq j.
\]

From (2.7), the estimator of conditional cumulative entropy \( \mathcal{E}_n(X_i \mid X_j; t_1, t_2) \) is obtained as

\[
\mathcal{E}_n(X_i \mid X_j; t_1, t_2) = \int_t^\infty \int_{t_1}^\infty \sum_{k=1}^n w^{(x_j - X_{jk})} I(X_{ik} \geq x_i) \sum_{k=1}^n w^{(x_j - X_{jk})} I(X_{ik} \geq x_i) \times \log \left[ \sum_{k=1}^n w^{(x_j - X_{jk})} I(X_{ik} \geq x_i) \right] \; dx_i \; dx_j, \quad i \neq j = 1, 2
\]

Finally, we obtain the estimator of the joint cumulative residual entropy functions as

\[
\mathcal{E}_n(X_1, X_2; t_1, t_2) = \mathcal{E}_n(X_1; t_1, t_2) + \mathcal{E}_n(X_2 \mid X_1; t_1, t_2)
\]

\[
\mathcal{E}_n(X_2, X_1; t_1, t_2) = \mathcal{E}_n(X_2; t_1, t_2) + \mathcal{E}_n(X_1 \mid X_2; t_1, t_2)
\]

Using Glivenko-Cantelli theorem for bivariate survival functions, we can prove the consistency and weak convergence of the estimates.

5. Simulation studies

For empirical studies on the performance of the estimators, we carried out a series of 1000 simulations each of size \( n = 20, 50, 100 \) from a bivariate Pareto distribution with survival function

\[
F(t_1, t_2) = (t_1 + t_2 - 1)^{-\alpha}, \quad t_1, t_2 \geq 1
\]

for \( \alpha = 3, 6, \) and \( 10 \). We used the algorithms of Johnson (1987) for generating the observations. Simulation results are tabulated in Tables 2–4. In the tables, the bias (upper parenthesis) and mean squared errors (MSEs) (lower parenthesis) of the estimates are given at pairs of time points \((t_1, t_2)\), where \( t_1 \) and \( t_2 \) take values 1.09 and 1.57 corresponding to marginal survival probabilities 0.75 and 0.25. Results are also given at the point \((1, 1)\). A careful observation of the tables reveals that MSE decreases as sample size increases.
Table 2. Bias (upper parenthesis) and Mean squared error (lower parenthesis) for \( \overline{E}(X_i; t_1, t_2) \) at the selected five \((t_1, t_2)\) pairs for family (5.1), \( \alpha = 3, 6, 9. \)

| \( \alpha \) | \((t_1, t_2)\) | Sample size
|---|---|---|
| 3 | (1,1) | \((-0.5643, -0.4024)\) \((0.8014, 0.4613)\) \((-0.3013, -0.2986)\) \((0.1175, 0.1578)\) \((-0.1925, -0.1804)\) \((0.0739, 0.0333)\) |
| | (1.09, 1.09) | \((-0.4274, -0.3901)\) \((0.4559, 0.2535)\) \((-0.2807, -0.1416)\) \((0.1341, 0.1335)\) \((-0.1146, -0.1093)\) \((0.0406, 0.0231)\) |
| | (1.09, 1.57) | \((-0.0233, -0.6018)\) \((0.3054, 0.6620)\) \((-0.2070, -0.3966)\) \((0.1126, 0.2204)\) \((-0.1380, -0.1464)\) \((0.0269, 0.0241)\) |
| | (1.57, 1.09) | \((-0.1698, -0.6527)\) \((0.3468, 0.5973)\) \((-0.3962, -0.3278)\) \((0.1838, 0.1455)\) \((-0.2041, -0.1995)\) \((0.0622, 0.0404)\) |
| | (1.57, 1.57) | \((-0.0110, -0.3071)\) \((0.3368, 0.2324)\) \((-0.3081, -0.3282)\) \((0.1011, 0.1387)\) \((-0.1769, -0.1480)\) \((0.0372, 0.0330)\) |
| 6 | (1,1) | \((-0.4870, -0.2451)\) \((0.3272, 0.2361)\) \((-0.2458, -0.2478)\) \((0.1349, 0.1172)\) \((-0.1855, -0.1833)\) \((0.0416, 0.0463)\) |
| | (1.09, 1.09) | \((-0.3083, -0.4989)\) \((0.3743, 0.3281)\) \((-0.0917, -0.2779)\) \((0.1747, 0.1079)\) \((-0.1986, -0.1545)\) \((0.0774, 0.0309)\) |
| | (1.09, 1.57) | \((-0.4561, -0.4582)\) \((0.3022, 0.4275)\) \((-0.4447, -0.2666)\) \((0.2158, 0.1162)\) \((-0.2516, -0.2287)\) \((0.0847, 0.0556)\) |
| | (1.57, 1.09) | \((-0.4030, -0.4041)\) \((0.2657, 0.2478)\) \((-0.2673, -0.3432)\) \((0.1038, 0.1993)\) \((-0.2257, -0.2039)\) \((0.0563, 0.0695)\) |
| | (1.57, 1.57) | \((-0.1766, -0.1574)\) \((0.3265, 0.4868)\) \((-0.3082, -0.2886)\) \((0.1229, 0.1883)\) \((-0.2697, -0.2455)\) \((0.0859, 0.0614)\) |
| 9 | (1,1) | \((-0.3563, -0.6066)\) \((0.4615, 0.6340)\) \((-0.3311, -0.3209)\) \((0.1287, 0.2064)\) \((-0.1518, -0.2233)\) \((0.0269, 0.0550)\) |
| | (1.09, 1.09) | \((-0.4504, -0.2791)\) \((0.3670, 0.3619)\) \((-0.2333, -0.2544)\) \((0.1094, 0.1360)\) \((-0.2682, -0.2390)\) \((0.0838, 0.0710)\) |
| | (1.09, 1.57) | \((-0.1527, -0.4257)\) \((0.3488, 0.3007)\) \((-0.2261, -0.3002)\) \((0.1067, 0.1314)\) \((-0.2232, -0.1303)\) \((0.0563, 0.0637)\) |
| | (1.57, 1.09) | \((-0.1633, -0.4403)\) \((0.6873, 0.2661)\) \((-0.3586, -0.2627)\) \((0.1552, 0.1559)\) \((-0.1206, -0.1752)\) \((0.0253, 0.0331)\) |
| | (1.57, 1.57) | \((-0.6207, -0.4799)\) \((0.4889, 0.5809)\) \((-0.3978, -0.3026)\) \((0.1808, 0.1322)\) \((-0.2187, -0.2513)\) \((0.0645, 0.0699)\) |
Table 3. Bias (upper parenthesis) and Mean squared error (lower parenthesis) for \( \hat{e}(X_1 | x_j; t_1, t_2) \) at the selected five \((t_1, t_2)\) pairs for family (5.1), \( \alpha = 3, 6, 9. \)

