Stochastic Comparisons of Residual Entropy of Order Statistics and Some Characterization Results

Ramesh C. Gupta [1], H.C. Taneja [2], and Richa Thapliyal [2]

[1] *Department of Mathematics and Statistics, University of Maine, Orono, ME 04469-5752, USA*

[2] *Department of Applied Mathematics, Delhi Technological University, Delhi-110042, India*

rcgupta@maine.edu, hctaneja@rediffmail.com, richa31aug@gmail.com

Received 20 January 2013

Accepted 2 February 2014

In this paper, we have presented some results for the residual and past entropies of order statistics. Results on the stochastic comparisons based on residual entropy of order statistics are presented. Characterization results for these dynamic entropies based on the sufficient condition for the uniqueness of the solution of an initial value problem have been considered.

Keywords: Order statistics, Residual entropy, Hazard rate, Past entropy, Stochastic comparisons, Characterization results.

1. Introduction

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

$$H(F) = -\int_0^\infty f(x)\ln f(x)dx = -E[\ln f(X)]\ , \tag{1.1}$$

where by convention $0\ln 0 = 0$.

The above differential entropy plays a central role in information theory. It measures the uncertainty associated with the distribution function F. It is obvious that this function $H(F)$ does not characterize the distribution function. In this connection, Ebrahimi (2001) presented some results that identify certain conditions based on the entropy under which the two random variables are stochastically equal.

The question now arises if the entropy of order statistics characterizes the distribution. For this purpose, Baratpour *et al.* (2007, 2008) explored some properties of the entropy of order statistics and record values and established some characterization results.

If we think X as life of a new unit, then $H(F)$ 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 life of the unit. In such situations, if the unit has survived upto time t, we consider the uncertainty of the residual lifetime distribution as

$$\begin{aligned} H(F;t) &= -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \ln \frac{f(x)}{\overline{F}(t)}\, dx \\ &= 1 - \frac{1}{\overline{F}(t)} \int_t^\infty f(x) \ln \lambda_F(x) dx \\ &= 1 - E[\ln \lambda_F(X) \mid X > t], \end{aligned} \tag{1.2}$$

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

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

$$\begin{aligned} \overline{H}(F;t) &= -\int_0^t \frac{f(x)}{F(t)} \ln \frac{f(x)}{F(t)} dx \\ &= 1 - \frac{1}{F(t)} \int_0^t f(x) \ln \tau_F(x) dx\,, \end{aligned} \tag{1.3}$$

where $\tau_F(x)$ is the reversed hazard rate of X given by $f(x)/F(x)$. Lately, the reversed hazard rate has drawn 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 assumptions. Under the additional assumptions Belzunce *et al.* (2004) proved that $H(F;t)$ characterize the distribution. Recently Gupta (2009) presented a general result to determine whether $H(F;t)$ or $\overline{H}(F;t)$ determine the distribution.

Let $X_1, X_2, \ldots, X_n$ be n independent and identically distributed observations from a distribution F, where $F(\cdot)$ is differentiable with a density $f(\cdot)$ which is positive in an interval and zero elsewhere. The order statistics of the sample is defined by the arrangement of $X_1, X_2, \ldots, X_n$ from the smallest to largest denoted as $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$. These 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 atleast $(n-k+1)$-out-of-n components are working. Series and parallel systems are particular cases of these system corresponding to $k=1$ and $k=n$, respectively.

In this paper, we shall explore the properties of the residual entropy of order statistics. We shall also consider the characterization results based on the entropy function of the order statistics based on residual lifetime distribution and the past life distribution. The organization of this paper is as follows: In Section 2, we present some basic results for the residual entropy and the past entropy based on order statistics together with some examples. Section 3 deals with the stochastic comparisons based on the residual entropy of order statistics. Section 4 contains some characterization results based on the entropy function of the residual lifetime distribution and the past life distribution. The proofs are based on using sufficient condition for the uniqueness of the solution of an initial value problem (IVP) encountered in the study of certain differential equation. Finally some conclusions and comments are given in Section 5.

