



Communications in Statistics - Theory and Methods

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/Ista20>

Characterizations of Life Distributions Using Conditional Expectations of Doubly (Interval) Truncated Random Variables

S. M. Sunoj^a, P. G. Sankaran^a & S. S. Maya^a

^a Department of Statistics, Cochin University of Science and Technology, Cochin, India

Published online: 27 Apr 2009.

To cite this article: S. M. Sunoj, P. G. Sankaran & S. S. Maya (2009) Characterizations of Life Distributions Using Conditional Expectations of Doubly (Interval) Truncated Random Variables, *Communications in Statistics - Theory and Methods*, 38:9, 1441-1452, DOI: [10.1080/03610920802455001](https://doi.org/10.1080/03610920802455001)

To link to this article: <http://dx.doi.org/10.1080/03610920802455001>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Characterizations of Life Distributions Using Conditional Expectations of Doubly (Interval) Truncated Random Variables

S. M. SUNOJ, P. G. SANKARAN, AND S. S. MAYA

Department of Statistics, Cochin University of Science and Technology,
Cochin, India

In this article, we study reliability measures such as geometric vitality function and conditional Shannon's measures of uncertainty proposed by Ebrahimi (1996) and Sankaran and Gupta (1999), respectively, for the doubly (interval) truncated random variables. In survival analysis and reliability engineering, these measures play a significant role in studying the various characteristics of a system/component when it fails between two time points. The interrelationships among these uncertainty measures for various distributions are derived and proved characterization theorems arising out of them.

Keywords Generalized failure rate; Geometric vitality function; Shannon's measure of uncertainty.

Mathematics Subject Classification 62E10; 62N05.

1. Introduction

The standard practice in modeling statistical data is either to derive the appropriate model based on the physical properties of the system or to choose a flexible family of distributions and then find a member of the family that is appropriate to the data. In both situations it would be helpful if we find characterization theorems that explain the distribution using important measures of indices. For example, in reliability theory and survival analysis, identification of probability models is often achieved through studying the characteristics of measures such as failure rate, mean residual life, vitality function, coefficient of variation, etc. There are several investigations concerning these reliability measures to characterize different probability models (see Gupta and Kirmani, 2000; Kotz and Shanbhag, 1980; Nair and Sankaran, 1991). Similarly, considerable attention has also been paid to the identification probability models based on conditional expectations of left and right truncated data (see Navarro and Ruiz, 2004; Navarro et al., 1998a; Zoroa et al., 1990; and

Received October 8, 2007; Accepted September 5, 2008

Address correspondence to S. M. Sunoj, Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India; E-mail: smsunoj@cusat.ac.in

the references therein). Motivated this, in the present note, an attempt is made to derive some new characterizations to certain probability distributions and families of distributions using some important information measures which are useful for modeling and analysis of lifetime data.

Similar to vitality function, geometric vitality function has also been found a useful tool in the analysis of lifetime data (see Nair and Rajesh, 2000). For a non negative random variable (rv) X representing the lifetime of a component with an absolutely continuous distribution function $F(t)$ and $E(\log X) < \infty$, the geometric vitality function of a left truncated rv is given by

$$\log G(t) = E(\log X | X > t). \quad (1.1)$$

In reliability theory, (1.1) gives the geometric mean of lifetimes of components which has survived t units of time. For various properties and applications of (1.1), one could refer to Nair and Rajesh (2000).

In modeling and analysis of lifetime data, it is also well known that a basic uncertainty measure of a rv X with probability density function $f(t)$ is the Shannon information measure (see Shannon, 1948) given by

$$H(X) = -E(\log f(X)) = -\int_0^{\infty} f(x) \log f(x) dx. \quad (1.2)$$

Clearly, (1.2) gives the expected uncertainty contained in $f(t)$ about the predictability of an outcome of X . Motivated by this, Ebrahimi (1996) modified (1.2) to measure uncertainty in the residual lifetime distribution, referred as the residual Shannon's measure of uncertainty as follows. Let X be a non negative rv representing the lifetime of a unit or a system, then the rv $X - t | X \geq t$ represents the residual life of a unit with age t , the residual Shannon's measure of uncertainty is defined as

$$\begin{aligned} H(t) = H(X - t | X > t) &= -\int_t^{\infty} \frac{f(x)}{\bar{F}(t)} \log \left(\frac{f(x)}{\bar{F}(t)} \right) dx \\ &= 1 - \frac{1}{\bar{F}(t)} \int_t^{\infty} f(x) \log h(x) dx \end{aligned} \quad (1.3)$$

