



Statistical Theory and Related Fields

ISSN: 2475-4269 (Print) 2475-4277 (Online) Journal homepage: https://www.tandfonline.com/loi/tstf20

Some results on quantile-based Shannon doubly truncated entropy

Vikas Kumar, Gulshan Taneja & Samsher Chhoker

To cite this article: Vikas Kumar, Gulshan Taneja & Samsher Chhoker (2019) Some results on quantile-based Shannon doubly truncated entropy, Statistical Theory and Related Fields, 3:1, 59-70, DOI: 10.1080/24754269.2019.1586282

To link to this article: https://doi.org/10.1080/24754269.2019.1586282



Published online: 15 Mar 2019.



Submit your article to this journal 🕝

Article views: 128



View related articles



View Crossmark data 🗹

Citing articles: 1 View citing articles

Some results on quantile-based Shannon doubly truncated entropy

Vikas Kumar^a, Gulshan Taneja^b and Samsher Chhoker^b

^aDepartment of Applied Sciences, UIET, M. D. University, Rohtak, India; ^bDepartment of Mathematics, M. D. University, Rohtak, India

ABSTRACT

Sunoj et al. [(2009). Characterization of life distributions using conditional expectations of doubly (Intervel)truncated random variables. *Communications in Statistics – Theory and Methods, 38*(9), 1441–1452] introduced the concept of Shannon doubly truncated entropy in the literature. Quantile functions are equivalent alternatives to distribution functions in modelling and analysis of statistical data. In this paper, we introduce quantile version of Shannon interval entropy for doubly truncated random variable and investigate it for various types of univariate distribution functions. We have characterised certain specific lifetime distributions using the measure proposed. Also we discuss one fascinating practical example based on the quantile data analysis.

ARTICLE HISTORY

Received 11 April 2018 Revised 8 February 2019 Accepted 20 February 2019

KEYWORDS

Quantile function; interval entropy; geometric vitality function; characterisation

1. Introduction

Let *X* be a non-negative absolutely continuous random variable representing the lifetime of a component with cumulative distribution function (CDF) $F(t) = P(X \le t)$ and survival function $\overline{F}(t) = P(X > t) = 1 - F(t)$. In modelling and analysis of lifetime data, the average amount of uncertainty associated with a nonnegative continuous random variable *X* is given in the differential entropy function

$$H(X) = -\int_0^\infty f(x)\log f(x)\,\mathrm{d}x,$$

a continuous counterpart of the Shannon (1948) entropy in the discrete case and f(x) is the probability density function (pdf) of the random variable X. While the concept of entropy has found increased application, little attention has yet been given to the practical problems of estimating entropy. Gong, Yang, Gupta, and Nearing (2014) discussed a method for computing robust and accurate estimates of entropy that accounts for several important characteristics of hydrological data sets. Since this entropy is not applicable to a system which has survived for some unit of time or for used item. The residual lifetime of the system when it is still operating at time t is $X_t = (X - t | X > t)$; which has the probability density $f(x; t) = f(x)/\overline{F}(t), x \ge t > 0$. Ebrahimi (1996) proposed the entropy of the residual lifetime X_t as

$$H(X;t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)} \,\mathrm{d}x, \quad t > 0.$$
(1)

In some practical situations, uncertainty is related to past life time rather than future. In this situation,

CONTACT Vikas Kumar 🖾 vikas_iitr82@yahoo.co.in

the random variable $X_t^* = (t - X | X \le t)$, which is known as inactivity time is suitable to describe the time elapsed between the failure of a system and the time when it is found to be 'down'. Based on this idea, Di Crescenzo and Longobardi (2002, 2004) have considered the entropy of the inactivity time X_t^* given as

$$\bar{H}(X;t) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} \, \mathrm{d}x.$$
 (2)

In many situations, we only have information between two points, and in this case statistical measures are studied under the condition of doubly truncated random variables. The doubly truncated measures are applicable to engineering systems when the observations are measured after it starts operating and before it fails. If the random variable X denotes the lifetime of a unit, then the random variable $X_{t_1,t_2} = (X - t_1 \mid t_1 \le X \le t_2)$ is called the doubly truncated (interval) residual lifetime, which in special case $t_2 \rightarrow \infty$ tends to residual lifetime random variable X_t . Also, we can use the doubly truncated past lifetime random variable $X_{t_1, t_2}^* = (t_2 - t_2)$ $X | t_1 \le X \le t_2$), which in special case $t_1 = 0$, it tends to past lifetime random variable X_t^* . Another extension of Shannon entropy is based on a doubly truncated (interval) random variable, which is as follows,

$$H(X; t_1, t_2) = -\int_{t_1}^{t_2} \frac{f(x)}{F(t_2) - F(t_1)} \log \frac{f(x)}{F(t_2) - F(t_1)} \, \mathrm{d}x. \quad (3)$$