2. 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 $U_1, U_2, \ldots, U_n$ be a random sample from a uniform distribution $[0,1]$ and $W_1 < W_2 < \cdots < W_n$ be the order statistics, then W_i, $i = 1, 2, \ldots, n$ has a beta distribution with density function

$$g_i(w) = \frac{1}{B(i, n-i+1)} w^{i-1}(1-w)^{n-i}, \quad 0 \leqslant w \leqslant 1, \tag{2.1}$$

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

The entropy of W_i is given by

$$H_n(W_i) = \ln B(i, n-i+1) - (i-1)[\psi(i) - \psi(n+1)] - (n-i)[\psi(n-i+1) - \psi(n+1)], \tag{2.2}$$

where, $\psi(z) = \frac{d\log\Gamma(z)}{dz}$ is the digamma function.

The entropy of the order statistics $X_{i:n}$ is given by

$$H(X_{i:n}) = H_n(W_i) - E_{g_i}[\ln(f(F^{-1}(W_i)))] , \tag{2.3}$$

where $g_i(w)$ is given by (2.1), refer to Ebrahimi *et al.* (2004).

The residual entropy of order statistics $X_{i:n}$ is given by

$$\begin{aligned} H(X_{i:n};t) &= -\int_t^\infty \frac{f_{i:n}(x)}{\overline{F}_{i:n}(t)} \ln\left(\frac{f_{i:n}(x)}{\overline{F}_{i:n}(t)}\right) dx \\ &= 1 - \frac{1}{\overline{F}_{i:n}(t)} \int_t^\infty f_{i:n}(x) \ln \lambda_{F_{i:n}}(x) dx , \end{aligned} \tag{2.4}$$

where $\lambda_{F_{i:n}}(x)$ is the hazard rate of the i^{th} order statistics.

Similarly, the entropy of the order statistics of the past life distribution is given by

$$\begin{aligned} \overline{H}(X_{i:n};t) &= -\int_0^t \frac{f_{i:n}(x)}{F_{i:n}(t)} \ln \frac{f_{i:n}(x)}{F_{i:n}(t)} dx \\ &= 1 - \frac{1}{F_{i:n}(t)} \int_0^t f_{i:n}(x) \ln \tau_{F_{i:n}}(x) dx , \end{aligned} \tag{2.5}$$

where $\tau_{F_{i:n}}$ is the reversed hazard rate of the i^{th} order statistics.

2.1. *Some Special cases*

In reliability engineering $(n-k+1)$-out-of-n systems are very important kind of structures. A $(n-k+1)$-out-of-n system functions if and only if at least $(n-k+1)$ components out of n components function. If $X_1, X_2, \ldots, X_n$ denote the independent lifetimes of the components of such system, then the lifetime of the system is equal to the order statistic $X_{k:n}$. The special case of $k = 1$ and n, that is for sample minima and maxima correspond to series and parallel systems respectively. Here, we calculate the residual and the past entropy of first order (sample minima) and n^{th} order (sample maxima) statistics, for an exponentially distributed random variable.

The residual entropy of $X_{1:n}$ can be obtained using (2.4) as

$$H(X_{1:n};t) = \left(\frac{n-1}{n}\right) - \ln n + \ln \overline{F}(t) - n\int_{F(t)}^{1} (1-u)^{n-1} \ln [f(F^{-1}(u))]du \, . \tag{2.6}$$

We calculate the residual entropy of the first order statistics $X_{1:n}$ for an exponentially distributed random variable with p.d.f. $f(x) = \theta e^{-\theta x}$, $\theta > 0$, $x \geqslant 0$. The p.d.f. of the i^{th} order statistics $X_{i:n}$, for $i = 1, 2, \dots, n$ is given by

$$f_{i:n}(x) = \frac{1}{B(i, n-i+1)} [F(x)]^{i-1} [1-F(x)]^{n-i} f(x), \tag{2.7}$$

for details refer to Arnold *et al.* (1992).

Putting $i = 1$ in (2.7) and using (2.6), we get that the residual entropy of the first order statistics for an exponentially distributed random variable is $1 - \ln n\theta$.

