

Entropy and Residual Entropy Functions and Some Characterization Results

Ramesh C. Gupta
Department of Mathematics and Statistics
University of Maine, Orono, ME 04469-5752, U.S.A.
Ramesh.Gupta@umit.maine.edu

H.C. Taneja
Department of Applied Mathematics,
Delhi Technological University, Delhi-110042, India
hctaneja@rediffmail.com

Abstract

In this paper, we have developed conditions under which the entropy function and the residual entropy function characterize the distribution. We have also studied some stochastic comparisons based on the entropy measure and established relations between entropy comparisons and comparisons with respect to other measures in reliability. Conditions for decreasing (increasing) uncertainty in a residual life distribution are obtained. Some relations between the classes of distribution in reliability and the classes of distribution, based on the monotonic properties of uncertainty, in a residual life distribution are obtained.

Keywords and Phrases: Stochastic comparisons, IFR (DFR) distributions, Order statistics, Residual life distribution, Decreasing (increasing) uncertainty residual life.

1. Introduction

Let X be a non-negative random variable with distribution function $F(x)$ and probability density function $f(x)$. Shannon (1948) introduced a measure of uncertainty associated with distribution function $F(x)$ as

$$H(F) = -\int_0^{\infty} f(x) \log f(x) dx = -E[\log f(X)]. \quad (1)$$

The above differential entropy plays a central role in information theory and a large number of research work has been reported on it in the literature.

The entropy measure $H(F)$ measures the uncertainty associated with the distribution function $F(x)$. As $H(F)$ increases, $f(x)$ approaches a uniform distribution. Consequently, it becomes more difficult to predict an outcome of a draw from $f(x)$. In fact, a very sharply peaked distribution has a very low entropy, whereas if the probability is spread out, the entropy is much higher. In this sense, $H(F)$ is a measure of uncertainty associated with $f(x)$.

The entropy $H(F)$ has been computed for various distributions. For example Verdugo Lazo and Rathie (1978) presented a table for a large class of continuous distributions. Ahmed and Gokhale (1989) derived expression of entropy for several continuous multivariate distributions including the multivariate normal, multivariate Pareto, multivariate exponential and multivariate logistic distributions.

It is obvious that the function $H(F)$ does not characterize the distribution. In this connection Ebrahimi (2001) presented results that identify certain conditions based on the entropy under which two random variables are stochastically equal. Ebrahimi et al. (2004) also studied some stochastic comparisons in terms of entropy measures and established relations between entropy comparisons and comparisons with respect to some other measures used in reliability studies. The entropy of order statistics based on a random sample from a distribution function $F(x)$ has been studied by various authors including Ebrahimi et al. (2004), Yari and Borzadaran (2010), Tahmasebi and Behboodian (2010), Baratpour et al. (2007) and Baratpour et al. (2008). Some of these authors have characterized the parent distributions based on Shannon entropy of order statistics and record values.

If we think of X as the lifetime of a new unit, then Shannon entropy $H(F)$ can be useful for measuring the associated uncertainty. However, for a used unit $H(F)$ is no longer useful for measuring the uncertainty about the remaining lifetime of the unit. In such a situation, if the unit has survived up to time t , we consider the uncertainty of the residual life distribution as

$$H(F; t) = - \int_t^{\infty} \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)}{\bar{F}(t)} dx \quad (2)$$

$$= 1 - \int_t^{\infty} \frac{f(x)}{\bar{F}(t)} \log \lambda_F(x) dx$$

where $\bar{F}(t)$ is the survival function corresponding to $F(\cdot)$ and $\lambda_F(x)$ is its failure rate defined by $\lambda_F(x) = f(x)/\bar{F}(x)$. After the component has survived up to time t , $H(F; t)$ measures the expected uncertainty contained in the conditional density of $X - t$ given $X > t$, about the predictability of remaining lifetime of the unit. Clearly for $t = 0$, $H(F; 0) = - \int_0^{\infty} f(x) \log f(x) dx$ represents the Shannon uncertainty contained in X .

