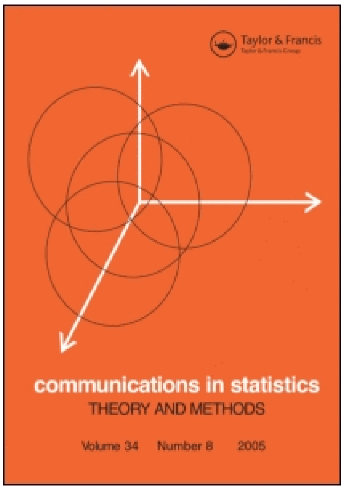

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

Characterizations Based on Cumulative Residual Entropy of First-Order Statistics

S. Baratpour[a]
[a] Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

Online publication date: 22 September 2010

To cite this Article Baratpour, S.(2010) 'Characterizations Based on Cumulative Residual Entropy of First-Order Statistics', Communications in Statistics - Theory and Methods, 39: 20, 3645 — 3651

To link to this Article: DOI: 10.1080/03610920903324841

URL: http://dx.doi.org/10.1080/03610920903324841

Communications in Statistics—Theory and Methods, 39: 3645–3651, 2010

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920903324841

Characterizations Based on Cumulative Residual Entropy of First-Order Statistics

S. BARATPOUR

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

Two different distributions may have equal cumulative residual entropy (CRE), thus a distribution cannot be determined by its CRE. In this article, we explore properties of the CRE and study conditions under which the CRE of the first-order statistics can uniquely determines the parent distribution. Weibull family is characterized through ratio of the CRE of the first-order statistics to its expectation. We have also some bounds for the CRE of residual lifetime of a series system.

Keywords Cumulative residual entropy; Order statistics; Residual lifetime distribution; Shannon information; Series system; Weibull family.

Mathematics Subject Classification Primary 62G30; Secondary 62E10, 62B10, 94A17.

1. Introduction

In information theory, entropy is a measure of the uncertainty associated with a random variable. This concept was introduced by Shannon (1948). Shannon entropy represents an absolute limit on the best possible lossless compression of any communication. Shannon entropy of a discrete random variable X with possible values $\{x_1, x_2, \ldots, x_n\}$ and probability mass function p is defined as

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i).$$

The formula

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

where f denotes the probability density function (pdf) of the continuous random variable X, is an extension of the Shannon entropy and is usually referred to as the

Received November 19, 2008; Accepted September 9, 2009
Address correspondence to S. Baratpour, Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, P.O. Box 91775-1159, Mashhad, Iran; E-mail: baratpur@math.um.ac.ir

differential entropy. Shannon entropy has been used as a major tool in information theory on in almost every branch of science and engineering. Numerous entropy and information indices, among them Renyi entropy, were developed and used in various disciplines and contexts.

Rao et al. (2004) introduced a new measure of information that extends the Shannon entropy to continuous random variables, and called it cumulative residual entropy (CRE). He showed that it is more general than the Shannon entropy and possesses more general mathematical properties than the Shannon entropy. It can easily computed from sample data and its estimation asymptotically converges to the true value. CRE has applications in reliability engineering and computer vision, for more details see Rao (2005). This measure is based on the cumulative distribution function (cdf) F and is defined as follows:

$$CRE(X) = -\int_0^{+\infty} p(|X| > x) \log p(|X| > x) \mathrm{d}x.$$

In reliability theory, CRE is based on survival function $\overline{F}(x)$, and is defined as

$$CRE(X) = -\int_0^{+\infty} \overline{F}(x) \log \overline{F}(x) \mathrm{d}x.$$

In this article, we suppose X is a positive continuous random variable. If we use change of variable $u = \overline{F}(x)$, then

$$CRE(X) = -\int_0^1 \frac{u \log u}{f(F^{-1}(1-u))} \mathrm{d}u, \tag{1}$$

where F^{-1} is the inverse function of F.

Suppose that $X_1, \ldots, X_n$ are independent and identically distributed (iid) observations from cdf $F(x)$ and pdf $f(x)$. The order statistics of the sample is defined by the arrangement of $X_1, \ldots, X_n$ from the smallest to the largest, denoted as $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$. These statistics have been used in a wide range of problems, including robust statistical estimation, detection of outliers, characterization of probability distributions and goodness-of-fit tests, entropy estimation, analysis of censored samples, quality control and strength of materials; for more details, see Arnold et al. (1992), David and Nagaraja (2003), and references therein. Also, in reliability theory, order statistics are used for statistical modeling. The mth order statistics in a sample of size n represents the life length of a $(n-m+1)$-out of-n system.