Given that a system has survived up to time t_1 and has been found to be down at time t_2 , then $H(X; t_1, t_2)$ measure the uncertainty about its lifetimes between t_1 and t_2 . Different aspects and properties of $H(X; t_1, t_2)$ have been studied by Sunoj, Sankaran, and Maya (2009) and Misagh and Yari (2011, 2012). For various results on doubly truncated random variable, we refer to Sankaran and Sunoj (2004)), Khorashadizadeh, Rezaei Roknabadi, and Mohtashami Borzadaran (2013), Kayal and Moharana (2016), and Kundu (2017).

All the theoretical investigations and applications using these information measures are based on the distribution function. A probability distribution can be specified either in terms of its distribution function or by the quantile function. Although both convey the same information about the distribution with different interpretations, the concepts and methodologies based on distribution functions are traditional. When traditional approach are either difficult or fails to obtain desired results then quantile-based study were carried out. However, as Gilchrist (2000) discussed, there are many distinct properties for quantile functions that are not shared by the distribution functions, which makes the former attractive in certain practical situations. For inference purposes, quantile-based statistics are often more robust than those based on moments in the distribution function approach. Furthermore, there exist many simple quantile functions, that serve very well in empirical model building, for which distribution function are not in tractable forms, refer to van Staden and Loots (2009), Hankin and Lee (2006) and Nair, Sankaran, and Vinesh Kumar (2012). In many cases, QF is more convenient as it is less influenced by extreme observations and thus provides a straightforward analysis with a limited amount of information. In such cases, conventional tools of analysis using distribution functions are difficult to apply. An alternative approach to the study is to use the quantile functions (QFs), defined by

$$Q(u) = F^{-1}(u) = \inf\{x \mid F(x) \ge u\}, \quad 0 \le u \le 1.$$
(4)

When *F* is continuous, we have from (4), FQ(u) = u, where FQ(u) represents the composite function F(Q(u)). Defining the density quantile function by fQ(u) = f(Q(u)) and quantile density function by q(u) = Q'(u), where prime denotes the differentiation, we have

$$q(u)fQ(u) = 1,$$

refer to Nair, Sankaran, and Balkrishanan (2013). Several researchers have studied information theoretic measures based on quantile function. Sunoj and Sankaran (2012) have considered the quantile version of Shannnon entropy and its residuals form, defined as

$$H = \int_0^1 \log q(p) \,\mathrm{d}p \tag{5}$$

and

$$H(u) = \log(1-u) + (1-u)^{-1} \int_{u}^{1} \log q(p) \, \mathrm{d}p \quad (6)$$

respectively. Sunoj, Sankaran, and Nanda (2013) have considered the quantile past entropy, which is defined as

$$\bar{\mathbf{H}}(u) = \log u + u^{-1} \int_0^u \log q(p) \, \mathrm{d}p.$$
 (7)

Readers can refer to Nanda, Sankaran, and Sunoj (2014), Baratpour and Khammar (2018), Sankaran and Sunoj (2017), Guoxin (2018), and Kumar (2018) for more works on this line.

Motivated with the usefulness of the quantile function and the interval entropy, in the present note, we introduced a quantile version of Shannon interval entropy and derived some new characterisations to certain probability distributions as well as study its important properties. The proposed measure has several advantage. The measure proposed for doubly truncated random variable appears in quasar survey, where an investigator assumes that the apparent magnitude is doubly truncated. Also, the times to progression for patients with certain disease who received chemotherapy, experienced tumour progression and subsequently died, are doubly truncated. Secondly, quantile functions (QFs) have several properties that are not shared by distribution functions. Application of these properties give some new results and better insight into the properties of the measure that are difficult to obtain in the conventional approach. However, the use of QFs in the place of F provides new models, alternative methodology, easier algebraic manipulations, and methods of analysis in certain cases and some new results that are difficult to derive by using distribution function.

The paper is organised as follows. In Section 2, we consider the quantile version of Shannon interval entropy. In Section 3, the quantile interval entropy has been derived in case of some specific distributions. In Section 4, we study characterisation results concerning *Quantile Interval Entropy (QIE)* and also characterise a few specific lifetime distributions. Finally, conclusions have been given along with comments.

