Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/050315

On Dynamic Cumulative Residual Entropy of Order Statistics

Osman Kamari[*]

College of Administration and Economics, University of Human Development, Sulaymaniyah, Iraq.

Received: 6 May. 2016, Revised: 22 Oct. 2016, Accepted: 27 Oct. 2016
Published online: 1 Nov. 2016

Abstract: The entropy functions are useful tools to measure the uncertainty in a random variable. Dynamic Cumulative Residual Entropy (DCRE) introduced by Asadi and Zohrevand [1] as a useful dynamic measure of Cumulative Residual Entropy. They studied some properties and applications of these measures. In this paper, Dynamic Cumulative Residual Entropy is proposed based on order statistics and under conditions is showed a characterization result that Dynamic Cumulative Residual Entropy of order statistics can determine the distribution function uniquely. Then some properties for DCRE of order statistics is presented.

Keywords: Dynamic Cumulative Residual Entropy, Order statistics, Mean residual lifetime, Survival function

1 Introduction

The concept of Shannon entropy as a basic measure of uncertainty for a random variable was introduced by Shannon [2]. Suppose X be a continuous random variable having probability density function f and cumulative distribution function F. Therefore, Shannon entropy $H(f)$ of X is defined as follows:

$$H(f) = -E[\log f(x)] = -\int_{-\infty}^{+\infty} f(x) log f(x)\, dx, \tag{1.1}$$

Study of duration is a subject of interest in many fields of science such as reliability, survival analysis, economics and business. In reliability theory and survival analysis, the additional life time given that the component has survived up to time t is called the residual life function of the component. If X be the life of a component, then $X_t = (X - t \mid X > t)$ is called the residual life function.
If a component is known to have survived to age t then Shannon entropy is no longer useful to measure the uncertainty of remaining lifetime of the component. Therefore, Ebrahimi [3] defined the entropy for residual lifetime $X_t = (X - t \mid X > t)$ as a dynamic form of uncertainty called the residual entropy at time t and defined as

$$H(X;t) = -\int_{t}^{+\infty} \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)}\, dx, \tag{1.2}$$

That $\overline{F}(t) = P(X > t) = 1 - F(t)$ is the survival (reliability) function of X.

Rao et al. [4] defined Cumulative Residual Entropy(CRE) as an alternative measure of uncertainty to Shannon entropy in that the probability density function is replaced by survival function and obtained some properties and applications of that in reliability engineering and computer vision. Also they showed CRE overcomes some problems with Shannon entropy. Such as CRE possesses more general mathematical properties than Shannon entropy and easily is computed from sample data. The measure is more consistent since it is based on distribution function than the density function which is a derivative of the distribution function. For more details see [4] and [5]. CRE for a non-negative univariate random variable

[*] Corresponding author e-mail: osman.kamari@uhd.edu.iq

is by:

$$\xi(X) = -\int_0^{+\infty} \overline{F}(x) \log \overline{F}(x)\, dx, \tag{1.3}$$

Analogous to the residual entropy, Asadi and Zohrevand [1] defined a dynamic form of CRE, that is CRE of X_t. This function is called Dynamic Cumulative Residual Entropy(DCRE) and defined as

$$\xi(X;t) = -\int_t^{+\infty} \frac{\overline{F}(x)}{\overline{F}(t)} \log \frac{\overline{F}(x)}{\overline{F}(t)}\, dx, \tag{1.4}$$

It's clear that $\xi(X;0) = \xi(X)$.

Suppose $X_1, X_2, ..., X_n$ be a random sample, the order statistics of the sample is defined by the arrangement $X_1, X_2, ..., X_n$ from the minimum to the maximum by $X_{(1)}, X_{(2)}, ..., X_{(n)}$. Order statistics are widely used in reliability theory and survival analysis to study $(n-k+1)$ out of n system which works if and only if at least $(n-k+1)$ out of n components are working. Series and parallel systems are particular cases of these system corresponding to $k = 1$ and $k = n$, respectively. For more details see [9] , [10].

Baratpour et al. [6], [7] presented some properties of the entropy of order statistics and record values and established some characterization results. Also, Baratpour [8] have derived characterizations result based on Cumulative residual entropy of first order statistics. Park and Kim [11] defined the cumulative residual entropy of first r order statistics.