We calculate the past entropy of the n^{th} order statistics $X_{n:n}$ for an exponentially distributed random variable with p.d.f. $f(x) = \theta e^{-\theta x}$, $\theta > 0$, $x \geqslant 0$. Note that using (2.7), $f_{n:n}(x) = nF^{n-1}(x)f(x)$ and hence $F_{n:n}(x) = F^n(x)$. Putting these values in (2.8) we get

$$\begin{aligned}\overline{H}(X_{n:n};t) = &-\ln n - (n-1)\ln F(t) + \left(\frac{n-1}{n}\right) \\ &- \frac{n}{F^n(t)} \int_0^{F(t)} u^{n-1} \ln f(F^{-1}(u))du + n\ln F(t) \, .\end{aligned} \tag{2.8}$$

It can be easily seen that $f(F^{-1}(u)) = \theta(1-u)$. Putting this value in (2.9) and let $t \to \infty$, we get

$$\lim_{t\to\infty} \overline{H}(X_{n:n};t) = 1 - \ln n\theta + \gamma + \psi(n), \tag{2.9}$$

where $\gamma = -\psi(1) \approx 0.5772$, which is same as derived by Ebrahimi *et al.* (2004).

2.2. *Upper bound for dynamic entropies*

We derive the upper bound for the dynamic entropies under the condition that the pdf for the i^{th} order statistics is less than 1. Note that

$$\begin{aligned}H(X_{i:n};t) &= -\int_t^{\infty} \frac{f_{i:n}(x)}{\overline{F}_{i:n}(t)} \ln \frac{f_{i:n}(x)}{\overline{F}_{i:n}(t)} dx \\ &= \ln \overline{F}_{i:n}(t) - \frac{1}{\overline{F}_{i:n}(t)} \int_t^{\infty} f_{i:n}(x) \ln f_{i:n}(x) dx.\end{aligned}$$

We know that, for $t \geqslant 0$, $\ln \overline{F}_{i:n}(t) \leqslant 0$. Using this we get

$$\begin{aligned}H(X_{i:n};t) &\leqslant -\frac{1}{\overline{F}_{i:n}(t)} \int_t^{\infty} f_{i:n}(x) \ln f_{i:n}(x) dx \\ &\leqslant -\frac{1}{\overline{F}_{i:n}(t)} \int_0^{\infty} f_{i:n}(x) \ln f_{i:n}(x) dx \, .\end{aligned}$$

Hence

$$H(X_{i:n};t) \leqslant \frac{H(X_{i:n})}{\overline{F}_{i:n}(t)} \tag{2.10}$$

with equality when $t \to 0$.

Next, we calculate an upper bound for the past entropy. We have

$$\begin{aligned}\overline{H}(X_{i:n};t) &= -\int_0^t \frac{f_{i:n}(x)}{F_{i:n}(t)} \ln\left(\frac{f_{i:n}(x)}{F_{i:n}(t)}\right) \\ &= \ln F_{i:n}(t) - \frac{1}{F_{i:n}(t)} \int_0^t f_{i:n}(x) \ln f_{i:n}(x) dx\,.\end{aligned}$$

For $t > 0$, we have $\ln F_{i:n}(t) \leqslant 0$

$$\begin{aligned}\overline{H}(X_{i:n};t) &\leqslant -\frac{1}{F_{i:n}(t)} \int_0^t f_{i:n}(x) \ln f_{i:n}(x) dx \\ &\leqslant -\frac{1}{F_{i:n}(t)} \int_0^\infty f_{i:n}(x) \ln f_{i:n}(x) dx.\end{aligned}$$

Hence

$$\overline{H}(X_{i:n}(t)) \leqslant \frac{H(X_{i:n})}{F_{i:n}(t)}. \tag{2.11}$$

The equality is obtained when $t \to \infty$.

3. Stochastic Comparisons Based on Residual Entropy of Order Statistics

We have the following definitions:

1. A non-negative random variable X is said to have increasing (decreasing) failure rate IFR (DFR) if $\lambda_X(t) = \frac{f_X(t)}{\overline{F}_X(t)}$ is increasing (decreasing).
2. A random variable X is said to be less than Y in dispersion ordering (denoted by ($X \overset{\text{disp}}{\leqslant} Y$) if $F^{-1}(u) - F^{-1}(v) \leqslant G^{-1}(u) - G^{-1}(v), \forall 0 \leqslant v \leqslant u \leqslant 1$.
3. A random variable X is said to be less than Y in likelihood ratio ordering (denoted by $X \overset{\text{lr}}{\leqslant} Y$) if $\frac{f_X(x)}{g_Y(x)}$ is non increasing in x.
4. A random variable X is said to be less than Y in the failure rate ordering (denoted by $X \overset{\text{fr}}{\leqslant} Y$) if $\lambda_F(x) \geqslant \lambda_G(x)$, for all $x \geqslant 0$, where $\lambda_F(x)$ and $\lambda_G(x)$ are the failure rates of X and Y, respectively.
5. A random variable X is said to be less than Y in the stochastic ordering (denoted by $X \overset{\text{st}}{\leqslant} Y$) if $\overline{F}(x) \leqslant \overline{G}(x)$ for all x, where $\overline{F}(x)$ and $\overline{G}(x)$ are the survival functions of X and Y respectively.
6. A random variable X is said to be smaller than Y in residual entropy ordering (denoted by $X \overset{\text{re}}{\leqslant} Y$) if $H(F;t) \leqslant H(G;t)$, for all $t > 0$.

It is well known that $X \overset{\text{lr}}{\leqslant} Y \Rightarrow X \overset{\text{fr}}{\leqslant} Y \Rightarrow X \overset{\text{st}}{\leqslant} Y$.

We want to compare the residual entropy of order statistics. For that we present the following results.

Theorem 3.1. *Let $X \overset{lr}{\leqslant} Y$ and $\lambda_F(x)$ or $\lambda_G(x)$ be non increasing in x. Then $X \overset{re}{\leqslant} Y$.*

Proof. Refer to Ebrahimi and Pellerey (1995), Theorem 2.3. □

The above result can be strengthened for failure rate ordering as follows.

Theorem 3.2. *Let $X \overset{fr}{\leqslant} Y$ and $\lambda_F(x)$ or $\lambda_G(x)$ be non increasing in x. Then $X \overset{re}{\leqslant} Y$.*

Ebrahimi and Kirmani (1996) proved the following result for dispersive ordering.

Theorem 3.3. *Let $X \overset{disp}{\leqslant} Y$ and $\lambda_F(x)$ or $\lambda_G(x)$ be increasing in x. Then $X \overset{re}{\leqslant} Y$.*

Infact Bagai and Kochar (1986) proved the following.

Theorem 3.4. *Let $X \overset{disp}{\leqslant} Y$ and $\lambda_F(x)$ or $\lambda_G(x)$ be increasing in x. Then $X \overset{fr}{\leqslant} Y$.*

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a distribution F and let $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ be the order statistics. It is known that $X_{1:n} \overset{\mathrm{lr}}{<} X_{2:n} \overset{\mathrm{lr}}{<} \cdots \overset{\mathrm{lr}}{<} X_{n:n}$, see Chan *et al.* (1991).
Ma (1998), under some weaker conditions, extended the above result when $X_1, X_2, \ldots, X_n$ are independent, but not necessarily identically distributed.
Boland *et al.* (1994) showed that for independent but not necessarily identically distributed random variables $X_{k:n} \overset{\mathrm{fr}}{\leqslant} X_{k+1:n}$, for $k = 1, 2, \ldots, n-1$.
Let us now compare the random variable X to $X_{i:n}$, $i = 1, 2, \ldots, n$, in the likelihood ratio. For that Raqab and Amin (1996) and Khaledi and Kochar (1999) proved that $X_{i:m} \overset{\mathrm{lr}}{\leqslant} X_{j:n}$ whenever $i \leqslant j$ and $m - i \geqslant n - j$.
In particular $X_{1:n} \overset{\mathrm{lr}}{\leqslant} X$ and $X \overset{\mathrm{lr}}{\leqslant} X_{n:n}$. Using the above results we have the following theorem.

Theorem 3.5. *Suppose $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ be the order statistics based on a random sample of size n from a distribution $F(\cdot)$. Suppose X has a decreasing failure rate. Then $X \overset{re}{\leqslant} X_{n:n}$.*

Proof. As shown above $X \overset{\mathrm{lr}}{\leqslant} X_{n:n}$. Since X has a decreasing failure rate, it follows from Theorem 3.1 that $X \overset{\mathrm{re}}{\leqslant} X_{n:n}$. Note that if X has a decreasing failure rate, then $X_{n:n}$ has a decreasing failure rate, see Takahasi (1988). □

Now we compare $X_{i:n}$ and $X_{j:n}$, $i < j$. We have the following result.