Ebrahimi (1996) defined the concept of dynamic Shannon entropy and obtained some properties of that. Since then several attempts have been made to extend that, see Asadi et al. (2005a,b). Asadi and Zohrevand (2007) proposed a dynamic form of CRE and obtained some of its properties. The CRE for the residual lifetime distribution of a system with survival function $\overline{F}_t(x) = P(X - t > x \mid X > t) = \frac{\overline{F}(x+t)}{\overline{F}(t)}$, is

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

It is clear that, for $t = 0$, $CRE(X; t) = CRE(X)$.

The aim of this article is characterizing the parent distributions based on the CRE of first-order statistics. It is shown that the equality of the CRE in first-order statistics can determine uniquely the parent distribution. The rest of this article is organized as follows. Section 2 contains some characterizations based on the first-order statistics ($X_{1:n}$); also, we characterize Weibull distribution based on the ratio of the $CRE(X_{1:n})$ to the $E(X_{1:n})$. In Sec. 3, we also have some characterization based on CRE of the residual lifetime distribution.

2. Characterization Based on First-Order Statistics

Let $X_{1:n}$ be the first-order statistic in a random sample of size n from a positive and continuous random variable X with cdf F and pdf f, then the cdf of $X_{1:n}$ is given by

$$F_{X_{1:n}}(x) = 1 - \overline{F}^n(x).$$

Thus,

$$CRE(X_{1:n}) = -n \int_0^\infty \overline{F}^n(x) \log \overline{F}(x) \mathrm{d}x.$$

By change of variable $u = \overline{F}(x)$, we have

$$CRE(X_{1:n}) = -n \int_0^1 \frac{u^n \log u}{f(F^{-1}(1-u))} \mathrm{d}u. \tag{2}$$

First, let us look at the following examples.

Example 2.1. Suppose X has a Pareto(α, β) distribution, with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$. That is, the pdf $f(x)$ is given by $f(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}}$, $x \geq \beta$. Using (1) and (2), the CRE(X) and the CRE($X_{1:n}$) are as follows:

$$\begin{aligned} CRE(X) &= \frac{\alpha\beta}{(\alpha-1)^2}, \quad \alpha > 1 \\ &= +\infty, \quad \alpha \leq 1, \end{aligned}$$

and

$$\begin{aligned} CRE(X_{1:n}) &= \frac{n\alpha\beta}{(n\alpha-1)^2}, \quad \alpha > \frac{1}{n} \\ &= +\infty, \quad \alpha \leq \frac{1}{n}. \end{aligned}$$

Let $\alpha > 1$ and $\Delta_1 = CRE(X) - CRE(X_{1:n})$, then $\Delta \geq 0$, that means for $\alpha > 1$, uncertainty of X is more than $X_{1:n}$. Similarly, this property is obtained for every α and β if we replace $CRE(X)$ and $CRE(X_{1:n})$ by $H(X)$ and $H(X_{1:n})$, respectively. We can also show that for $n > \frac{1}{\alpha}$, Δ_1 is an increasing function of n.

Example 2.2. A non negative random variable X is Weibull distributed, if its cdf is

$$F(x) = 1 - \exp(-\lambda^q x^q), \quad \lambda > 0, \quad q > 0, \quad x > 0,$$

where λ and q are, respectively, scale and shape parameters. We can show that $E(X) = \frac{1}{\lambda}\Gamma(1+\frac{1}{q})$ and $CRE(X) = \frac{1}{\lambda q}\Gamma(1+\frac{1}{q})$. Thus, $\frac{CRE(X)}{E(X)} = \frac{1}{q}$. If $q = 1$, then X is standard exponential distributed and this ratio is equal 1.