where $h(t) = f(t)/\bar{F}(t)$ is the failure rate. It is well known that $H(t)$ has much relevance in characterizing, ordering, and classifying life distributions according to the behavior of $H(t)$ (see Asadi and Ebrahimi, 2000; Belzunce et al., 2004; Ebrahimi and Pellerey, 1995; Nair and Rajesh, 1998). Analogously, Di Crescenzo and Longobardi (2002) recently introduced a useful measure

$$\bar{H}(t) = H(t - X | X < t) = -\int_0^t \frac{f(x)}{F(t)} \log \left(\frac{f(x)}{F(t)} \right) dx \quad (1.4)$$

known as a measure of past entropy, to measure the uncertainty in the inactivity time $(t - X | X < t)$.

In continuation of the residual Shannon's measure of uncertainty proposed by Ebrahimi (1996), Sankaran and Gupta (1999) introduced another conditional measure of uncertainty, which is also quite useful in the study of aging pattern of

the system. For a non negative rv X , the conditional measure of uncertainty due to Sankaran and Gupta (1999) is given by

$$M(t) = -E(\log f(X) | X > t) = -\frac{1}{\bar{F}(t)} \int_t^\infty f(x) \log f(x) dx. \tag{1.5}$$

Later, Rajesh and Nair (2000) studied this concept and proved some characterizations of certain probability distributions.

In survival studies and in life testing, often one has information about the lifetime only between two time points. That is, individuals whose event time lies within a certain time interval are only observed. Thus, an individual whose event time is not in this interval is not observed and therefore information on the subjects outside this interval is not available to the investigator. Accordingly, Kotlarski (1972) studied the conditional expectation for the doubly (interval) truncated random variables. Later, Navarro and Ruiz (1996) generalized the failure rate and the conditional expectation to the doubly truncated random variables. It is shown that generalized failure rate (GFR) and the conditional expectation for doubly truncated random variables determine the distribution uniquely. For the various relationships between GFR and conditional expectation, characterizations and their applications we may refer to Ruiz and Navarro (1996), Betensky and Martin (2003), Navarro and Ruiz (2004), Sankaran and Sunoj (2004), and Bairamov and Gebizlioglu (2005).

The aim of the present article is to further investigate these reliability measures to the doubly truncated rv's. In Sec. 2, we define geometric vitality function and two measures of uncertainty for the doubly truncated rv's and examine its properties and different relationships. Finally in Sec. 3, some of the existing characterizations by relationships between these measures of uncertainty and GFR functions are extended to model various probability distributions and families of distributions.

2. Definitions and Properties

2.1. Geometric Vitality Function

If X is a non negative rv the geometric vitality function for doubly truncated rv $(X | t_1 \leq X \leq t_2)$, where $(t_1, t_2) \in D = \{(u, v) \in \mathfrak{R}^{+2}; F(u) < F(v)\}$ is defined as

$$G(t_1, t_2) = E(\log X | t_1 < X < t_2), \tag{2.1}$$

which gives the geometric mean life of a rv truncated at two points t_1 and t_2 . It is clear that when $t_2 \rightarrow \infty$ (2.1) reduces to (1.1). The following properties are immediate from the definition (2.1),

$$\lim_{\substack{t_1 \rightarrow 0 \\ t_2 \rightarrow \infty}} G(t_1, t_2) = E(\log X), \quad \text{and} \tag{2.2}$$

$$m(t_1, t_2) \geq G(t_1, t_2) \quad \text{for all } (t_1, t_2) \in D, \tag{2.3}$$

where $m(t_1, t_2) = E(X | t_1 < X < t_2)$. Denoting the GFR functions as $h_1(t_1, t_2) = \frac{f(t_1)}{\bar{F}(t_1) - \bar{F}(t_2)}$ and $h_2(t_1, t_2) = \frac{f(t_2)}{\bar{F}(t_1) - \bar{F}(t_2)}$ of Navarro and Ruiz (1996), (2.1) is related to

$h_i = h_i(t_1, t_2); i = 1, 2$ as

$$h_1(t_1, t_2) = \frac{\frac{\partial G(t_1, t_2)}{\partial t_1}}{G(t_1, t_2) - \log t_1} \tag{2.4}$$

and

$$h_2(t_1, t_2) = \frac{\frac{\partial G(t_1, t_2)}{\partial t_2}}{\log t_2 - G(t_1, t_2)}, \tag{2.5}$$

for all $(t_1, t_2) \in D$. Table 1 provides the relationship between geometric vitality function and GFR functions for various lifetime models.