Theorem 3.6. *Suppose $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ be the order statistics based on a random sample of size n from a distribution $F(\cdot)$. Suppose X has a decreasing failure rate. Then $X_{i:n} \overset{re}{\leqslant} X_{j:n}$, $i < j$.*

Proof. Using the result of Chan *et al.* (1991) we have $X_{i:n} \overset{\mathrm{lr}}{\leqslant} X_{j:n}$. This implies that $X_{i:n} \overset{\mathrm{fr}}{\leqslant} X_{j:n}$. Since X has a decreasing failure rate, $X_{i:n}$ has a decreasing failure rate, see Takahasi (1988). Using Theorem 3.2, we conclude that $X_{i:n} \overset{\mathrm{re}}{\leqslant} X_{j:n}$. □

The following result deals with the likelihood ratio ordering of order statistics of different sample sizes.

Theorem 3.7. *Suppose $X_1, X_2, \ldots, X_n, X_{max(m,n)}$ are independent and identically distributed random variables where m and n are positive integers. Then $X_{j:m} \overset{lr}{\leqslant} X_{i:n}$ whenever $j \leqslant i$ and $m - j \geqslant n - i$.*

Proof. See Shaked and Shanthikumar (2007). □

In particular $X_{n:n} \overset{\text{lr}}{\leqslant} X_{n+1:n+1}$ and $X_{1:n+1} \overset{\text{lr}}{\leqslant} X_{1:n}$. Using the above result, we have the following theorem.

Theorem 3.8. *Suppose $X_1, X_2, \ldots, X_{n+1}$ are independent and identically distributed random variables from a distribution function $F(\cdot)$. Suppose X has a decreasing failure rate. Then*

(i) $X_{n:n} \overset{re}{\leqslant} X_{n+1:n+1}$
(ii) $X_{1:n+1} \overset{re}{\leqslant} X_{1:n}$.

Proof. Using the above result, we have $X_{n:n} \overset{\text{lr}}{\leqslant} X_{n+1:n+1}$. Since X has decreasing failure rate, $X_{n:n}$ and $X_{n+1:n+1}$ have increasing failure rate. Using Theorem 3.1 we conclude that $X_{n:n} \overset{\text{re}}{\leqslant} X_{n+1:n+1}$. Similarly we can prove (ii). □

In the case of dispersive ordering, we have the following result.

Theorem 3.9. *Suppose $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ be the order statistics from a sample of size n from a distribution $F(\cdot)$. Suppose X has an increasing failure rate and $X_{i:n} \overset{disp}{<} X_{j:n}$, $i < j$. Then $X_{i:n} \overset{re}{<} X_{j:n}$.*

Proof. Since X has increasing failure rate, $X_{i:n}$ has increasing failure rate, see Takahasi (1988). Using Theorem 3.3 and the fact that $X_{i:n} \overset{\text{disp}}{<} X_{j:n}$, we can conclude that $X_{i:n} \overset{\text{re}}{<} X_{j:n}$. □

4. Some Characterization Results

In this section, we present some characterization results based on dynamic entropy of order statistics. Baratpour *et al.* (2007) have studied characterizations based on Shannon entropy of order statistics using Stone-Weistrass Theorem. We study characterizations with a different approach given below:

Consider a problem of finding sufficient condition for the uniqueness of the solution of the initial value problem (IVP)

$$\frac{dy}{dx} = f(x,y), \quad y(x_0) = y_0, \tag{4.1}$$

where f is a given function of two variables whose domain is a region $D \subset \mathbb{R}^2$, (x_0, y_0) is a specified point in D, y is the unknown function. By the solution of the IVP on an interval $I \subset \mathbb{R}$, 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) $\phi'(x) = f(x, \phi(x_0))$, for all $x \in I$. The following theorem together with other results will help in proving our characterization result.

Theorem 4.1. *Let the function f be defined and continuous in a domain $D \subset \mathbb{R}^2$, and let f satisfy a Lipschitz condition (with respect to y) in D, namely*

$$|f(x,y_1) - f(x,y_2)| \leqslant k\,|y_1 - y_2|, \quad k > 0, \tag{4.2}$$

for every point (x, y_1) and (x, y_2) in D. Then the function $y = \phi(x)$ satisfy 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 $D \subset \mathbb{R}^2$, we now present a sufficient condition which guarantees that the Lipschitz condition is satisfied in D.