By (2), we have

$$\begin{aligned} CRE(X_{1:n}) &= \frac{n}{\lambda q}\int_0^1 u^{n-1}(-\log u)^{\frac{1}{q}}\mathrm{d}u \\ &= \frac{1}{\lambda q n^{\frac{1}{q}}}\Gamma\left(1+\frac{1}{q}\right). \end{aligned}$$

On the other hand, $E(X_{1:n}) = \frac{1}{\lambda n^{\frac{1}{q}}}\Gamma(1+\frac{1}{q})$. Thus, $\frac{CRE(X_{1:n})}{E(X_{1:n})} = \frac{1}{q}$. This result shows that for all n, in Weibull family, this ratio is constant. If $q = 1$, then for standard exponential distribution this ratio is equal 1, for all n. We also have

$$\Delta_2 = CRE(X) - CRE(X_{1:n}) = \frac{1}{\lambda q}\Gamma\left(1+\frac{1}{q}\right)\left(1-\frac{1}{n^{\frac{1}{q}}}\right) \geq 0,$$

for all n and is increasing in n.

In the following theorem, we show that only in Weibull family the ratio $\frac{CRE(X_{1:n})}{E(X_{1:n})}$ is constant. The following lemma is used in the following theorem. It is known in the literature as Müntz-Szász Theorem, which is often invoked in moment-based characterization theorems; see Kamps (1998) and Borwein and Erdelyi (1995, Sec. 4.2).

Lemma 2.1. *For any increasing sequence of positive integers $\{n_j, j \geq 1\}$, the sequence of polynomials $\{x^{n_j}\}$ is complete on $L(0, 1)$, if and only if $\sum_{j=1}^{\infty} n_j^{-1} = +\infty$.*

Theorem 2.1. *Suppose that $X_1, \ldots, X_n$ are positive, independent and identically distributed (iid) observations from an absolutely continuous cdf $F(x)$ and pdf $f(x)$. Then F belong to Weibull family, if and only if $\frac{CRE(X_{1:n})}{E(X_{1:n})} = c$ (>0), for all $n = n_j, j \geq 1$, such that $\sum_{j=1}^{+\infty} n_j^{-1} = +\infty$.*

Proof. By Example 2.1, necessity is trivial, hence it remains to prove the sufficiency part. By using change of variable $\overline{F}(x) = u$ in $E(X_{1:n}) = \int_0^{\infty} nxf(x)\overline{F}^{n-1}(x)\mathrm{d}x$, we have

$$E(X_{1:n}) = n\int_0^1 F^{-1}(1-u)u^{n-1}\mathrm{d}u. \tag{3}$$

Using (2) and (3), we have

$$\frac{CRE(X_{1:n})}{E(X_{1:n})} = -\frac{\int_0^1 \frac{u^n \log u}{f(F^{-1}(1-u))}\mathrm{d}u}{\int_0^1 F^{-1}(1-u)u^{n-1}\mathrm{d}u}. \tag{4}$$

If (4) coincides c, we can conclude that

$$\int_0^1 u^{n-1}\left[\frac{u\log u}{f(F^{-1}(1-u))} + cF^{-1}(1-u)\right]\mathrm{d}u = 0. \tag{5}$$

If (5) holds for $n = n_j$, $j \geq 1$, such that $\sum_{j=1}^{+\infty} n_j^{-1} = +\infty$, then from Lemma 2.1, we have

$$\frac{(1-v)\log(1-v)}{f(F^{-1}(v))} + cF^{-1}(v) = 0 \quad a.e, v \in (0, 1).$$

Since $\frac{d}{dv}F^{-1}(v) = \frac{1}{f(F^{-1}(v))}$, it then follows:

$$(1-v)\log(1-v)\frac{d}{dv}F^{-1}(v) + cF^{-1}(v) = 0 \quad a.e, v \in (0, 1).$$

After solving this differential equation, we can result that $F^{-1}(v) = c_1[-\log(1-v)]^c$, $v \in (0, 1)$, thus $F(x) = 1 - \exp\left(-(\frac{x}{c_1})^{\frac{1}{c}}\right)$, $x > 0$. This means that F belong to the Weibull family.

Theorem 2.2. *Let X and Y be two positive random variables with pdfs $f(x)$ and $g(x)$ and absolutely continuous cdfs $F(x)$ and $G(x)$, respectively. Then F and G belong to the same family of distributions, but for a change in location, if and only if*

$$CRE(X_{1:n}) = CRE(Y_{1:n}),$$

for $n = n_j$, $j \geq 1$ such that $\sum_{j=1}^{+\infty} n_j^{-1}$ is infinite.