Theorem 2.1. *The geometric vitality function determines distribution uniquely.*

Proof. The proof follows Theorem 2.6 in Navarro et al. (1998b) since $G(t_1, t_2) = m_{\log X}(\log t_1, \log t_2)$ where $m_Z(t_1, t_2) = E(Z | t_1 < Z < t_2)$.

Remark 2.1. In the absolutely continuous case it can also be proved by using (2.4) and (2.5).

2.2. Measure of Uncertainty

Defining a rv $(X | t_1 < X < t_2)$ which represents the lifetime of a unit which fails between t_1 and t_2 where $(t_1, t_2) \in D$, a measure of uncertainty for the doubly

Table 1
Relationships between geometric vitality function and GFR functions

Distribution	$\bar{F}(x)$	$G(t_1, t_2)$
Exponential	$\exp(-\lambda x); x > 0,$ $\lambda > 0$	$\frac{1}{\lambda} [h_1(t_1, t_2) \log t_1 - h_2(t_1, t_2) \log t_2 + R(t_1, t_2)]$
Finite range	$(1 - ax)^b;$ $0 < x < \frac{1}{a},$ $b > 0, a > 0$	$\frac{1}{ab} [(1 - at_1)h_1(t_1, t_2) \log t_1$ $- (1 - at_2)h_2(t_1, t_2) \log t_2 + R(t_1, t_2) - a]$
Pareto II	$(1 + px)^{-q};$ $x > 0,$ $p > 0, q > 0$	$\frac{1}{pq} [(1 + pt_1)h_1(t_1, t_2) \log t_1$ $- (1 + pt_2)h_2(t_1, t_2) \log t_2 + R(t_1, t_2) + p]$
Power	$1 - (x/\alpha)^\beta;$ $0 \leq x \leq \alpha,$ $\alpha > 0, \beta > 0$	$\frac{1}{\beta} [t_2 h_2(t_1, t_2) \log t_2 - t_1 h_1(t_1, t_2) \log t_1 - 1]$
Pareto I	$(a/x)^b; x > a,$ $a > 0, b > 0$	$\frac{1}{b} [t_1 h_1(t_1, t_2) \log t_1 - t_2 h_2(t_1, t_2) \log t_2 - 1]$

where $R(t_1, t_2) = E[\frac{1}{X} | t_1 < X < t_2]$.

truncated rvs is given by

$$\begin{aligned}
 H(t_1, t_2) &= H(X | t_1 < X < t_2) = - \int_{t_1}^{t_2} \frac{f(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \log \left(\frac{f(x)}{\bar{F}(t_1) - \bar{F}(t_2)} \right) dx \\
 &= 1 - \frac{1}{\bar{F}(t_1) - \bar{F}(t_2)} \int_{t_1}^{t_2} f(x)(\log h(x)) dx \\
 &\quad + \frac{1}{\bar{F}(t_1) - \bar{F}(t_2)} [\bar{F}(t_2)(\log \bar{F}(t_2)) - \bar{F}(t_1)(\log \bar{F}(t_1))] + \log(\bar{F}(t_1) - \bar{F}(t_2))
 \end{aligned}
 \tag{2.6}$$

which can also be written as

$$\begin{aligned}
 H(t_1, t_2) &= 1 - \frac{1}{\bar{F}(t_1) - \bar{F}(t_2)} \int_{t_1}^{t_2} f(x)(\log \bar{h}(x)) dx \\
 &\quad - \frac{1}{\bar{F}(t_1) - \bar{F}(t_2)} [F(t_2)(\log F(t_2)) - F(t_1)(\log F(t_1))] + \log(F(t_2) - F(t_1))
 \end{aligned}
 \tag{2.7}$$

where $\bar{h}(x) = \frac{f(x)}{F(x)}$, the reversed failure rate function (see Block et al., 1998). By using (1.3), (1.4), and (2.6), Shannon’s measure (1.2) can be decomposed as

$$\begin{aligned}
 H &= F(t_1)\bar{H}(t_1) + (\bar{F}(t_1) - \bar{F}(t_2))H(t_1, t_2) + \bar{F}(t_2)H(t_2) \\
 &\quad - [F(t_1) \log F(t_1) + (\bar{F}(t_1) - \bar{F}(t_2)) \log(\bar{F}(t_1) - \bar{F}(t_2)) + \bar{F}(t_2) \log \bar{F}(t_2)].
 \end{aligned}
 \tag{2.8}$$