Lemma 4.1. *Suppose that the function f is continuous in a convex region $D \subset \mathbb{R}^2$. 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). □

We now present our two characterization results.

Theorem 4.2. *Let X be a non-negative continuous random variable with distribution function $F(\cdot)$. Let the residual entropy of k^{th} order statistics based on a random sample of size n, be denoted by $H(F_{k:n};t) < \infty$, $t \geqslant 0$. Then $H(F_{k:n};t)$ characterizes the distribution.*

Proof. Suppose there exists two functions F_1 and F_2 such that

$$H(F_{1k:n};t) = H(F_{2k:n};t),$$

for all $t \geqslant 0$ and for all $k \leqslant n$. Then

$$H'(F_{ik:n};t) = \lambda_{F_{ik:n}}(t)[H(F_{ik:n};t) - 1 + \ln \lambda_{F_{ik}}(t)], \quad i = 1, 2. \tag{4.3}$$

where $\lambda_{F_{ik:n}}$ is the hazard rate of the k^{th} order statistics, for $i = 1, 2$. Differentiating the above equation with respect to t and simplifying, we get

$$\lambda'_{F_{ik:n}}(t) = \frac{\lambda_{F_{ik:n}}(t)}{\lambda_{F_{ik:n}}(t) + H'(F_{ik:n};t)}[H''(F_{ik:n}(t)) - \lambda_{F_{ik:n}}(t)H'_{F_{ik:n}}(t)] \tag{4.4}$$

Suppose now

$$H(F_{1k:n};t) = H(F_{2k:n};t) = g(t), \tag{4.5}$$

Then for all t,
$\lambda'_{F_{1k:n}}(t) = \psi(t, \lambda_{F_{1k:n}}(t))$ and $\lambda'_{F_{2k:n}}(t) = \psi(t, \lambda_{F_{2k:n}}(t))$, where

$$\psi(t,y) = \frac{y}{y + g'(t)}[g''(t) - y\, g'(t)].$$

It follows from Theorem 4.1 and Lemma 4.1 that $\lambda_{F_{1k:n}}(t) = \lambda_{F_{2k:n}}(t)$.
This proves our main characterization result. □

Theorem 4.3. *Let X be a non-negative continuous random variable with distribution function $F(\cdot)$. Let the past entropy of k^{th} order statistics based on random sample of size n, be denoted by $\overline{H}(F_{k:n};t) < \infty$, $t \geqslant 0$. Then $\overline{H}(F_{k:n};t) < \infty$ characterizes the distribution.*

Proof. Suppose that there are two functions F_1 and F_2 such that

$$\overline{H}(F_{1k:n};t) = \overline{H}(F_{2k:n};t)$$

for all $k \geqslant 0$ and for all $k \leqslant n$.

Then

$$\overline{H}'(F_{ik:n};t) = \tau_{F_{ik:n}}(t)[-\overline{H}(F_{ik:n};t) + 1 - \ln \tau_{F_{ik:n}}(t)], \quad i = 1, 2 \tag{4.6}$$

where $\tau_{F_{ik:n}}(t)$ is the reversed hazard rate of the k^{th} order statistics.

Differentiating (4.7) with respect to t and simplifying, we get

$$\tau'_{F_{ik:n}}(t) = \frac{\tau_{F_{ik:n}}(t)}{\overline{H}'(F_{ik:n};t) - \tau_{F_{ik:n}}(t)}[\overline{H}''(F_{ik:n};t) + \tau_{F_{ik:n}}(t)\overline{H}'(F_{ik:n};t)], \quad i = 1,2. \tag{4.7}$$

Suppose now

$$\overline{H}(F_{1k:n};t) = \overline{H}(F_{2k:n};t) = g(t), \text{ say.}$$

Then, for all $t \geqslant 0$,

$$\tau'_{F_{1k:n}}(t) = \psi(t, \tau_{F_{1k:n}}(t)), \quad \tau'_{F_{2k:n}}(t) = \psi(t, \tau_{F_{2k:n}}(t)),$$

where

$$\psi(t,y) = \frac{y}{g'(t) - y}[g''(t) + yg'(t)].$$

It follows from Theorem 4.1 and Lemma 4.1 that $\tau_{F_{1k:n}}(t) = \tau_{F_{2k:n}}(t)$. Since this result is true for all $t \geqslant 0$ and for all $k \leqslant n$, our characterization theorem is proved. □