Proof. The necessity is trivial, hence it remains to prove the sufficiency part. By (2), if $CRE(X_{1:n}) = CRE(Y_{1:n})$, then we have

$$\int_0^1 u^n \log u \left[\frac{1}{f(F^{-1}(1-u))} - \frac{1}{g(G^{-1}(1-u))}\right] \mathrm{d}u = 0. \tag{6}$$

If (6) holds for $n = n_j$, $j \geq 1$, such that $\sum_{j=1}^{+\infty} n_j^{-1} = \infty$, then from Lemma 2.1 we can conclude that $f(F^{-1}(t)) = g(G^{-1}(t))$, $0 < t < 1$. Since $\frac{d}{dv}F^{-1}(t) = \frac{1}{f(F^{-1}(t))}$, we have $\frac{d}{dv}F^{-1}(t) = \frac{d}{dv}G^{-1}(t)$, $0 < t < 1$. It then follows that $F^{-1}(t) = G^{-1}(t) + d$, $0 < t < 1$. This means F and G belong to the same family of distributions, but for a location shift.

Baratpour et al. (2007, 2008) obtained Similar properties based on Shannon entropy and Renyi entropy of order statistics and record values.

3. Characterizations Based on CRE of Residual Lifetime of Series Systems

An important method of improving the reliability of a system is to build redundancy to it. A common structure of redundancy is the k-out-of-n systems and an important special of it is series systems (see, e.g., Xie and Lai, 1996).

A series system consisting of n components, is a system which functions if and only if all of its n components function. Let $X_1, X_2, \ldots$ denote the lifetimes of n component of a series system. We assume that X_is are continuous and iid random variables with common distribution function F and survival function $\overline{F}$. Let also $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the ordered lifetimes of the components. Then $X_{1:n}$ represents the lifetime of that series system with survival function $\overline{F}_{X_{1:n}}(x) = \overline{F}^n(x)$,

$x > 0$. The survival function of $X_{1:n} - t$ given that $X_{1:n} > t$, is $\overline{F}_{X_{1:n,t}}(x) = \left(\frac{\overline{F}(t+x)}{\overline{F}(t)}\right)^n$, where $X_{1:n} - t$ is called residual lifetime of system. Now the CRE for the residual lifetime distribution of a series system with survival function $\overline{F}_{X_{1:n,t}}(x)$ is

$$\begin{aligned} CRE(X_{1:n,t}) &= -\int_0^\infty \overline{F}_{X_{1:n,t}}(x) \log \overline{F}_{X_{1:n,t}}(x) \mathrm{d}x \\ &= -\int_t^\infty \left(\frac{\overline{F}(x)}{\overline{F}(t)}\right)^n \log \left(\frac{\overline{F}(x)}{\overline{F}(t)}\right)^n \mathrm{d}x \\ &= -\frac{1}{(\overline{F}(t))^n} \int_t^\infty (\overline{F}(x))^n \log(\overline{F}(x))^n \mathrm{d}x + n \log \overline{F}(t) \int_t^\infty \left(\frac{\overline{F}(x)}{\overline{F}(t)}\right)^n \mathrm{d}x \\ &= -\frac{1}{(\overline{F}(t))^n} \int_t^\infty (\overline{F}(x))^n \log(\overline{F}(x))^n \mathrm{d}x + n \log \overline{F}(t) m_{X_{1:n}}(t), \end{aligned} \tag{7}$$

where $m_{X_{1:n}}(t) = E(X_{1:n} - t \mid X_{1:n} > t)$ is the mean residual lifetime (MRL) of system.

Now we get some lower bounds for $CRE(X_{1:n})$. By (7) and noting that $\log \overline{F}(t) \le 0$, we conclude that

$$\begin{aligned} CRE(X_{1:n,t}) &\le -\frac{1}{(\overline{F}(t))^n} \int_t^\infty (\overline{F}(x))^n \log(\overline{F}(x))^n \mathrm{d}x \\ &\le -\frac{1}{(\overline{F}(t))^n} \int_0^\infty (\overline{F}(x))^n \log(\overline{F}(x))^n \mathrm{d}x = \frac{1}{(\overline{F}(t))^n} CRE(X_{1:n}). \end{aligned}$$