2. Quantile interval entropy

Defining a doubly truncated random variable $(X | u_1 \le X \le u_2)$ which represents the lifetime of a unit between u_1 and u_2 , where $(u_1, u_2) \in D = \{(u_1, u_2); Q(u_1) < Q(u_2)\}$. Corresponding to (4), a measure of uncertainty for the doubly truncated random variable in term of quantile function (4) is defined as

$$\begin{aligned} H(u_1, u_2) &= -\int_{u_1}^{u_2} \frac{f(Q(p))}{(u_2 - u_1)} \log \frac{f(Q(p))}{u_2 - u_1} \, \mathrm{d}Q(p) \\ &= \log(u_2 - u_1) \\ &+ \frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \log(q(p)) \, \mathrm{d}p. \end{aligned} \tag{8}$$

The important quantile measures useful in reliability analysis are *hazard quantile function* and *reversed* hazard quantile, defined as $A(u) = ((1-u)q(u))^{-1}$ and $\overline{A}(u) = (uq(u))^{-1}$, respectively, corresponding to the hazared rate $a(x) = f(x)/\overline{F}(x)$ and reversed hazared rate $\overline{a}(x) = f(x)/F(x)$ of X. In doubly truncation, Ruiz and Navarro (1996) defined the generalised hazard function (GHF) given by $h_1(t_1, t_2) =$ $f(t_1)/(F(t_2) - F(t_1))$ and $h_2(t_1, t_2) = f(t_2)/(F(t_2) F(t_1))$, respectively. Thus, generalised quantile hazard functions (GQHF) are defined as

$$h_1(u_1, u_2) = \frac{1}{(u_2 - u_1)q(u_1)} \text{ and}$$

$$h_2(u_1, u_2) = \frac{1}{(u_2 - u_1)q(u_2)}, \tag{9}$$

respectively. Equation (8) can be rewritten as

$$\begin{split} \dot{H}(u_1, u_2) &= 1 - \frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \log A(p) \, dp \\ &+ \frac{1}{(u_2 - u_1)} [(1 - u_2) \log(1 - u_2) \\ &- (1 - u_1) \log(1 - u_1) + \log(u_2 - u_1)], \end{split}$$
(10)

where (10) and (11) are the expression of quantile entropy in terms of hazard quantile function and reversed hazard quantile function, respectively. Using (6), (7) and (8), the quantile entropy (5) can be decomposed as

$$\begin{split} \dot{\mathbf{H}} &= u_1 \bar{\mathbf{H}}(u_1) + (u_2 - u_1) \mathbf{H}(u_1, u_2) + (1 - u_2) \mathbf{H}(u_2) \\ &- [u_1 \log u_1 + (u_2 - u_1) \log(u_2 - u_1) \\ &+ (1 - u_2) \log(1 - u_2)]. \end{split}$$
(12)

The identity (12) can be interpreted by decomposing the uncertainty about the failure of item into in the following way. into four parts:

- (i) The uncertainty about the failure time in (0, *u*₁) given that the item has failed before *u*₁,
- (ii) The uncertainty about the failure time in the interval (u_1, u_2) given that the item has failed after u_1 but before u_2 ,
- (iii) The uncertainty about the failure time in (u_2, ∞) given that it has failed after u_2 ,
- (iv) The uncertainty of the item that has failed before u_1 or in between u_1 and u_2 or after u_2 ,Differentiating $H(u_1, u_2)$ with respect to u_1

and u_2 , we have

$$\frac{\partial \mathcal{H}(u_1, u_2)}{\partial u_1} = \dot{\mathfrak{h}}_1(u_1, u_2) \left(\log \dot{\mathfrak{h}}_1(u_1, u_2) + \mathcal{H}(u_1, u_2) - 1 \right)$$
(13)

and

$$\frac{\partial \dot{\mathcal{H}}(u_1, u_2)}{\partial u_2} = -\dot{\mathfrak{h}}_2(u_1, u_2) \left(\log \dot{\mathfrak{h}}_2(u_1, u_2) + \dot{\mathcal{H}}(u_1, u_2) - 1 \right).$$
(14)

When $H(u_1, u_2)$ is increasing in u_1 and u_2 , then, (13) and (14) together imply

$$(1 - \log \mathfrak{h}_1(u_1, u_2))$$

$$\leq \mathfrak{H}(u_1, u_2) \leq (1 - \log \mathfrak{h}_2(u_1, u_2)).$$

Nair and Rajesh (2000) gave some applications of geometric vitality function. Sunoj et al. (2009) discussed few properties of this measure and showed that it determines the distribution function uniquely. Next, we define the quantile-based geometric vitality function.