The identity (2.8), which is similar to the one given in Di Crescenzo and Longobardi (2002), can be interpreted in the following way. The uncertainty about the failure of an item can be decomposed into four parts: (i) the uncertainty about the failure time in $(0, t_1)$ given that the item has failed before t_1 ; (ii) the uncertainty about the failure time in the interval (t_1, t_2) given that the item has failed after t_1 but before t_2 ; (iii) the uncertainty about the failure time in $(t_2, +\infty)$ given that it has failed after t_2 ; and (iv) the uncertainty of the rv which determines if the item has failed before t_1 or in between t_1 and t_2 or after t_2 .

On differentiating $H(t_1, t_2)$ with respect to t_1 and t_2 , we get

$$\frac{\partial H(t_1, t_2)}{\partial t_1} = h_1(t_1, t_2)(\log h_1(t_1, t_2) + H(t_1, t_2) - 1)
 \tag{2.9}$$

and

$$\frac{\partial H(t_1, t_2)}{\partial t_2} = h_2(t_1, t_2)(1 - \log h_2(t_1, t_2) - H(t_1, t_2)).
 \tag{2.10}$$

When $H(t_1, t_2)$ is increasing in t_1 and in t_2 , then, (2.9) and (2.10) together imply

$$1 - \log h_1(t_1, t_2) \leq H(t_1, t_2) \leq 1 - \log h_2(t_1, t_2).
 \tag{2.11}$$

Thus, when the uncertainty measure is increasing, then it lies between $(1 - \log h_1(t_1, t_2))$ and $(1 - \log h_2(t_1, t_2))$. We can also write the bounds in (2.11) as

$$h_2(t_1, t_2) \leq \exp(1 - H(t_1, t_2)) \leq h_1(t_1, t_2).$$

Table 2 provides the relationships between $H(t_1, t_2)$, the conditional expectation $m(t_1, t_2) = E(X | t_1 < X < t_2)$ and GFR functions $h_i(t_1, t_2); i = 1, 2$ for various distributions.

2.3. Conditional Measure of Uncertainty

As an extension of (1.5), we define the conditional measure of uncertainty for the doubly truncated rv as

$$\begin{aligned} M(t_1, t_2) &= -E[\log f(X) | t_1 < X < t_2] \\ &= \frac{-1}{\bar{F}(t_1) - \bar{F}(t_2)} \int_{t_1}^{t_2} f(x)(\log f(x))dx, \end{aligned} \tag{2.12}$$

where $(t_1, t_2) \in D$. Using (2.12), $M(t_1, t_2)$ can be easily related to $H(t_1, t_2)$ through the relation

$$M(t_1, t_2) = H(t_1, t_2) - \log(\bar{F}(t_1) - \bar{F}(t_2)). \tag{2.13}$$

Differentiating (2.13) with respect to t_1 and t_2 , respectively, provide the relationships with GFR functions, which are given by

$$\frac{\partial M(t_1, t_2)}{\partial t_1} = \frac{\partial H(t_1, t_2)}{\partial t_1} + h_1(t_1, t_2)$$

Table 2
Relationships between $H(t_1, t_2)$, the conditional expectation, and GFR functions for various distributions

Distribution	$\bar{F}(x)$	$H(t_1, t_2)$
Exponential	$\exp(-\lambda x); x > 0,$ $\lambda > 0$	$\lambda m(t_1, t_2) - \lambda t_2 - \log h_2(t_1, t_2)$ or $\lambda m(t_1, t_2) - \lambda t_1 - \log h_1(t_1, t_2)$
Finite range	$(1 - ax)^b; 0 < x < \frac{1}{a},$ $b > 0, a > 0$	$-(b - 1)E[\log(1 - ax) t_1 < X < t_2]$ $-\log \frac{ab}{((1-at_1)^b - (1-at_2)^b)}$
Pareto II	$(1 + px)^{-q}; x > 0,$ $p > 0, q > 0$	$(q + 1)E[\log(1 + px) t_1 < X < t_2]$ $-\log \frac{pq}{((1+pt_1)^{-q} - (1+pt_2)^{-q})}$
Power	$1 - (x/\alpha)^\beta; 0 \leq x \leq \alpha,$ $\alpha > 0, \beta > 0$	$1 + \log G(t_1, t_2) + t_1 h_1(t_1, t_2) \log(t_1/\alpha)$ $-t_2 h_2(t_1, t_2) \log(t_2/\alpha) - \log \frac{\beta}{((t_2/\alpha)^\beta - (t_1/\alpha)^\beta)}$
Pareto I	$(a/x)^b; x > a,$ $a > 0, b > 0$	$1 + \log G(t_1, t_2) + t_2 h_2(t_1, t_2) \log(a/t_2)$ $-t_1 h_1(t_1, t_2) \log(a/t_1) - \log \frac{b}{((a/t_1)^b - (a/t_2)^b)}$