| \( \alpha \) | \( \alpha | t_1, t_2 \) | \sample size\ | \( 20 \) | \( 50 \) | \( 100 \) |
|---|---|---|---|---|---|
| 3 | \((1, 1)\) | \(-0.7357, -0.4027\) | \(-0.3044, -0.2958\) | \(-0.2481, -0.2044\) |
|  | (1.09, 1.09) | \(-0.4519, -0.3813\) | \(-0.3106, -0.4337\) | \(-0.2040, -0.1909\) |
|  | (1.09, 1.57) | \(-0.0067, -0.4936\) | \(-0.2562, -0.2249\) | \(-0.2566, -0.1585\) |
|  | (1.57, 1.09) | \(-0.2418, -0.4455\) | \(-0.2428, -0.1394\) | \(-0.1975, -0.2228\) |
|  | (1.57, 1.57) | \(-0.2774, -0.4133\) | \(-0.2958, -0.2771\) | \(-0.2448, -0.2030\) |
| 6 | \((1, 1)\) | \(-0.4130, -0.5213\) | \(-0.2755, -0.2787\) | \(-0.1615, -0.1583\) |
|  | (1.09, 1.09) | \(-0.4945, -0.1444\) | \(-0.3450, -0.2798\) | \(-0.1651, -0.2255\) |
|  | (1.09, 1.57) | \(-0.3789, -0.2927\) | \(-0.2482, -0.3143\) | \(-0.1474, -0.2062\) |
|  | (1.57, 1.09) | \(-0.2523, -0.7432\) | \(-0.3152, -0.3567\) | \(-0.1453, -0.1316\) |
|  | (1.57, 1.57) | \(-0.4284, -0.3559\) | \(-0.2318, -0.3145\) | \(-0.2199, -0.2198\) |
| 9 | \((1, 1)\) | \(-0.3493, -0.4839\) | \(-0.3978, -0.3954\) | \(-0.2311, -0.2383\) |
|  | (1.09, 1.09) | \(-0.3373, -0.4132\) | \(-0.1899, -0.2961\) | \(-0.2208, -0.2320\) |
|  | (1.09, 1.57) | \(-0.4409, -0.3504\) | \(-0.3488, -0.3065\) | \(-0.2102, -0.2216\) |
|  | (1.57, 1.09) | \(-0.4891, -0.2322\) | \(-0.2531, -0.2992\) | \(-0.1859, -0.1807\) |
|  | (1.57, 1.57) | \(-0.3075, -0.3922\) | \(-0.2970, -0.3006\) | \(-0.2464, -0.2772\) |
Table 4. Bias (upper parenthesis) and Mean squared error (lower parenthesis) for \( \hat{\xi}(X_1 \mid X_1; t_1, t_2) \) at the selected five \((t_1, t_2)\) pairs for family (5.1), \( \alpha = 3, 6, 9. \)