Thus, for all t,

$$CRE(X_{1:n}) \ge (\overline{F}(t))^n CRE(X_{1:n,t}).$$

By non negativity of CRE, from (7) we conclude that

$$\begin{aligned} m_{X_{1:n}}(t) &\le -\frac{1}{n|\log \overline{F}(t)|(\overline{F}(t))^n} \int_t^\infty (\overline{F}(x))^n \log(\overline{F}(x))^n \mathrm{d}x \\ &\le -\frac{1}{n|\log \overline{F}(t)|(\overline{F}(t))^n} \int_0^\infty (\overline{F}(x))^n \log(\overline{F}(x))^n \mathrm{d}x \\ &= \frac{1}{n|\log \overline{F}(t)|(\overline{F}(t))^n} CRE(X_{1:n}). \end{aligned}$$

Thus for all t,

$$CRE(X_{1:n}) \ge n|\log \overline{F}(t)|(\overline{F}(t))^n m_{X_{1:n}}(t)$$

Theorem 3.1. *Let X and Y be two positive random variables with pdfs $f(x)$ and $g(x)$ and absolutely continuous cdfs $F(x)$ and $G(x)$, respectively. Then F and G belong to the same family of distributions, but for a change in location and scale, if and only if for $t > 0$*

$$CRE(X_{1:n}, t) = CRE(Y_{1:n}, t),$$

for $n = n_j$, $j \ge 1$ such that $\sum_{j=1}^{+\infty} n_j^{-1}$ is infinite.

Proof. The necessity is trivial, hence it remains to prove the sufficiency part. If for all $n = n_j$, $j \geq 1$ such that $\sum_{j=1}^{+\infty} n_j^{-1}$ is infinite, $CRE(X_{1:n,}t) = CRE(Y_{1:n,}t)$, then by Theorem 2.2, $X \mid X > t$ and $Y \mid Y > t$ have a same distribution but for a change in location parameter, that is $f_t(x) = g_t(x + c)$, where f_t and g_t are, respectively, pdfs of $X \mid X > t$ and $Y \mid Y > t$. Thus, $f(x) = \frac{\overline{F}(t)}{\overline{G}(t)} g(x + c)$, that means F and G belong to the same family of distributions, but for a change in location and scale.

Remark 3.1. Similar result given in this article holds for last order statistic ($X_{n:n}$) if in definition of CRE, we substitute $\overline{F}$ by F, which needs to define other new uncertainty measure.

Acknowledgment

This research was supported by a grant from Ferdowsi University of Mashhad (No. MS87090BRA).

References

Arnold, B. C., Balakrishnan, N., Nagaraja, H. N. (1992). *A First Course in Order Statistics*. New York: John Wiley & Sons.

Asadi, M., Ebrahimi, N., Hamedani, G. H., Soofi, E. (2005a). Minimum dynamic discrimination information models. *J. Appl. Probab.* 42(3):643–660.

Asadi, M., Ebrahimi, N., Soofi, E. (2005b). Dynamic generalized information measures. *Statist. Probab. Lett.* 71:89–98.

Asadi, M., Zohrevand, Y. (2007). On the dynamic cumulative residual entropy. *J. Statist. Plann. Infer.* 137:931–1941.

Baratpour, S., Ahmadi, J., Arghami, N. R. (2007). Some characterization based on entropy of order statistics and record values. *Commun. Statist. Theor. Meth.* 36:47–57.

Baratpour, S., Ahmadi, J., Arghami, N. R. (2008). Some characterization based on Renyi entropy of order statistics and record values. *J. Statist. Plann. Infer.* 138:2544–2551.

Borwein, P., Erdelyi, T. (1995). *Polynomials and Polynomials Inequalities*. New York: Springer.

David, H. A., Nagaraja, H. N. (2003). *Order Statistics*. 3rd ed. New York: John Wiley & Sons.

Ebrahimi, N. (1996). How to measure uncertainty in the residual life distributions. *Sankhys, Ser. A* 58:48–57.

Kamps, U. (1998). Characterizations of distributions by recurrence relations and identities for moments of order statistics. In: Balakrishnan, N., Rao, C. R., eds. *Order Statistics: Theory and Methods*. Handbook of Statistics. Vol. 16. Amsterdam: Elsevier, pp. 291–311.

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

Rao, M., Chen. Y., Vemuri, B. C., Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE Trans. Inform. Theor.* 50(6):1220–1228.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tec. J.* 27:379–432.

Xie, M., Lai, C. D. (1996). On the increase of the lifetime redundancy. *Asia-Pacific J. of Operational Research* 13(2):171–179.