The purpose of this paper is determination distribution function using Dynamic Cumulative Residual Entropy of the i^{th} order statistic. The paper is organized as follows: In section 2, Dynamic Cumulative Residual Entropy is defined based on order statistics. Section 3 includes Characterization property based on i^{th} order statistics. Some properties of CRE and DCRE for i^{th} order statistics is presented in section 4. In section 5, DCRE of type I censored data is given. Following a brief conclusion in the end of paper is given.

2 Dynamic Cumulative Residual Entropy of Order Statistics

Suppose $X_1, X_2, ..., X_n$ be a random sample with distribution function F, the order statistics of the sample is defined by the arrangement $X_1, X_2, ..., X_n$ from the minimum to the maximum by $X_{(1)}, X_{(2)}, ..., X_{(n)}$. The cumulative distribution function of the $i^{th}(i = 1, 2, ..., n)$ order statistics is given by

$$F_i(x) = \sum_{k=i}^{n} \binom{n}{k} [F(x)]^i [1 - F(x)]^{(n-i)} \tag{2.1}$$

First order statistics $X_{(1)}$ and last order statistics $X_{(n)}$ are very important special case of order statistics in practice. In many statistical applications the interest is centered on estimating the maximum or the minimum. (See [12]).

Then cumulative distribution function of $X_{(1)}$ and $X_{(n)}$ are respectively as

$$F_{X_{(1)}}(x) = 1 - (1 - F(x))^n, \tag{2.2}$$

$$F_{X_{(n)}}(x) = (F(x)^n, \tag{2.3}$$

Park and Kim [11] defined the Cumulative Residual Entropy of the i^{th} order statistics as

$$\xi(F_{(i)}) = -\int_0^{\infty} \overline{F}_{(i)}(x) \log \overline{F}_{(i)}(x)\, dx, \tag{2.4}$$

Then Dynamic Cumulative Residual Entropy of the i^{th} order statistics is defined as follows:

$$\xi(F_{(i)};t) = -\int_t^{\infty} \frac{\overline{F}_{(i)}(x)}{\overline{F}_{(i)}(t)} \log \frac{\overline{F}_{(i)}(x)}{\overline{F}_{(i)}(t)}\, dx, \tag{2.5}$$

It's clear that $\xi(F_{(i)};0) = \xi(F_{(i)})$.

If $i = 1$, we have

$$\xi(F_{(1)};t) = -\int_t^{\infty} \frac{\overline{F}_{(1)}(x)}{\overline{F}_{(1)}(t)} \log \frac{\overline{F}_{(1)}(x)}{\overline{F}_{(1)}(t)} dx, \tag{2.6}$$

Using relation (2.2) and changing variable $\overline{F}(X) = U$ in relation (2.5), we can rewrite by

$$\xi(F_{(1)};t) = -\frac{n}{(\overline{F}(t))^n} \int_0^{\overline{F}(t)} \frac{u^n \log u}{f(F^{-1}(1-u)} du + \frac{n \log \overline{F}(t)}{(\overline{F}(t))^n} \int_0^{\overline{F}(t)} \frac{u^n}{f(F^{-1}(1-u)} du, \tag{2.7}$$

In the following example, DCRE of the first order statistics is calculated for an exponentially distributed random variable .

Example2.1 If X is distributed with parameter of (θ) and survival function of $\overline{F}(t) = 1 - e^{-\theta x}$, then easily using relation ((2.7) we can obtain $\xi(F_{(1)};t) = \frac{1}{n\theta}$.

3 Characterization property based on DCRE of i^{th} order statistics

Gupta, et al. [13] have studied characterizations based on dynamic Shannon entropy of order statistics. In this section characterization property on the Dynamic Cumulative Residual Entropy of the i^{th} order statistics is studied by using the sufficient condition for the uniqueness of the solution of initial value problem

$$y' = f(x,y), y(x_0) = y_0,$$

That f is a function of two variables whose domain is a region $S \subset R^2$, (x_0, y_0) is a point in S and y is the unknown function. By the solution of the initial value problem on an interval $I \subset R$, we mean a function $\phi(x)$ satisfies the conditions: ϕ is differentiable on I, the growth of ϕ lies in S, $\phi(x_0) = y_0$ and $\phi'(x) = f(x, \phi(x))$ for all $x \in I$.

The following theorem and lemma are used in proving of theorem 3.2. see [14].

Theorem 3.1 Suppose that f be a continuous function defined in a domain $S \subset R^2$ is said to satisfy Lipschitz condition with respect to y on the domain S, that is

$$|\ f(x,y_1) - f(x,y_2)\ | \leq k\ |\ y_1 - y_2\ |, k > 0$$

For every pair of points (x,y_1) and (x,y_2) in D. Then $y = \phi(x)$ satisfy the initial value problem $y' = f(x,y)$ and $\phi(x_0) = y_0$ is unique.