Definition 2.1: Let *X* be a non-negative random variable then geometric vitality quantile function (GVQF) for a doubly truncated random variable is defined by

$$G(u_1, u_2) = E(\log X | u_1 < X < u_2)$$

= $\frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} fQ(p)q(p) \log Q(p) dp$
= $\frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \log Q(p) dp.$ (15)

This gives the geometric mean life of a doubly truncated random variable between the points u_1 and u_2 . Relationships between *geometric vitality quantile function* (15) for doubly truncated random variables and *generalised hazard quantile function* (9) are given in Table 1.

where

$$R(u_1, u_2) = E\left(\frac{1}{X} \left| u_1 < X < u_2\right)\right.$$
$$= \frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \frac{1}{Q(p)} \, \mathrm{d}p.$$
(16)

2.1. Relationship between $H(u_1, u_2)$ and quantile condition measure of uncertainty

Based on residual life distribution, Sankaran and Gupta (1999) have introduced a new measure of uncertainty known as conditional measure of uncertainty, which is defined as $M(X;t) = -E(\log f(X) | X > t) = -(1/\overline{F}(t)) \int_t^{\infty} f(x) \log f(x) dx$. The doubly truncated situation was considered in Sunoj et al. (2009) given as follows $M(X;t_1,t_2) = -E(\log f(X) | t_1 < X < t_2) =$

Table 1. Relationship between GVQF and GHQF.

Distribution	Quantile Functions	G(u ₁ , u ₂)
Exponential	$-\lambda^{-1}\log(1-u)$	$\frac{1}{\lambda} [\underline{h}_1(u_1, u_2) \log Q(u_1) - \underline{h}_2(u_1, u_2) \log Q(u_2) + R(u_1, u_2)]$
Finite Range	$\frac{1}{a}(1-(1-u)^{1/b})$	$\frac{1}{(ab)}[(1 - aQ(u_1))\dot{\mathfrak{h}}_1(u_1, u_2)\log Q(u_1) - (1 - aQ(u_2))\dot{\mathfrak{h}}_2(u_1, u_2)\log Q(u_2) + R(u_1, u_2) - a]$
Pareto-II	$\frac{1}{p}((1-u)^{1/q}-1)$	$\frac{1}{(pq)}[(1+pQ(u_1))\dot{\mathfrak{h}}_1(u_1,u_2)\log Q(u_1) - (1+pQ(u_2))\dot{\mathfrak{h}}_2(u_1,u_2)\log Q(u_2) + R(u_1,u_2) + p]$
Power	au ^{1/b}	$\log a + \frac{1}{b} \left[Q(u_2) \mathfrak{h}_2(u_1, u_2) \log \left(\frac{Q(u_2)}{a} \right) - Q(u_1) \mathfrak{h}_1(u_1, u_2) \log \left(\frac{Q(u_1)}{a} \right) - 1 \right]$
Pareto-1	$a(1-u)^{-1/b}$	$\log a + \frac{1}{b} \left[Q(u_2) \dot{\mathfrak{h}}_2(u_1, u_2) \log \left(\frac{a}{Q(u_2)} \right) - Q(u_1) \dot{\mathfrak{h}}_1(u_1, u_2) \log \left(\frac{a}{Q(u_1)} \right) + 1 \right]$

Table 2. Relation between $M(u_1, u_2)$	and GHQF for various distributions.
--	-------------------------------------

Distribution	Quantile Functions	Ӎ(<i>u</i> ₁ , <i>u</i> ₂)
Exponential	$-\lambda^{-1}\log(1-u)$	$\lambda m(u_1, u_2) - \log \lambda$
Finite range	$\frac{1}{a}(1-(1-u)^{1/b})$	$-\log ab + (-b+1)E[\log(1-aX) u_1 < X < u_2]$
Pareto -II	$\frac{1}{p}((1-u)^{1/q}-1)$	$(q+1)E[\log(1+pQ(u)) u_1 < X < u_2] - \log pq$
Power	$au^{1/b}$	$\left(\frac{b-1}{b}\right)\left[1+Q(u_1)\mathfrak{h}_1(u_1,u_2)\log\left(\frac{Q(u_1)}{a}\right)-Q(u_2)\mathfrak{h}_2(u_1,u_2)\log\left(\frac{Q(u_2)}{a}\right)\right]+\log\left(\frac{a}{b}\right)$
Pareto -l	$a(1-u)^{-1/b}$	$\left(\frac{b+1}{b}\right)\left[1-Q(u_1)\dot{\mathfrak{h}}_1(u_1,u_2)\log\left(\frac{a}{Q(u_1)}\right)+Q(u_2)\dot{\mathfrak{h}}_2(u_1,u_2)\log\left(\frac{a}{Q(u_2)}\right)\right]+\log\left(\frac{a}{b}\right)$
Weibull	$(-\log(1-u))^{1/p}$	$-\log p - (p-1)G(u_1, u_2) + E[(X)^p u_1 < X < u_2]$
Rayleigh	$(-\log(1-u))^{1/2}$	$-\log 2 - \log G(u_1, u_2) + E[(X)^2 u_1 < X < u_2]$