Analogous to the residual entropy, the entropy of $X | X \leq t$ called the past entropy at time t , has also drawn attention in the literature in the past decade or so. The past entropy is given by

$$\bar{H}(F; t) = - \int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} dx \quad (3)$$

$$= 1 - \int_0^t \frac{f(x)}{F(t)} \log \tau_F(x) dx, \quad (4)$$

where $\tau_F(x)$ is the reversed hazard rate function of X , given by $\tau_F(x) = f(x)/F(x)$. Lately the reversed hazard rate has attracted considerable attention, see for example Block et al. (1998), Di Crescenzo and Longobardi (2002, 2004) and Gupta and Gupta (2007).

A natural question arises: whether $H(F;t)$ and $\overline{H}(F;t)$ characterize the distribution. In this connection, Ebrahimi (1996) proved a characterization of a lifetime distribution in terms of the residual entropy. This was followed by Belzunce et al. (2004) who pointed out that the proof of Ebrahimi (1996) was not valid without some additional assumption. Under the additional assumptions Belzunce et al. (2004) proved that $H(F;t)$ characterizes the distribution. Gupta (2009) further studied this problem to determine, if indeed such assumptions are necessary. Gupta (2009) presented a general result to determine whether $H(F;t)$ and $\overline{H}(F;t)$ determine the distribution. Ebrahimi and Pellery (1995) studied some stochastic comparisons between two residual entropy measures and established relation between some other stochastic comparisons used in reliability theory. This was followed by Ebrahimi and Kirmani (1996 a) who further studied this measure and obtained some new properties. They also obtained conditions for decreasing (increasing) uncertainty in a residual life distribution.

The organization of this paper is as follows: In Section 2, we present conditions based on the entropy function under which two random variables are stochastically equal. We also study some stochastic comparisons based on the entropy measures and establish relations between entropy comparisons and comparisons with respect to some other measures used in reliability theory. Section 3 contains the entropy of order statistics based on a random sample from a distribution. Some characterization results based on order statistics are also presented. In Section 4, the concepts of residual entropy and the past entropy are introduced and some characterization results based on these functions are presented. Some stochastic comparisons based on residual entropy are presented in Section 5. Conditions for decreasing (increasing) uncertainty in a residual life distribution are obtained. Finally in Section 6, we present some conclusions and comments.

2. Characterization Based on Entropy Function

It is obvious that function $H(F)$ does not characterize a distribution function. In this section we present condition under which the function $H(F)$ characterizes the distribution. For this purpose, we present some stochastic comparison results as follows:

Let X and Y be two absolutely continuous random variables with distribution functions F and G . Let \overline{F} and \overline{G} denote their survival functions with probability density functions f and g , respectively. Then

- (i) The random variable X is said to be stochastically smaller than Y (denoted by $X \leq^{st} Y$) if $E(\phi(X)) \leq E(\phi(Y))$ for all increasing functions $\phi: R \rightarrow R$, whenever these expectations exist. This is equivalent to $\overline{F}(x) \leq \overline{G}(x)$ for all x .

(ii) The random variable X is said to be smaller than Y in the increasing convex order (denoted by $X \leq^{icx} Y$) if $E(\phi(X)) \leq E(\phi(Y))$ for all increasing convex functions $\phi: R \rightarrow R$, whenever these expectations exist. This is equivalent to $\int_t^\infty \bar{F}(x)dx \leq \int_t^\infty \bar{G}(x)dx$ for all t , provided the integrals exist.

(iii) The random variable X is said to be smaller than Y in the convex order (denoted by $X \leq^{cx} Y$) if $E(\phi(X)) \leq E(\phi(Y))$ for all convex functions $\phi: R \rightarrow R$, whenever these expectations exist. This is equivalent to $\int_t^\infty F(x)dx \leq \int_t^\infty G(x)dx$ for all t , provided the integrals exist and $E(X) = E(Y)$.

Note that

$$X \leq^{st} Y \Rightarrow X \leq^{icx} Y$$

and

$$X \leq^{cx} Y \Rightarrow X \leq^{icx} Y$$

see Shaked and Shanthikumar(2007). We now present the following result.