and

$$\frac{\partial M(t_1, t_2)}{\partial t_2} = \frac{\partial H(t_1, t_2)}{\partial t_2} - h_2(t_1, t_2).$$

The relationships between the conditional measure of uncertainty for doubly truncated random variables and GFR functions for some useful probability models are given in Table 3.

3. Characterizations

In this section, we prove certain characterization theorems for some important life distributions and for certain family of distributions using GFR functions, geometric vitality function (2.1), and conditional Shannon’s measure of uncertainties (2.6) and (2.12).

Theorem 3.1. *Let X be a rv with support $(0, \infty)$ admitting an absolutely continuous distribution function $F(x)$. Then a relationship of the form*

$$G(t_1, t_2) = \frac{1}{k} [(1 + Ct_1)h_1(t_1, t_2) \log t_1 - (1 + Ct_2)h_2(t_1, t_2) \log t_2 + R(t_1, t_2) + C], \tag{3.1}$$

where $R(t_1, t_2) = E(\frac{1}{X} | t_1 < X < t_2)$ and k, C are constants holds for all $(t_1, t_2) \in D$ if and only if X follows exponential with $\bar{F}(x) = \exp(-\lambda x); x > 0, \lambda > 0$ for $C = 0$,

Table 3
Relationships between $M(t_1, t_2)$ and GFR functions for various distributions

Distribution	$\bar{F}(x)$	$M(t_1, t_2)$
Exponential	$\exp(-\lambda x);$ $x > 0, \lambda > 0$	$\lambda m(t_1, t_2) - \log \lambda$
Finite range	$(1 - ax)^b;$ $0 < x < \frac{1}{a},$ $b > 0, a > 0$	$-\log ab - (b + 1)E[\log(1 - aX) t_1 < X < t_2]$
Pareto II	$(1 + px)^{-q};$ $x > 0,$ $p > 0, q > 0$	$(q + 1)E[\log(1 + pX) t_1 < X < t_2] - \log pq$
Power	$1 - (x/\alpha)^\beta;$ $0 \leq x \leq \alpha,$ $\alpha > 0, \beta > 0$	$(\frac{\beta-1}{\beta})[1 + t_1 h_1 \log t_1 - t_2 h_2 \log t_2 + \log \beta - \beta \log \alpha]$
Pareto I	$(a/x)^b;$ $x > a,$ $a > 0, b > 0$	$(\frac{b+1}{b})[1 + t_1 h_1 \log t_1 - t_2 h_2 \log t_2 - \log b - b \log a]$
Weibull	$\exp(-x^p);$ $x > 0, p > 0$	$-\log p - (p - 1) \log G(t_1, t_2) + E[X^p t_1 < X < t_2]$
Rayleigh	$\exp(-x^2); x > 0$	$-\log 2 - \log G(t_1, t_2) + E[X^2 t_1 < X < t_2]$

Pareto distribution with $\bar{F}(x) = (1 + px)^{-q}$; $x > 0, p > 0, q > 0$ for $C > 0$ and finite range distribution with $\bar{F}(x) = (1 - ax)^b$; $0 < x < \frac{1}{a}, b > 0$ for $C < 0$.

Proof. Assume that the relation (3.1) holds. Then from the definitions of $R(t_1, t_2)$, $h_i(t_1, t_2)$, and $G(t_1, t_2)$, we can write (3.1) as

$$\int_{t_1}^{t_2} f(x) \log x \, dx = \frac{1}{k} \left[(1 + Ct_1)f(t_1) \log t_1 - (1 + Ct_2)f(t_2) \log t_2 + \int_{t_1}^{t_2} \frac{1}{x} f(x) dx + C(\bar{F}(t_1) - \bar{F}(t_2)) \right]. \quad (3.2)$$

Differentiating (3.2) with respect to t_i , $i = 1, 2$ and simplifying we get

$$\frac{f'(t_i)}{f(t_i)} = -\frac{(k + C)}{(1 + Ct_i)}, \quad \text{for } (t_1, t_2) \in D$$

or

$$\frac{d}{dt} \log f(t) = -\frac{(k + C)}{(1 + Ct)}. \quad (3.3)$$

From (3.3) we get that X follows exponential, Pareto II and finite range distributions according as $C = 0$, $C > 0$, and $C < 0$. The converse part is obtained from Table 1.