 $((-1)/(F(t_2) - F(t_1))) \int_{t_1}^{t_2} f(x) \log f(x) dx$. Using (4), the quantile-based condition measure of uncertainty for the doubly truncated random variable defined as

$$\begin{aligned}
& M(u_1, u_2) \\
&= -\frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} f(Q(u)) \log f(Q(u)) \, \mathrm{d}Q(u) \\
&= \frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \log q(u) \, \mathrm{d}u.
\end{aligned}$$
(17)

Using (17) in (8), we obtain

$$\mathcal{M}(u_1, u_2) = \mathcal{H}(u_2 - u_1) - \log(u_2 - u_1).$$
(18)

Differentiation of (18) with respect to u_1 and u_2 , respectively, provides the relationships with GHQF, which is given as $\partial M(u_1, u_2)/\partial u_1 = \partial H(u_1, u_2)/\partial u_1 +$ $h_1(u_1, u_2)$, and $\partial M(u_1, u_2)/\partial u_2 = \partial H(u_1, u_2)/\partial u_2$ $h_2(u_1, u_2)$. The various relationships between the quantile condition measure of uncertainty (17) and *GHQF* (9) for some commonly used probability models are given in Table 2.

3. Quantile interval entropy for various univariate distributions

In reliability theory, while studying the lifetime of a component or a system, a flexible model which is used in the literature is that of a generalised Pareto distribution (GPD) with survival function

$$\bar{F}(x) = \left(\frac{b}{ax+b}\right)^{1/a+1}, \quad x > 0 \ b > 0, \ a > -1.$$

It plays an important role in extreme value theory and other branches of statistics. The GPD, as a family of distributions, includes the exponential distribution when $a \rightarrow 0$, the Pareto type-II distribution or Lomax distribution for a > 0, which is used in the investigation of city population, occurrence of natural resources, insurance risk, size of human settlements, reliability modelling and business failure. It has been an important model in many socio-economic studies. The GPD becomes power distribution for -1 < a < 0. Next, let us discuss some examples on expression for quantile interval entropy function for some commonly used univariate distribution.

Example 3.1: If *X* is a random variable following the GPD with quantile function and quantile density function are given, respectively, by $Q(p) = (b/a)((1-p)^{-a/(a+1)}-1)$ and $q(p) = (b/(a+1))(1-p)^{-((2a+1)/(a+1))}$. Hence quantile interval entropy (8) for GPD is given by

 $\begin{aligned} H(u_1, u_2) \\ &= \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \log q(p) \, \mathrm{d}p \\ &= \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log \left[\left(\frac{b}{a+1} \right) (1-p)^{-} \left(\frac{2a+1}{a+1} \right) \right] \mathrm{d}p, \\ &= \log(u_2 - u_1) + \frac{\log \left(\frac{b}{a+1} \right)}{(u_2 - u_1)} \int_{u_1}^{u_2} \mathrm{d}p \\ &- \left(\frac{2a+1}{(a+1)(u_2 - u_1)} \right) \int_{u_1}^{u_2} \log(1-p) \, \mathrm{d}p, \end{aligned}$

which gives,

$$\begin{aligned}
\dot{H}(u_1, u_2) &= \log(u_2 - u_1) + \log\left(\frac{b}{a+1}\right) \\
&- \left(\frac{2a+1}{a+1}\right) \frac{1}{(u_2 - u_1)} \left[(1 - u_2)\right] \\
&\left(1 - \log(1 - u_2)\right) - (1 - u_1) \\
&\left(1 - \log(1 - u_1)\right) \right].
\end{aligned}$$
(19)

When $u_1 = u$ and $u_2 = 1$, then (19) reduces to $\dot{H}(u) = \log(b/(a+1)) + ((2a+1)/(a+1)) - (a/(a+1)) \log(1-u)$, the quantile residual entropy (6) for GPD. Also when $u_1 = 0$ and $u_2 = u$, then (19) reduces to $\bar{H}(u) = \log(b/(a+1)) + ((2a+1)/(a+1)) + \log u + ((2a+1)/(a+1))[((1-u)/u)\log(1-u)+1]$, the quantile past entropy (7) for GPD.