Lemma 3.1 Let f be a continuous function in a convex region $S \subset R^2$. Suppose $\frac{\partial f}{\partial y}$ exists and it's continuous in S. Then f satisfies Lipschitz condition in S.

Now in the following theorem characterization property on DCRE of i^{th} order statistics is presented.

Theorem 3.2 Suppose that X be a non-negative continuous random variable with cumulative distribution function of F and with $\xi(F_{(i)};t) < \infty$, $t \geq 0$. Then $\xi(F_{(i)};t)$ uniquely determines F.

Proof: let F_1, F_2 are two functions such that

$$\xi(F_{1(i)};t) = \xi(F_{2(i)};t), t \geq 0, i \leq n$$

Differentiating both sides of the relation (2.5) with respect to t and relationship between hazard rate function and mean residual lifetime function for the i^{th} order statistics given by

$$r_{F_i}(t) = \frac{m'_{F_i}(t) + 1}{m_{F_i}(t)}, \tag{3.1}$$

we have

$$m'_{F_i}(t) = \frac{m_{F_i}(t) - \xi(F_{(i)};t) + m_{F_i}(t)\frac{d\xi(F_{(i)};t)}{dt}}{\xi(F_{(i)};t) - m_{F_i}(t)} \tag{3.2}$$

Suppose

$$\xi(F_{1(i)};t)=\xi(F_{2(i)};t)=h(t), t\geq 0, i\leq n$$

that F_1 and F_2 are to distribution functions. Then for $t\geq 0$ from relation (3.2) we conclude that
$m'_{F_{1_{(i)}}}(t)=\psi(t,m_{F_{1_{(i)}}}(t)), m'_{F_{2_{(i)}}}(t)=\psi(t,m_{F_{2_{(i)}}}(t))$
where

$$\psi(t,y)=\frac{y-h(t)+yh'(t)}{h(t)-y}$$

By applying theorem 3.1 and lemma 3.1 we obtain
$m_{F_{1_{(i)}}}(t)=m_{F_{2_{(i)}}}(t), t\geq 0, i\leq n$
Which using relation (3.1) gives $r_{F_{1_{(i)}}}(t)=r_{F_{2_{(i)}}}(t), t\geq 0, i\leq n$ Hence the proof is complete.

4 Some Properties on CRE and DCRE for i^{th} Order Statistics

Mean residual lifetime play important role in reliability and survival analysis. The following property shows the relation between $\xi_{F_(i)}(t)$ and $m_{F_(i)}(t)$.

Proposition 1 Suppose $\xi_{F_(i)}<\infty$ denotes the CRE of the i^{th} order statistics, then

$$\xi_{F_(i)}=E[m_{F_{(i)}}(X)].$$

Proof: Refer to [1], theorem 2.1.

Proposition 2 Let $\xi_{F_(i)}<\infty$ denotes the DCRE of i^{th} order statistics, then

$$\xi(F_{(i)};t)=E[m_{F_{(i)}}(X)|X_{(i)}\geq t]$$

Proof: the proof of the theorem follows the same steps as used to prove P 1.

In the following proposition the upper bound is derived for the DCRE of the i^{th} order statistics.

Proposition 3 The upper bound for the DCRE of i^{th} order statistics is as follows

$$\xi(F_{(i)};t)\leq\frac{\xi(F_{(i)}}{\overline{F}_{(i)}(t)}$$

Proof: we can write $\xi(F_{(i)};t)$ as follows

$$\xi(F_{(i)};t)=-\frac{1}{\overline{F}_{(i)}(t)}\int_t^\infty\overline{F}_{(i)}(x)\log\overline{F}_{(i)}(x)\,dx+\frac{\log\overline{F}_{(i)}(t)}{\overline{F}_{(i)}(t)}\int_t^\infty\overline{F}_{(i)}(x)\,dx, t\geq 0$$

Since $\log\overline{F}_{(i)}(t)\leq 0$, we have

$$\xi(F_{(i)};t)\leq-\frac{1}{\overline{F}_{(i)}(t)}\int_t^\infty\overline{F}_{(i)}(x)\log\overline{F}_{(i)}(x)\,dx\leq-\frac{1}{\overline{F}_{(i)}(t)}\int_0^\infty\overline{F}_{(i)}(x)\log\overline{F}_{(i)}(x)\,dx,$$

hence

$$\xi(F_{(i)};t)\leq\frac{\xi(F_{(i)}}{\overline{F}_{(i)}(t)}$$

if $t\to 0$ then $\xi(F_{(i)};t)=\xi(F_{(i)})$.