Theorem 3.2. Let X be a rv with support $(0, \infty)$ admitting an absolutely continuous distribution function $F(x)$, a relationship of the form

$$M(t_1, t_2) - \lambda m(t_1, t_2) = k, \quad (3.4)$$

where $\lambda > 0$ and k is a constant, holds for all $(t_1, t_2) \in D$ if, and only if, X follows exponential distribution with $\bar{F}(x) = \exp(-\lambda x)$; $x > 0$.

Proof. Assume (3.4) holds. From the definition (2.12) of $M(t_1, t_2)$, we can write

$$-\int_{t_1}^{t_2} f(x)(\log f(x)) dx - \lambda \int_{t_1}^{t_2} x f(x) dx = k(\bar{F}(t_1) - \bar{F}(t_2)). \quad (3.5)$$

Differentiating (3.5) with respect to t_i , $i = 1, 2$ gives

$$\log f(t_i) = -k - \lambda t_i, \quad i = 1, 2,$$

or $f(t) = \lambda \exp(-\lambda t)$, which provides the result. For converse part, see Table 3.

Theorem 3.3. Let X be a rv with support $(0, \infty)$ admitting an absolutely continuous distribution function $F(x)$, a relationship of the form

$$M(t_1, t_2) - (c + 1)G(t_1, t_2) = k, \quad (3.6)$$

where k and c are constants and $c > 0$, holds for $a < t_1 < t_2$ with $F(t_1) < F(t_2)$ if and only if X follows a Pareto type I with $\bar{F}(x) = \left(\frac{a}{x}\right)^c$; $x > a, a > 0$.

Proof. Assume (3.6) holds, from the definition of $M(t_1, t_2)$

$$-\int_{t_1}^{t_2} f(x)(\log f(x))dx - (c + 1) \int_{t_1}^{t_2} f(x) \log x dx = k(\bar{F}(t_1) - \bar{F}(t_2)). \tag{3.7}$$

Differentiating (3.7) with respect to $t_i, i = 1, 2$ and simplifying we get $f(x) = kx^{-(c+1)}$, which implies X follows a Pareto Type I. The converse part is obtained from Tables 1 and 3.

Theorem 3.4. *Let X be a rv with support $(0, \infty)$ admitting an absolutely continuous distribution function $F(x)$, a relationship of the form*

$$M(t_1, t_2) + (\beta - 1)G(t_1, t_2) = k, \quad \text{a constant} \tag{3.8}$$

holds for $0 < t_1 < t_2 < \alpha$ with $F(t_1) < F(t_2)$ and $\beta > 1$ if and only if X follows Power distribution with $\bar{F}(x) = 1 - (\frac{x}{\alpha})^\beta; 0 \leq x \leq \alpha, \alpha > 0, \beta > 0.$

The proof is similar to that of Theorem 3.3.

Now we prove a characterization theorem using $M(t_1, t_2)$ for one-parameter log exponential family defined by

$$f(x) = \frac{C(x)x^\theta}{A(\theta)}, \quad x \in (0, \infty), \quad \theta > 0 \tag{3.9}$$

where $C(x)$ is non negative function of x and $A(\theta)$ is non negative function of θ satisfying $A(\theta) = \int_0^\infty x^\theta C(x)dx.$

Theorem 3.5. *Let X be a rv with support $(0, \infty)$ admitting an absolutely continuous distribution function $F(x)$, then the distribution of X belongs to one-parameter log exponential family if and only if*

$$M(t_1, t_2) = \log A(\theta) - \theta G(t_1, t_2) - m_c(t_1, t_2), \tag{3.10}$$

where $m_c(t_1, t_2) = E[\log C(X) | t_1 < X < t_2], (t_1, t_2) \in D.$

Proof. Assume (3.10) holds. From the definition (2.12), we get

$$\begin{aligned} -\int_{t_1}^{t_2} f(x) \log f(x)dx &= (\bar{F}(t_1) - \bar{F}(t_2)) \log A(\theta) - \theta \int_{t_1}^{t_2} f(x) \log x dx \\ &+ \int_{t_1}^{t_2} f(x)(\log C(x))dx. \end{aligned} \tag{3.11}$$

Differentiating (3.11) with respect to $t_i, i = 1, 2$ and simplifying, we get (3.9). The proof of the second part of the theorem is direct.