Next we present a characterization result of the linear mean residual family of distributions based on Theorem 4.2. The linear mean residual life of a distribution is given by $\mu_F(t) = a + bt$, $a > 0$, $b > -1$. It can be verified that the corresponding failure rate is given by $\lambda_F(t) = \frac{1+b}{a+bt}$. This model has been studied among others by Hall and Wellner (1981), Oakes and Dasu (1990) and Gupta and Kirmani (1998). It includes the exponential distribution for $b = 0$ and the power distribution for $-1 < b < 0$.

Theorem 4.4. *Let X be a non negative absolutely continuous random variable with hazard rate $\lambda_F(t)$ and residual entropy $H(F;t)$. Then $H(X_{1:n};t) = 1 + \frac{b}{1+b} - \ln \lambda_{F_{1:n}}(t)$ if and only if $\mu_{F_{1:n}}(t) = a + bt$ and hazard rate $\lambda_{F_{1:n}}(t) = \frac{1+b}{a+bt}$.*

Proof. We have

$$\begin{aligned} H(X_{1:n};t) &= 1 - E[\ln \lambda_{F_{1:n}}(t)|X > t] \\ &= 1 - \ln(1+b) + \int_t^\infty \ln(a+bx)\frac{f(x)}{\overline{F}(t)}dx. \end{aligned}$$

To evaluate $\int_t^\infty \ln(a+bx)f(x)dx$, we proceed as follows.
Consider

$$\begin{aligned} E[(a+bX)^r|X > t] &= \int_t^\infty (a+bx)^r \frac{f(x)}{\overline{F}(t)}dx \\ &= \frac{(1+b)}{(b+1-br)}(a+bt)^r, \quad r < 1 + \frac{1}{b}. \end{aligned} \tag{4.8}$$

Taking the derivative of the above equation with respect to r and evaluating at $r = 0$, we get

$$E[\ln(a+bx)|X > t] = \frac{b}{1+b} + \ln(a+bt).$$

Thus

$$
\begin{aligned}
H(X_{1:n};t) &= 1-\ln(1+b)+\frac{b}{1+b}+\ln(a+bt) \\
&= 1+\frac{b}{1+b}-\ln\left(\frac{1+b}{a+bt}\right) \\
&= 1+\frac{b}{1+b}-\ln\lambda_{F_{1:n}}(t).
\end{aligned} \tag{4.9}
$$

To prove the converse, equation (4.4) gives

$$
\begin{aligned}
H'(X_{1:n};t) &= \lambda_{F_{1:n}}(t)[H(F_{1:n};t)-1+\ln\lambda_{F_{1:n}}(t)] \\
&= \lambda_{F_{1:n}}(t)[1+b-\ln\lambda_{F_{1:n}}(t)-1+\ln\lambda_{F_{1:n}}(t)].
\end{aligned}
$$

This gives

$$
\lambda'_{F_{1:n}}(t)+\frac{b}{1+b}\lambda^2_{F_{1:n}}(t)=0,
$$

whose solution is

$$
\lambda_{F_{1:n}}(t)=\frac{1+b}{a+bt}.
$$

□

5. Conclusions and comments

The concept of entropy as studied by Shannon (1948) in information theory plays a crucial role in many applications. For a system, which is observed at time t the residual and past entropies measure the uncertainty about the remaining and the past life of the distribution respectively. Entropy measures based on order statistics are crucial for measuring uncertainty in statistical modeling. The dynamic entropy measures based on order statistics characterize uniquely the underlying distribution and are also bounded above in terms of Shannon entropy of order statistics. The stochastic comparisons and the characterization results studied in this paper can be of wide interest and may find applications in statistical modeling.

References

[1] Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., "A first Course in Order Statistics", *New York: John Wiley and Sons*, (1992).

[2] Bagai, I. and Kochar, S.C., "On tail ordering and comparison of failure rates", *Communications in Statistics: Theory and Methods*, **15**, 1377–1388, (1986).