Example 3.2: If random variable *X* having the Pareto-II distribution with quantile function and quantile density function are given, respectively, by $Q(p) = a((1-p)^{-1/b} - 1)$ and $q(p) = (a/b)(1-p)^{-(1+b)/b}$. Then quantile interval entropy (8) becomes

$$\begin{split} \dot{H}(u_1, u_2) &= \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \\ \int_{u_1}^{u_2} \log\left(\frac{a}{b}(1 - p)^{(-(1+b))/b}\right) dp \\ &= \log(u_2 - u_1) + \frac{\log\left(\frac{a}{b}\right)}{(u_2 - u_1)} \int_{u_1}^{u_2} dp \\ &- \frac{(1+b)}{b(u_2 - u_1)} \int_{u_1}^{u_2} \log(1 - p) dp \\ &= \log(u_2 - u_1) + \log\left(\frac{a}{b}\right) \\ &- \frac{(1+b)}{b(u_2 - u_1)} [(1 - u_2)(1 - \log(1 - u_2)) \\ &- (1 - u_1)(1 - \log(1 - u_1))]. \end{split}$$

Example 3.3: If a random variable *X* follows the rescaled beta distribution distribution with quantile and quantile density functions are given, respectively, by $Q(p) = R(1 - (1 - p)^{1/c})$ and $q(p) = (R/c)(1 - p)^{(1-c)/c}$. Then quantile interval entropy for the rescaled beta distribution is given as

$$\begin{aligned} H(u_1, u_2) &= \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log\left(\frac{R}{c}(1 - p)^{(1 - c)/c}\right) dp, \\ &= \log(u_2 - u_1) + \frac{\log\left(\frac{R}{c}\right)}{(u_2 - u_1)} \int_{u_1}^{u_2} dp \\ &+ \frac{(1 - c)}{c(u_2 - u_1)} \int_{u_1}^{u_2} \log(1 - p) dp \end{aligned}$$

which gives,

$$\dot{H}(u_1, u_2) = \log(u_2 - u_1) + \log\left(\frac{R}{c}\right) \\ + \left(\frac{1 - c}{c}\right) \left(\frac{1}{u_2 - u_1}\right) \\ [(1 - u_2)(1 - \log(1 - u_2)) \\ - (1 - u_1)(1 - \log(1 - u_1))]$$

Example 3.4: If a random variable *X* having half logistic distribution with quantile and quantile density functions as $Q(p) = \sigma \log((1 + p)/(1 - p))$ and $q(p) = 2\sigma/(1 - p)(1 + p)$, respectively. Then quantile interval entropy (8) for half logistic distribution is given as

$$\begin{aligned} H(u_1, u_2) &= \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log\left(\frac{2\sigma}{(1 - p)(1 + p)}\right) \, \mathrm{d}p, \\ &= \log(u_2 - u_1) + \log 2\sigma \\ &- \frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \log(1 - p)(1 + p) \, \mathrm{d}p. \end{aligned}$$

after some algebraic simplifications, we obtain

$$\begin{split} H(u_1, u_2) &= \log(u_2 - u_1) + \log(2\sigma) \\ &- \frac{1}{(u_2 - u_1)} (1 - u_2)(1 - \log(1 - u_2)) \\ &+ \frac{1}{(u_2 - u_1)} (1 - u_1)(1 - \log(1 - u_1)) \\ &- \frac{1}{(u_2 - u_1)} (1 + u_2)(\log(1 + u_2) - 1) \\ &+ \frac{1}{(u_2 - u_1)} (1 + u_1)(\log(1 + u_1) - 1). \end{split}$$

$$(20)$$

Put $u_1 = u$ and $u_2 = 1$, in (20), we have the quantile residual entropy for the half logistic distribution which is given by

$$H(u) = 3 + \log 2\sigma - \frac{2}{(1-u)}\log 2 + \log(1-u) + \frac{2}{(1-u)}\log(1+u)$$

and for $u_1 = 0$ and $u_2 = u$ then we get quantile past entropy for the half logistic distribution as

$$\bar{\mathcal{H}}(u) = 2 + \log u + \log 2\sigma - \left(\frac{1+u}{u}\right)\log(1+u) + \left(\frac{1-u}{u}\right)\log(1-u).$$

Example 3.5: If a random variable *X* having log logistic distribution with quantile function and quantile density function are respectively given as $Q(p) = (1/a)((p/(1-p))^{1/b})$ and $q(p) = (1/ab)(p/(1-p))^{1/b}((p(1-p))^{-1})$. Then quantile interval entropy for log logistic distribution is given as

$$\begin{split} \dot{H}(u_1, u_2) \\ &= \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log\left(\frac{1}{ab} \left(\frac{p}{1 - p}\right)^{1/b} \left(p(1 - p)\right)^{-1}\right) dp, \\ &= \log(u_2 - u_1) + \log\left(\frac{1}{ab}\right) + \frac{1}{(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log\left(\left(\frac{p}{1 - p}\right)^{1/b} \left(p(1 - p)\right)^{-1}\right) dp, \\ &= \log(u_2 - u_1) + \log\left(\frac{1}{ab}\right) + \frac{(1 - b)}{b(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log p \, dp - \frac{(1 + b)}{b(u_2 - u_1)} \int_{u_1}^{u_2} \log(1 - p) \, dp \end{split}$$