In the following, we present a characterization theorem using $M(t_1, t_2)$ for the one-parameter exponential family defined by

$$f(x) = \frac{a(x)\theta^x}{b(\theta)}, \quad x \in (0, \infty), \quad \theta > 0, \tag{3.12}$$

where $a(x)$ is a non negative function of x and $b(\theta)$ is a non negative function of θ satisfying $b(\theta) = \int_0^\infty a(x)\theta^x dx$.

Theorem 3.6. Let X be a rv with support $(0, \infty)$ admitting an absolutely continuous distribution function $F(x)$, the relationship

$$M(t_1, t_2) = \log b(\theta) - m(t_1, t_2) \log \theta - m_a(t_1, t_2) \quad (3.13)$$

where $m_a(t_1, t_2) = E[\log a(X) | t_1 < X < t_2]$, $(t_1, t_2) \in D$, holds if and only if the distribution of X belongs to one-parameter exponential family (3.12).

Proof. The proof is similar to that of the Theorem 3.5.

Length-biased sampling is frequently a convenient technique for the collection of positive-valued or lifetime data. Such problems may occur in clinical trials, reliability theory, survival analysis, and population studies, where a proper sampling frame is absent (see Navarro et al., 2001; Rao, 1965; Sunoj, 2004; Sunoj and Maya, 2006 and the references therein). Let X be a non negative rv denoting the life length of a component with probability density function (pdf) $f(t)$. Then a rv Y with density $f^l(t) = \frac{t}{\mu} f(t)$, where $\mu = E(X) < \infty$, is said to have length-biased distribution corresponding to X . Then the geometric vitality function of the length-biased model is given by

$$\begin{aligned} G^l(t_1, t_2) &= E[\log Y | t_1 < Y < t_2], \quad (t_1, t_2) \in D \\ &= \frac{1}{m(t_1, t_2)(\bar{F}(t_1) - \bar{F}(t_2))} \int_{t_1}^{t_2} x f(x) \log x dx, \quad (t_1, t_2) \in D \\ &= \frac{1}{m(t_1, t_2)(\bar{F}(t_1) - \bar{F}(t_2))} \left[t_1 \bar{F}(t_1) \log t_1 - t_2 \bar{F}(t_2) \log t_2 \right. \\ &\quad \left. + \int_{t_1}^{t_2} \bar{F}(x) dx + \int_{t_1}^{t_2} \bar{F}(x) \log x dx \right], \quad (t_1, t_2) \in D \end{aligned}$$

In the following theorem, we characterize the exponential distribution using the functional relationship between geometric vitality functions of the length-biased and original rv's and GFR functions.

Theorem 3.7. For a non negative rv X , the relationship

$$\lambda m(t_1, t_2) G^l(t_1, t_2) - G(t_1, t_2) = 1 + t_1 \log t_1 h_1(t_1, t_2) - t_2 \log t_2 h_2(t_1, t_2) \quad (3.14)$$

holds for $(t_1, t_2) \in D$ if and only if X follows an exponential distribution.

Proof. Suppose that the relationship (3.14) holds. Then by definition,

$$\begin{aligned} &\frac{\lambda}{\bar{F}(t_1) - \bar{F}(t_2)} \int_{t_1}^{t_2} x f(x) \log x dx - \frac{1}{\bar{F}(t_1) - \bar{F}(t_2)} \int_{t_1}^{t_2} f(x) \log x dx \\ &= 1 + t_1 \frac{f(t_1)}{\bar{F}(t_1) - \bar{F}(t_2)} \log t_1 - t_2 \frac{f(t_2)}{\bar{F}(t_1) - \bar{F}(t_2)} \log t_2. \end{aligned} \quad (3.15)$$

Multiply both sides of (3.15) by $(\bar{F}(t_1) - \bar{F}(t_2))$ and on differentiation with respect to t_i , $i = 1, 2$, yields the required result. The converse part is straightforward.

Acknowledgment

The third author wishes to acknowledge the University Grants Commission, India for the financial assistance for carrying out this research work.