[3] Baratpour, S., Ahmadi, J., Arghami, N.R., "Some characterizations based on entropy of order statistics and record values", *Communications in Statistics: Theory and Methods*, **36**, 47–57, (2007).

[4] Baratpour, S., Ahmadi, J., Arghami, N.R., "Characterizations based on Renyi entropy of order statistics and record values", *Journal of Statistical Planning and Inference*, **138**, 2544–2551, (2008).

[5] Belzunce, F., Navarro, J., Ruiz, J.M., Aguila, Y., "Some results on residual entropy function", *Metrika*, **59**, 147–161, (2004).

[6] Block, H.W. Savits, T.H. Singh, H., "The reversed hazard rate function", *Probability in the Engineering and Informational Sciences*, **12**, 69–90, (1998).

[7] Boland, P.J., El-Neweihi, E., and Proschan, F., "Applications of hazard rate ordering in reliability and order statistics", *Journal of Applied Probability*, **31**, 180–192, (1994).

[8] Chan, W., Proschan, F. and Sethuraman, J., "Convex ordering among functions with applications to reliability and mathematical statistics", *In Block, H.W., Sampson, A.R., Savits, T.H. (eds.) Topics in Statistical Dependence. IMS Lecture Notes*, **16**, 121–134, (1991).

[9] David, H.A., Nagaraja, H.N. "Order Statistics", *New York: Wiley*, (2003).

[10] Di Crescenzo, A. and Longobardi, M., "Entropy-based measure of uncertainty in past lifetime distributions", *Journal of Applied Probability*, **39**, 434–440, (2002).

[11] Di Crescenzo, A. and Longobardi, M., "A measure of discrimination between past lifetime distributions", *Statistics Probability Letters*, **67**, 173–182, (2004).

[12] Ebrahimi, N., "How to measure uncertainty about residual lifetime", *Sankhya A*, **58**, 48–57, (1996).

[13] Ebrahimi, N., "Families of distributions characterized by entropy", *IEEE Tansactions on Information Theory*, **47**(5), 2042–2044, (2001).

[14] Ebrahimi, N. and Kirmani, S.N.U.A., "Some results on ordering of survival functions through unceratinty", *Statistics and Probability Letters*, **29**, 167–176, (1996).

[15] Ebrahimi, N. and Pellerey, F., "New partial ordering of survival functions based on the notion of uncertainty", *Journal of Applied Probability*, **32**, 202–211, (1995).

[16] Ebrahimi, N., Soofi, E.S., Zahedi, H., "Information properties of order statistics and spacings", *IEEE Transactions on Information Theory*, **50**, 177–183, (2004).

[17] Gupta, R.C., "Some characterization results based on residual entropy function", *Journal of Statistical Theory and Applications*, **8**(1), 45–59, (2009).

[18] Gupta, R.C. and Gupta, R.D., "Proportional reversed hazard rate model and its applications", *Journal of Statistical Planning and Inference*, **137**, 3525–3536, (2007).

[19] Gupta, R.C. and Kirmani, S.N.U.A., "Characterizations based on conditional mean function", *Journal of Statistical Planning and Inference*, **138**, 964–970, (2008).

[20] Hall, W.J. and Wellner, J.A., "Mean residual life", *Statistics and Related Topics, North-Holland, Amsterdam*, 169–184, (1981).

[21] Khaledi, B. and Kochar, S.C., "Stochastic orderings between distributions and their sample spaces II", *Statistics and Probability Letters*, **44**, 161–166, (1999).

[22] Ma, C., "Likelihood ratio ordering of order statistics", *Journal of Statistical Planning and Inference*, **70**, 255–261, (1998).

[23] Oakes, D. and Dasu, T., "A note on residual life", *Biometrika*, **77**, 409–410, (1990).

[24] Raqab, M.Z. and Amin, W.A., "Some ordering result on order statistics and record values", *IAPQR Transactions*, **21**, 1–8, (1996).

[25] Shaked, M. and Shanthikumar, J.G., "Stochastic Orders", *Springer Verlag*, (2007).

[26] Shannon, C.E., "A mathematical theory of communication", *Bell System Technical Journal*, 279–423 and 623–656, (1948).

[27] Takahasi, K., "A note on hazard rates of order statistics", *Communications in Statistics: Theory and Methods*, **17**(12), 4133–4136, (1988).