Theorem 2.1 (Ebrahimi (2001)) Let X and Y be two absolutely continuous random variables satisfying the conditions above with entropy function $H(F)$ and $H(G)$ respectively. Then

- (i) If (a) $H(F) \leq H(G)$, (b) $X \leq^{st} Y$ and (c) $g(x)$ is decreasing ($f(x)$ is increasing), then $X \leq^{st} Y$.
- (ii) If (a) $H(F) = H(G)$, (b) $X \leq^{icx} Y$ and (c) $\ln f(x)$ is increasing convex ($\ln g(x)$ is decreasing convex), then $X \leq^{st} Y$.
- (iii) If (a) $H(F) = H(G)$, (b) $X \leq^{cx} Y$ and (c) $\ln f(x)$ is convex ($\ln g(x)$ is concave), then $X \leq^{st} Y$.

We now present some results which ensure that $H(F) \leq H(G)$. For that we have the following definitions

1. A non-negative random variable X is said to have increasing (decreasing) failure rate IFR (DFR) if $\lambda_X(t) = f_X(t) / \bar{F}_X(t)$ is increasing (decreasing).
2. A random variable X is said to be less than Y in dispersion ordering (denoted by $X \leq^d Y$) if

$$F^{-1}(u) - F^{-1}(v) \leq G^{-1}(u) - G^{-1}(v), \quad \forall \quad 0 \leq v \leq u \leq 1.$$
3. A random variable X is said to be less than Y in likelihood ratio ordering (denoted by $X \leq^{lr} Y$) if $f_X(x) / g_Y(x)$ is non increasing in x .

4. A random variable X is said to be less than Y in entropy ordering (denoted by $X \leq^e Y$) if $H(F) \leq H(G)$.

5. A random variable X is said to be less than Y in failure rate ordering (denoted by $X \leq^{FR} Y$) if $\lambda_F(x) \geq \lambda_G(x)$ for all $x \geq 0$, where $\lambda_F(x)$ and $\lambda_G(x)$ are the failure rates of X and Y , respectively.

It is well known that

$$X \leq^d Y \Rightarrow X \leq^{st} Y$$

and

$$X \leq^{lr} Y \Rightarrow X \leq^{st} Y, \text{ and } X \leq^{lr} Y \Rightarrow X \leq^{FR} Y.$$

It is also known that $X \leq^d Y \Rightarrow X \leq^e Y$, see Oja (1981).

The following theorem gives conditions for two random variable X and Y to have entropy ordering.

Theorem 2.2 (Ebrahimi et al.(2004)) Suppose X and Y are two nonnegative random variables. If $X \leq^{st} Y$ and X is DFR, then $X \leq^e Y$.

3. Entropy of Order Statistics

Suppose X is a continuous random variable with distribution function $F_X(x)$. It is well known that $U = F_X(X)$ has a uniform distribution in $[0, 1]$. Let U_1, U_2, \dots, U_n be a random sample from a uniform distribution $[0, 1]$ and $W_1 < W_2 < \dots < W_n$ be the order statistics, then $W_i, i = 1, 2, \dots, n$ has a beta distribution with density function

$$g_i(w) = \frac{1}{B(i; n-i+1)} w^{i-1} (1-w)^{n-i}, \quad 0 \leq w \leq 1, \quad (3)$$

$$\text{where } B(a; b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

The entropy of a random variable X is given by

$$\begin{aligned} H(F) &= -\int_{-\infty}^{\infty} f_X(x) \log f_X(x) dx, \\ &= \int_0^1 \log F_X^{-1}(u) du. \end{aligned} \quad (4)$$

Hereafter the range of integration will not be shown and will be clear from the context.

Let X_1, X_2, \dots, X_n be a random sample of size n from a distribution $F_X(x)$. Let $Y_1 < Y_2 < \dots < Y_n$ be the order statistics. Then the entropy function of Y_i is given by

$$H(Y_i) = H_n(W_i) - \int f_i(y) \log f_X(y) dy \quad (5)$$

where $H_n(W_i)$ denotes the entropy of the beta distribution given by (5) and is given by

$$H_n(W_i) = \ln B(i; n-i+1) - (i-1)[\psi(i) - \psi(n+1)] - (n-i)[\psi(n-i+1) - \psi(n+1)] \quad (6)$$

where $\psi(z) = \frac{d}{dz} \ln \Gamma(z)$ is the digamma function. We now present two examples.

Example 3.1 Let X be a random variable with exponential distribution $F_X(x) = 1 - e^{(-\lambda x)}$, $x \geq 0$.