5 DCRE of type I Censored Data

If we have the Type I right censored variable, $min(X,C)$ that C is the censoring point assumed to be a constant, the survival function of $min(X,C)$ is

$$\bar{F}_C(x) = \begin{cases} \bar{F}_C(x), x < C \\ 0, x \geq C \end{cases}$$

Then the DCRE of the Type I censored variable can be defined as follows:

$$\xi_C(F;t) = -\int_t^C \frac{\overline{F}(x)}{\overline{F}(t)} \log \frac{\overline{F}(x)}{\overline{F}(t)} dx,$$

Example 5.1 Let $X \sim Exp(\theta)$ with survival function of $\overline{F}(x) = e^{-\theta x}$, hence $\xi_C(F;t)$ can be written as

$$\xi_C(F;t) = e^{-\theta(t-C)}[t - C - \frac{1}{\theta}] + \frac{1}{\theta}$$

6 Conclusion

The entropy functions are efficient measures of uncertainty in a random variable that are applied in a lot of fields such as reliability, finance, economics, insurance, medicine, and etc. Cumulative Residual Entropy is an alternative measure of uncertainty to Shannon entropy in that probability density function is replaced by survival function. The Dynamic form of Cumulative Residual Entropy measures the residual lifetime of the component has survived up to time t. The entropy measures based on order statistics have been studied widely and are crucial to measure uncertainty in statistical modeling. In this paper, Dynamic Cumulative Residual Entropy proposed based on order statistics and under conditions showed a characterization result that Dynamic Cumulative Residual Entropy of order statistics can determine the distribution function uniquely. The theoretical outcomes in this paper can be interest both from theoretical as well as practical point of view.

Acknowledgement

The author is thankful to referee for some useful comments which improved the presentation.

References

[1] M. Asadi, Y. Zohrevand. On the dynamic cumulative residual entropy. Journal of Statistical Planning and Inference **137**, 1931-1941 (2007).
[2] C. E. Shannon, A mathematical theory of communication. Bell System Technical J **27**, 379-423 (1948).
[3] N. Ebrahimi, How to measure uncertainty about residual lifetime. Sankhya A **58** 48-57(1996).
[4] M. Rao, Y. Chen, B. Vemuri, and F. Wang F. Cumulative residual entropy: a new measure of information. IEEE Transactions on Information Theory **50**, 1220-1228 (2004).
[5] M. Rao,More on a new concept of entropy and information. J. Theoret. Probab. **18**, 967-981(2005).
[6] S.Baratpour, J. Ahmadi, N. R. Arghami. Some characterizations based on entropy of order statistics and record values. Communications in Statistics-Theory and Methods **36**, 47-57(2007).
[7] S. Baratpour, J. Ahmadi, N. R. Arghami, Characterizations based on Renyi entropy of order statistics and record values. Journal of Statistical Planning and Inference. **138**, 2544-2551(2008).
[8] S. Baratpour, Characterizations Based on Cumulative Residual Entropy of First-Order Statistics, Communications in Statistics - Theory and Methods, **39**: 20, 3645 -3651(2010).
[9] B.C. Arnold, N. Balakrishnan, H. N. Nagaraja, A First Course in Order Statistics. New York: John Wiley & Sons.(1992).
[10] H. A. David, H. N. Nagaraja, Order Statistics. 3rd ed. New York: John Wiley & Sons.(2003).
[11] S. Park, and I. Kim, I, On cumulative residual entropy of order statistics, Statistics and probability letters, **94**, 170-175(2014).
[12] E. Castillo, A. S. Hadi, N. Balakrishnan, and J. M. Sarabia. Extreme Value and Related Models with Applications in Engineering and Science, John Wiley & Sons(2005).
[13] R.C. Gupta, H. C. Taneja, and Richa. Thapliyal. Stochastic Comparisons of Residual Entropy of Order Statistics and Some Characterization Results. Journal of Statistical Theory and Applications, **13** (1), 27-37(2014).
[14] R. C. Gupta, and S. N. U. A. Kirmani. Characterizations based on conditional mean function, Journal of Statistical Planning and Inference, **138**, 964-970(2008).