| \( \alpha \) | \( (t_1, t_2) \)       | 20                  | 50                  | 100                  |
|------------|------------------------|---------------------|---------------------|---------------------|
| 3          | (1, 1)                 | \((-0.3329, -0.4913)\) | \((-0.3131, -0.3201)\) | \((-0.2168, -0.2262)\) |
|            | (1.09, 1.09)           | \((-0.4812, -0.4670)\) | \((-0.2530, -0.2735)\) | \((-0.1401, -0.2146)\) |
|            | (1.09, 1.57)           | \((-0.1781, -0.4202)\) | \((-0.2905, -0.3891)\) | \((-0.1791, -0.2305)\) |
|            | (1.57, 1.09)           | \((-0.4304, -0.4875)\) | \((-0.3058, -0.2618)\) | \((-0.1882, -0.1383)\) |
|            | (1.57, 1.57)           | \((-0.5037, -0.3104)\) | \((-0.3006, -0.2672)\) | \((-0.2319, -0.1444)\) |
| 6          | (1, 1)                 | \((-0.4395, -0.4449)\) | \((-0.3475, -0.3784)\) | \((-0.2481, -0.1799)\) |
|            | (1.09, 1.09)           | \((-0.5444, -0.4869)\) | \((-0.3119, -0.3198)\) | \((-0.2132, -0.2632)\) |
|            | (1.09, 1.57)           | \((-0.4177, -0.5763)\) | \((-0.3038, -0.3137)\) | \((-0.1811, -0.2000)\) |
|            | (1.57, 1.09)           | \((-0.2572, -0.7139)\) | \((-0.2899, -0.2928)\) | \((-0.2055, -0.1624)\) |
|            | (1.57, 1.57)           | \((-0.5343, -0.3732)\) | \((-0.2934, -0.3453)\) | \((-0.1597, -0.2485)\) |
| 9          | (1, 1)                 | \((-0.4031, -0.6694)\) | \((-0.2804, -0.3640)\) | \((-0.2270, -0.1855)\) |
|            | (1.09, 1.09)           | \((-0.5365, -0.5728)\) | \((-0.2855, -0.2973)\) | \((-0.2096, -0.2026)\) |
|            | (1.09, 1.57)           | \((-0.4277, -0.4326)\) | \((-0.3952, -0.2961)\) | \((-0.2332, -0.1442)\) |
|            | (1.57, 1.09)           | \((-0.4438, -0.5852)\) | \((-0.3622, -0.2436)\) | \((-0.2038, -0.1176)\) |
|            | (1.57, 1.57)           | \((-0.4149, -0.4927)\) | \((-0.3305, -0.3066)\) | \((-0.2530, -0.2152)\) |

Sample size
Acknowledgements

The authors would like to express their gratitude to the referees and the editor for their valuable suggestions which have considerably improved the earlier version of the paper.

References

Asadi, M. and Zohrevand, Y. (2007). On dynamic cumulative residual entropy, *J. Stat. Plann. Inference*, **137**, 1931–1941.

Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution, *Sankhya A*, **58**, 48–56.

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

Ebrahimi, N., Kirmani, S. N. U. A. and Soofi, E. S. (2007). Multivariate dynamic information, *J. Multivar. Anal.*, **98**, 328–349.

Ebrahimi, N. and Pellerey, F. (1995). New partial ordering of survival functions based on the notion of uncertainty, *J. App. Prob.*, **32**, 202–211.

Johnson, M. E. (1987). *Multivariate Statistical Simulation*, Wiley, New York.

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

Nadarajah, S. and Zografos, K. (2005). Expressions for Renyi and Shannon entropies for bivariate distributions, *Information Sciences*, **170**, 173–189.

Nanda, K. N. and Paul, P. (2005). Some results on generalized residual entropy, *Information Sciences*, **176**, 27–47.

Rajesh, G. and Nair, K. R. M. (2000). Residual entropy of conditional distributions, *Statistical Methods*, **2**(1), 72–80.

Rajesh, G., Sathar, A. E. I. and Nair, K. R. M. (2009). Bivariate extension of residual entropy and some characterization results, *J. Indian Stat. Assoc.*, **47**, 91–107.

Rajesh, G., Sathar, A. E. I., Nair, K. R. M. and Reshmi, K. V. (2014). Bivariate extension of dynamic cumulative residual entropy, *Statistical Methodology*, **16**, 72–82.

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

Sathar, A. E. I., Nair, K. R. M. and Rajesh, G. (2009). Generalized bivariate residual entropy function and some characterization results, *S. Afr. Statist. J.*, **44**, 1–18.

Sathar, A. E. I., Rajesh, G. and Nair, K. R. M. (2010). Bivariate Geometric vitality function and some characterization results, *Bull. - Calcutta Stat. Assoc.*, **62**, 247–267.

Shannon, C. E. (1948). A mathematical theory of communication, *Bell System Tech. J.*, **279–423, 623–656.

Wang, F. and Vemuri, B. C., (2007). Non-rigid multimodal image registration using cross-cumulative residual entropy, *Int. J. Comput. Vis.*, **74**, 201–205.

Wang, F., Vemuri, B. C., Rao, M. and Chen, Y. (2003). A new and robust information theoretic measures and its application to image alignment, *Proc. of the International Conference of Information Processing in Medical Imaging (IPMF)*, 388–400.