Putting $i = 1$ in (8), we get

$$H_n(W_1) = 1 - \log n - \frac{1}{n}.$$

Using (7) to calculate the entropy of first order statistics for exponential distribution and substituting for $H_n(W_1)$ we obtain

$$H(Y_1) = 1 - \log n \lambda.$$

Similarly, for $i = n$, we obtain

$$H(Y_n) = 1 - \log n - \log \lambda + \psi(n) + \gamma,$$

where $\gamma = -\psi(1) = 0.5772 \dots$ is the Euler's constant, for details see Ebrahimi et al. (2004).

Example 3.2 Suppose X has a Pareto distribution with distribution function

$$F_X(x) = 1 - \left[1 + \left(\frac{x-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right]^{-\alpha}, \quad x > \mu, \theta > 0, \gamma > 0, \alpha > 0 \quad (7)$$

and p.d.f.

$$f_X(x) = \frac{\alpha \left(\frac{x-\mu}{\theta} \right)^{\frac{1}{\gamma}-1}}{\theta \gamma \left(1 + \left(\frac{x-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right)^{\alpha+1}}, \quad x > \mu$$

Using (1) we get

$$H(F) = \ln \left(\frac{\gamma \theta}{\alpha} \right) + (\gamma - 1) + [\psi(1) - \psi(\alpha)] + \frac{\alpha + 1}{\alpha}.$$

The p.d.f of the i^{th} order statistics is given by

$$g_{i:n}(y) = n \binom{n-1}{i-1} [F(y)]^{i-1} [1 - F(y)]^{n-i} f_X(y).$$

Using the above relation for Pareto distribution given by (9), we get

$$g_{i:n}(y) = \frac{n\theta}{\theta\gamma} \binom{n-1}{i-1} \left(\frac{y-\mu}{\theta} \right)^{\frac{1}{\gamma}-1} \left(1 + \left(\frac{y-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right)^{-\alpha(n-i+1)-1} \\ \left[1 - \left(1 + \left(\frac{y-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right)^{-\alpha} \right]^{i-1}, y > \mu.$$

The entropy expression for $g_{i:n}(y)$ is given by

$$H_{i:n}(Y) = -\ln \left[\frac{n\alpha}{\theta\gamma} \binom{n-1}{i-1} \right] + \left(1 - \frac{1}{\gamma} \right) E \left[\ln \left(\frac{Y-\mu}{\theta} \right) \right] \\ + [\alpha((n-i+1)+1)+1] E \left[\ln \left(1 + \frac{Y-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right] \\ + (1-i) E \left[\ln \left(1 - \left(1 + \frac{Y-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right)^{-\alpha} \right].$$

To calculate $H_{i:n}(Y)$, we need to calculate $E \left[\ln \left(\frac{Y-\mu}{\theta} \right) \right]$, $E \left[\ln \left(1 + \frac{Y-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right]$ and

$E \left[\ln \left(1 - \left(1 + \frac{Y-\mu}{\theta} \right)^{\frac{1}{\gamma}} \right)^{-\alpha} \right]$. These expressions, after some mathematical simplifications, give

$$H_{i:n}(Y) = -\ln \left[\frac{n\alpha}{\theta\gamma} \binom{n-1}{i-1} \right] + (\gamma-1)n \binom{n-1}{i-1} \\ \sum_{l=0}^{i-1} (-1)^l \binom{i-1}{l} \left[\frac{\psi(1) - \psi(\alpha(n-i+1+l))}{n-i+1+l} + [\alpha(n-i+1)+1] \right] \\ \left[\frac{\psi(n+1) - \psi(n-i+1)}{\alpha} \right] + (i-1)[\psi(n+1) - \psi(i)],$$

for details refer to Yari and Borzadaran (2010).

4. Characterization Based on Entropy of Order Statistics

We first present two characterizations of exponential distribution based on sample minimum or maximum.

Theorem 4.1 (Baratpour et al. (2007)) The hazard rate functions $\lambda_X(x; \theta) = \alpha(\theta)$ for some positive function $\alpha(\theta)$ if and only if

$$H(X, \theta) - H(X_{i:n}, \theta) = \log n \text{ for all } n \geq 1,$$

where $X_{i:n}$ is the first order statistics.