References

- Asadi, M., Ebrahimi, N. (2000). Residual entropy and its characterizations in terms of hazard function and mean residual life function. *Statist. Probab. Lett.* 49:263–269.
- Bairamov, I., Gebzlioglu, O. (2005). On the characterizations of distributions through the properties of conditional expectations of order statistics. In: *Advances on Models, Characterizations and Applications*. Boca Raton, FL: Chapman and Hall/CRC Press.
- Belzunce, F., Navarro, J., Ruiz, J. M., del Aguila, Y. (2004). Some results on residual entropy function. *Metrika* 59:47–161.
- Betensky, R. A., Martin, E. C. (2003). Commentary: Failure-rate functions for doubly truncated random variables. *IEEE Trans. Reliab.* 52(1):7–8.
- Block, H. W., Savits, T. H., Singh, H. (1998). The reversed hazard rate function. *Probab. Eng. Inform. Sci.* 12:69–90.
- Di Crescenzo, A., Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. *J. Appl. Probab.* 39:434–440.
- Ebrahimi, N. (1996). How to measure uncertainty in the residual lifetime distribution. *Sankhya A* 58:48–56.
- Ebrahimi, N., Pellerey, F. (1995). New partial ordering of survival functions based on notion of uncertainty. *J. Appl. Probab.* 32:202–211.
- Gupta, R. C., Kirmani, S. N. U. A. (2000). Residual coefficient of variation and some characterization results. *J. Statist. Plann. Infer.* 91:23–31.
- Kotlarski, I. I. (1972). On a characterization of some probability distributions by conditional expectations. *Sankhya A* 34:461–466.
- Kotz, S., Shanbhag, D. N. (1980). Some new approaches to probability distributions. *Adv. Appl. Probab.* 12:903–921.
- Nair, N. U., Sankaran, P. G. (1991). Characterization of the Pearson system of distributions. *IEEE Trans. Reliab.* 40:75–77.
- Nair, K. R. M., Rajesh, G. (1998). Characterization of probability distributions using the residual entropy function. *J. Ind. Statist. Assoc.* 36:157–166.
- Nair, K. R. M., Rajesh, G. (2000). Geometric vitality function and its applications to reliability. *IAPQR Trans.* 25(1):1–8.
- Navarro, J., Ruiz, J. M. (1996). Failure rate functions for doubly truncated random variables. *IEEE Trans. Reliab.* 45:685–690.
- Navarro, J., Ruiz, J. M. (2004). Characterization from relationships between failure rate functions and conditional moments. *Commun. Statist. Theor. Meth.* 33(12):3159–3171.
- Navarro, J., Franco, M., Ruiz, J. M. (1998a). Characterization through moments of the residual life and conditional spacing. *Sankhya A* 60:36–48.
- Navarro, J., Ruiz, J. M., Zoroa, P. (1998b). A unified approach to characterization problems using conditional expectations. *J. Statist. Plann. Infer.* 69:193–207.
- Navarro, J., del Aguila, Y., Ruiz, J. M. (2001). Characterizations through reliability measures from weighted distributions. *Statist. Pap.* 42:395–402.
- Rajesh, G., Nair, K. R. M. (2000). Residual entropy of conditional distributions. *Statist. Meth.* 2(1):72–80.

- Rao, C. R. (1965). On discrete distributions arising out of methods of ascertainment. In: Patil, G. P., ed. *Classical and Contagious Discrete Distributions*. Calcutta: Pergunon Press and Statistical Publishing Society, pp. 320–332. Also reprinted in *Sankhya A* 27:311–324.
- Ruiz, J. M., Navarro, J. (1996). Characterizations based on conditional expectations of the doubled truncated distribution. *Ann. Instit. Statist. Math.* 48(3):563–572.
- Sankaran, P. G., Gupta, R. P. (1999). Characterization of lifetime distributions using measure of uncertainty. *Calcutta Statist. Assoc. Bull.* 49:195–196.
- Sankaran, P. G., Sunoj, S. M. (2004). Identification of models using failure rate and mean residual life of doubly truncated random variables. *Statist. Pap.* 45:97–109.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27:379–423.
- Sunoj, S. M. (2004). Characterizations of some continuous distributions using partial moments. *Metron* LXII(3):353–362.
- Sunoj, S. M., Maya, S. S. (2006). Some properties of weighted distributions in the context of repairable systems. *Commun. Statist. Theor. Meth.* 35(2):223–228.
- Zoroa, P., Ruiz, J. M., Marin, J. (1990). A characterization based on conditional expectations. *Commun. Statist. Theor. Meth.* 19(8):3127–3135.