which gives,

$$\begin{split} \dot{H}(u_1, u_2) \\ &= \log(u_2 - u_1) + \log\left(\frac{1}{ab}\right) \\ &+ \left(\frac{1-b}{b}\right) \frac{1}{(u_2 - u_1)} \\ &[u_2\left(\log u_2 - 1\right) - u_1(\log u_1 - 1)] \\ &- \left(\frac{1+b}{b}\right) \frac{1}{(u_2 - u_1)} [(1 - u_2)\left(1 - \log(1 - u_2)\right) \\ &- (1 - u_1)\left(1 - \log(1 - u_1)\right)]. \end{split}$$
(21)

Substituting $u_1 = u$ and $u_2 = 1$ in (21), then we get quantile residual entropy for log logistic distribution given as

$$\begin{split} \dot{H}(u) &= 2 + \log\left(\frac{1}{ab}\right) + \left(\frac{b-1}{b}\right) \left(\frac{u}{1-u}\right) \log u \\ &- \left(\frac{1}{b}\right) \log(1-u), \end{split}$$

whereas the quantile past entropy for log logistic distribution is obtain, when we take $u_1 = 0$ and $u_2 = u$ is given as

$$\bar{\mathcal{H}}(u) = 2 - \log\left(\frac{1}{ab}\right) + \frac{1}{b}\log u + \left(\frac{1+b}{b}\right)\left(\frac{1-u}{u}\right)\log(1-u).$$

Example 3.6: Let *X* be a random variable having the exponential geometric distribution with quantile and quantile density functions are given, respectively, by $Q(p) = (1/\lambda)(\log((1-qp)/(1-p)))$ and $q(p) = (1/\lambda)((1-q)/(1-qp))(1-p)^{-1}$. Then quantile interval entropy for the exponential geometric distribution is given as

$$\begin{split} H(u_1, u_2) &= \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log\left[\frac{1}{\lambda} \left(\frac{1 - q}{1 - qp}\right) (1 - p)^{-1}\right] dp, \\ &= \log(u_2 - u_1) + \log\left(\frac{1 - q}{\lambda}\right) \\ &- \frac{1}{(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log\left[\left(1 - qp\right) (1 - p)\right] dp, \end{split}$$

which gives

$$\begin{aligned} H(u_1, u_2) \\ &= \log(u_2 - u_1) + \log\left(\frac{1 - q}{\lambda}\right) \\ &- \frac{1}{q(u_2 - u_1)} [(1 - qu_2)(1 - \log(1 - u_2))) \\ &- (1 - qu_1)(1 - \log(1 - u_1))] \\ &- \frac{1}{(u_2 - u_1)} [(1 - u_2)\left(1 - \log(1 - u_2)\right) \\ &- (1 - u_1)(1 - \log(1 - u_1))]. \end{aligned}$$
(22)

Particularly, if we put $u_1 = u$ and $u_2 = 1$ in (22), then we get quantile residual entropy for the exponential geometric distribution as

$$\begin{split} \dot{H}(u) &= \log 2 + \log(1-u) + \log\left(\frac{1-q}{\lambda}\right) \\ &+ \frac{1}{q} \left[(1-q) \log(1-q) \\ &- (1-qu) \log(1-qu) \right], \end{split}$$

and if we take $u_1 = 0$ and $u_2 = u$, we get quantile past entropy for the exponential geometric distribution as

$$\bar{\mathcal{H}}(u) = \log u + \log\left(\frac{1-q}{\lambda}\right) + \left(\frac{1+q}{q}\right)\left(\frac{1}{u}\right) + \log\left(\frac{1-u}{1-qu}\right).$$

Example 3.7: If *X* is a random variable following the quantile function and quantile density functions of linear hazard rate distribution are given, respectively, by $Q(p) = (1/(a+b))\log((a+bp)/a(1+p))$

and q(p) = ((a+b)/(b-a))(1/(a+bp)(1+p)).Hence quantile interval entropy $H(u_1, u_2)$ is given as

$$\begin{aligned} H(u_1, u_2) &= \log(u_2 - u_1) \\ &+ \frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \log q(p) \, \mathrm{d}p, \\ &= \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log\left(\left(\frac{a+b}{b-a}\right) \frac{1}{(a+bp)(1+p)}\right) \, \mathrm{d}p, \\ &= \log(u_2 - u_1) + \log\left(\frac{b+a}{b-a}\right) \\ &- \frac{1}{(u_2 - u_1)} \left[\int_{u_1}^{u_2} \log(1+p) \, \mathrm{d}p \\ &+ \int_{u_1}^{u_2} \log(a+bp) \, \mathrm{d}p \right]. \end{aligned}$$