Theorem 4.2 (Baratpour et al. 2007) The reversed hazard rate function $\tau(x, \theta) = \beta(\theta)$ for some positive function $\beta(\theta)$ if and only if

$$H(X, \theta) - H(X_{n:n}, \theta) = \log n \text{ for all } n \geq 1,$$

where $X_{n:n}$ is the maximum in a sample of size n .

We now present a general result for the characterization based on entropy of order statistics.

Theorem 4.3 (Baratpour et al. 2007) Let X and Y be two continuous random variables with distribution function $F(x)$ and $G(y)$, respectively, whose supports S_1 and S_2 have a common lower boundary b . Then for a fixed m , ($1 \leq m \leq n$), X and Y have the same distribution if and only if

$$H(X_{m:n}) = H(Y_{m:n}), \quad \forall \quad n \geq m. \quad (8)$$

Remark 1 The above result can be proved by using the Stone-Weierstrass Theorem.

Remark 2 Similar results based on entropy of record values can be obtained, see Baratpour et al. (2007).

Remark 3 For more general results based on entropy of order α , see Baratpour et al. (2008).

5. Characterization by Residual Entropy

As stated in the introduction, the residual entropy of a random variable X is given by

$$H(F; t) = 1 - \frac{1}{\bar{F}(t)} \int_t^\infty f(x) \log \lambda_F(x) dx,$$

where $\lambda_F(x)$ is the failure rate of X . The past entropy function is given by

$$\bar{H}(F; t) = 1 - \frac{1}{F(t)} \int_0^t f(x) \log \tau_F(x) dx,$$

where $\tau_F(x)$ is the reversed failure rate of X . We first present the following results due to Ebrahimi (1996).

Theorem 5.1 Let X be a nonnegative random variable with density function f and residual entropy $H(F; t) < \infty, t \geq 0$. Also assume that $H(F; t)$ is an increasing function of t , then $H(F; t)$ characterizes the distribution.

Proof: See Gupta (2009).

Remark If $H(F;t)$ is decreasing, it has been shown by Gupta (2009) that the proof given in Ebrahimi (1996) is not valid.

Remark Belzunce et al. (2004) provide another proof under the assumption that $H(F;t)$ is an increasing function of t .

The questions now remains as to whether the characterization result of the above Theorem holds without the additional assumption that $H(F;t)$ is an increasing function of t .

Before answering this question we present the following result. Consider the problem of finding a sufficient condition for the uniqueness of the solution of the initial value problem (IVP)

$$\frac{dy}{dx} = f(x, y), y(x_0) = y_0, \quad (11)$$

where f is given function of two variables whose domain is a region $D \subset R^2$, (x_0, y_0) is a specified point in D , y is the unknown function. By the solution of the IVP on an interval $I \subset R$, we mean a function $\phi(x)$ such that (i) ϕ is differentiable on I , (ii) the growth of ϕ lies in D , (iii) $\phi(x_0) = y_0$ and (iv) $\phi'(x) = f(x, \phi(x))$, for all $x \in I$. The following theorem together with other results will help in proving our characterization result.

Theorem 5.2 Let the function f be defined and continuous in a domain $D \subset R^2$, and let f satisfying a Lipschitz condition (with respect to y) in D , namely

$$|f(x, y_1) - f(x, y_2)| \leq k |y_1 - y_2|, k > 0, \quad (12)$$

for every point (x, y_1) and (x, y_2) in D . Then the function $y = \phi(x)$ satisfying the initial value problem $y' = f(x, y)$ and $\phi(x_0) = y_0, x \in I$, is unique.

Proof See Gupta and Kirmani (2008).

For any function $f(x, y)$ of two variables defined in $D \subset R^2$, we now present a sufficient condition which guarantees that the Lipschitz condition is satisfied in D .

Lemma 5.1 Suppose that the function f is continuous in a convex region $D \subset R^2$. Suppose further that $\frac{\partial f}{\partial y}$ exists and is continuous in D . Then the function f satisfies Lipschitz condition in D .

Proof See Gupta and Kirmani (1998).

Using the above results, Gupta (2009) proved the following characterization theorems.

Theorem 5.3. Let X be a non-negative continuous random variable with distribution function $F(\cdot)$ and residual entropy $H(F;t) < \infty, t \geq 0$. Then $H(F;t)$ characterizes the distribution.

Theorem 5.4 Let X be a non-negative continuous random variable with distribution function $F(\cdot)$ and past entropy function $\overline{H}(F;t) < \infty, t \geq 0$. Then $\overline{H}(F;t)$ characterizes the distribution.

Theorem 5.5 Let X be a non-negative continuous random variable with distribution function $F(\cdot)$ and cumulative residual entropy function $CRE(F;t) < \infty, t \geq 0$ where

$$CRE(F;t) = -\int_t^\infty \overline{F}_t(x) \log \overline{F}_t(x) dx, \quad (13)$$

and $\overline{F}_t = \frac{\overline{F}(x)}{\overline{F}(t)}, x > t$. Then $CRE(F;t)$ characterizes the distribution.

Remark For some examples, see Gupta (2009).

5.1 Stochastic comparison Based on Residual Entropy

We first present the following definitions.

Definition: A random variable X is said to have less uncertainty than Y if $H(F;t) < H(G;t)$ for all $t \geq 0$, and we write this as $X \leq^{\text{lu}} Y$.

Ebrahimi and Pellery (1995) proved the following:

Theorem 5.6 Let $X \leq^{\text{lr}} Y$ and $\lambda_F(x)$ or $\lambda_G(x)$ be non-increasing in x . Then $X \leq^{\text{lu}} Y$.

Example 5.1 Suppose $f(t) = \exp(-t)$ and $g(t) = e^{-t}$. It follows from the above theorem that $X \leq^{\text{lu}} Y$.

Remark: For more results and examples on the uncertainty ordering, see Ebrahimi and Pellery (1995). The above result can be strengthened for the failure rate ordering as follows.

Theorem 5.7 Let $X \leq^{\text{FR}} Y$ and let $\lambda_F(x)$ or $\lambda_G(x)$ be non increasing in x . Then $X \leq^{\text{lu}} Y$.

We now present the following definition for the monotonicity of $\overline{H}(F;t)$.

Definition: A random variable X is said to have decreasing (increasing) uncertainty of residual life (DURL (IURL)), if $H(F;t)$ is decreasing (increasing) in $t \geq 0$.

The following result gives the relationship between the monotonicity of the failure rate and the monotonicity of the residual entropy.

Theorem 5.8 *If X has increasing (decreasing) failure rate (IFR (DFR)), then X is DURL (IURL).*

Proof: See Ebrahimi (1996).

The above result can be strengthened for mean residual life function. For that we have the following definitions.

Definition The mean residual life function (MRLF) of a random variable X is defined as

$$\mu_F(t) = E(X - t | X > t) = \frac{\int_t^{\infty} \bar{F}(x) dx}{\bar{F}(t)}. \quad (14)$$

The relation between the failure rate $\lambda_F(t)$ and $\mu_F(t)$ is

$$\lambda_F(t) = \frac{1 + \mu_F'(t)}{\mu_F(t)}. \quad (15)$$

Definition A random variable X is said to have decreasing (increasing) mean residual life (DMRL(IMRL)) if $\mu_F(t)$ is decreasing (increasing) in t .

It is well known that $IFR \Rightarrow DMRL$. The following is an extension of the previous theorem for DMRL (IMRL) distributions.

Theorem 5.9 *If X has decreasing (increasing) mean residual life, then X is DURL (IURL).*

Proof See Ebrahimi and Kirmani (1996 b).

Remark Ebrahimi and Kirmani (1996) have presented an example to show that the converse of the above theorem is not true.

In many cases of practical interest one would like to know whether the DURL (IURL) property of X is inherited by a transformation of X . The following Theorem provides a partial answer.

Theorem 5.10 (a) *If X is IURL and ϕ is non-negative, increasing and convex, then $\phi(X)$ is also IURL.*

(b) *If X is DURL and ϕ is non-negative, increasing and concave, then $\phi(X)$ is also DURL.*

Proof See Ebrahimi and Kirmani (1996 b).

6. Conclusions and Comments

Entropy is an index that is used to measure dispersion, volatility risk and uncertainty. This concept was formerly introduced by Shannon (1948) in the information theory literature. If we think of X as the lifetime of a unit, then $H(F)$ can be useful for measuring the associated uncertainty. For a component, which has survived up to time t , $H(F;t)$ measures the uncertainty about the remaining life of the component. In this paper, we have developed condition under which the entropy function and the residual entropy function characterize the distribution. We have also studied some stochastic comparison based on the entropy measure and established relations between entropy comparisons and comparisons with respect to other measures in reliability. The concepts of residual entropy and the past entropy have been introduced. Conditions for decreasing (increasing) uncertainty in a residual life distribution are obtained. We hope that this survey paper will prove helpful to the researchers in studying the concept of entropy further.

References

1. Ahmed, N.A. and Gokhale, D.V. (1989). Entropy expressions and their estimators for multivariate distributions. *IEEE Transactions on Information Theory*, 35(1), 688-692.
2. Baratpour, S., Ahmadi, J., Arghami, N. R. (2007). Some characterizations based on entropy of order statistics and record values. *Communications in Statistics-Theory and Methods* 36: 47-57.
3. Baratpour, S., Ahmadi, J., Arghami, N. R. (2008). Characterizations based on Renyi entropy of order statistics and record values. *Journal of Statistical Planning and Inference*, 138, 2544-2551.
4. Belzunce, F., Navarro, J., Ruiz, J.M., Aguila, Y. (2004). Some results on residual entropy function. *Metrika* 59, 147-161.
5. Block, H. W. Savits, T. H. Singh, H. (1998). The reversed hazard rate function. *Probability in the Engineering and Informational Sciences*, 12, 69-90.
6. Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. *J. Appl. Prob.*, 39, 434-440.
7. Di Crescenzo, A. and Longobardi, M. (2004). A measure of discrimination between past lifetime distributions. *Statistics Probability Letters*, 67, 173-182.
8. Ebrahimi, N. (1996). How to measure uncertainty about residual lifetime. *Sankhya A*, 58 48-57.
9. Ebrahimi, N. (2001). Families of distributions characterized by entropy. *IEEE Transactions on Information Theory*, 47(5), 2042-2044.
10. Ebrahimi, N. and Kirmani, S.N.U.A. (1996 a). A characterisation of the proportional hazards model through a measure of discrimination between two residual life distributions. *Biomertika*, 83(1), 233-235.

11. Ebrahimi, N. and Kirmani, S.N.U.A. (1996 b). A measure of discrimination between two residual life-time distributions and its applications. *Annals of the Institute of Statistical Mathematics*, 48(2), 257-265.
12. Ebrahimi, N. and Pellerey, F. (1995). New partial ordering of survival functions based on the notion of uncertainty. *J. App. Prob.*, 32, 202-211.
13. Ebrahimi, N., Soofi, E.S., Zahedi, H. (2004). Information properties of order statistics and spacings. *IEEE Trans. Inform. Theor.* 50: 177-183.
14. Gupta, R.C. (2009). Some characterization results based on residual entropy function. *Journal of Statistical Theory and Applications*, 8(1), 45-59.
15. Gupta, R.C. and Gupta, R.D. (2007). Proportional reversed hazard rate model and its applications. *Journal of Statistical Planning and Inference*, 137, 3525-3536.
16. Gupta, R.C. and Kirmani, S.N.U.A. (2008). Characterizations based on conditional mean function. *Journal of Statistical Planning and Inference*, 138, 964-970.
17. Oja, H. (1981). On location, scale, skewness and Kurtosis of Univariate distributions. *Scandinavian Journal of Statistics*, 8, 154-168.
18. Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic orders*. Springer Verlag.
19. Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, pp. 279-423 and 623-656.
20. Tahmasebi, S. and Behboodian, J. (2010). Shannon entropy for the Feller-Pareto (FP) family of order Statistics of FP subfamilies. *Applied Mathematical Sciences*, 4 (10), 495-504.
21. Verdugo Lazo, A.G. and Rathie, P.N. (1978). On the entropy of continuous probability distributions. *IEEE Transactions on Information Theory*, IT 24(1), 120-122.
22. Yari, G. H. and Borzadaran, G. R. M. (2010). Entropy for pareto-types and its order statistics distributions. *Communications in information and systems* 10(3): 193-202.