After some algebraic simplifications, we have

$$\begin{split} \dot{H}(u_1, u_2) &= 2 + \log(u_2 - u_1) + \log\left(\frac{b+a}{b-a}\right) \\ &+ \left(\frac{1+u_1}{u_2 - u_1}\right) \log(1+u_1) \\ &- \left(\frac{1+u_2}{u_2 - u_1}\right) \log(1+u_2) \\ &+ \frac{1}{b(u_2 - u_1)} [(a+bu_1)\log(1+bu_1)] \\ &- (1+bu_2)\log(1+bu_2)]. \end{split}$$

Example 3.8: If *X* be a random variable following the Davies distribution (2006), that do not have any closed form expressions for distribution and density function, then QFs and quantile density functions are given, respectively, by $Q(p) = cp^{\lambda_1}(1-p)^{-\lambda_2}$ and $q(p) = cp^{(\lambda_1-1)}(1-p)^{(-\lambda_2-1)}(\lambda_2p + \lambda_1(1-p))$. Hence quantile interval entropy (8) for Davies distribution is given by

$$\begin{aligned} H(u_1, u_2) &= \log(u_2 - u_1) \\ &+ \frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \log q(p) \, \mathrm{d}p, \\ &= \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \\ &\int_{u_1}^{u_2} \log[cp^{(\lambda_1 - 1)}(1 - p)^{(-\lambda_2 - 1)} \\ &(\lambda_2 p + \lambda_1 (1 - p))] \mathrm{d}p, \end{aligned}$$
$$&= \log(u_2 - u_1) + \log c + \frac{1}{(u_2 - u_1)} \\ &\left[(\lambda_1 - 1) \int_{u_1}^{u_2} \log p \, \mathrm{d}p - (\lambda_2 + 1) \\ &\int_{u_1}^{u_2} \log(1 - p) \, \mathrm{d}p \right] + \frac{1}{(u_2 - u_1)} \end{aligned}$$

$$\int_{u_1}^{u_2} \log \left(\lambda_2 p + \lambda_1 (1-p) \right) \mathrm{d}p,$$

We get, after some algebraic calculations,

$$\begin{split} \dot{H}(u_1, u_2) &= (\lambda_2 - \lambda_1 + 1) + \log c + \log(u_2 - u_1) \\ &+ \frac{(\lambda_1 - 1)u_2}{(u_2 - u_1)} \log u_2 \\ &- \frac{u_1(\lambda_1 - 1)}{(u_2 - u_1)} \log u_1 + \frac{(\lambda_2 + 1)(1 - u_2)}{(u_2 - u_1)} \\ &\log(1 - u_2) + (\lambda_2 + 1) \\ &\frac{(1 - u_1)}{(u_2 - u_1)} \log(1 - u_1) \\ &+ \frac{(\lambda_2 u_2 + (1 - u_2)\lambda_1)}{(\lambda_2 - \lambda_1)(u_2 - u_1)} \log(\lambda_2 u_2 \\ &+ (1 - u_2)\lambda_1) - \frac{(\lambda_2 u_1 + (1 - u_1)\lambda_1)}{(\lambda_2 - \lambda_1)(u_2 - u_1)} \\ &\log(\lambda_2 u_1 + (1 - u_1)\lambda_1). \end{split}$$

If we substitute $u_1 = u$ and $u_2 = 1$, then (23) reduces to

$$\begin{split} H(u) &= (\lambda_2 - \lambda_1 + 1) + \log c \\ &- \frac{1}{(1 - u)(\lambda_2 - \lambda_1)\lambda_2 \log \lambda_2} - \lambda_2 \log(1 - u) \\ &- \frac{1}{(1 - u)(\lambda_1 - 1)u} \log u \\ &- \frac{(\lambda_1(1 - u) + \lambda_2 u)}{(1 - u)(\lambda_2 - \lambda_1)} \log(\lambda_1(1 - u) + \lambda_2 u), \end{split}$$

the quantile residual entropy (6) for Davies distribution. When we put $u_1 = 0$ and $u_2 = u$, then (23) reduces to

$$\begin{split} \bar{\mathcal{H}}(u) &= (\lambda_2 - \lambda_1 + 1) + \log c + \lambda_1 \log u \\ &+ \frac{(\lambda_2 + 1)(1 - u)}{u} \log(1 - u) \\ &+ \frac{(\lambda_2 u + (1 - u)\lambda_1)}{(\lambda_2 - \lambda_1)u} \log(\lambda_2 u + (1 - u)\lambda_1) \\ &- \frac{\lambda_1}{(\lambda_2 - \lambda_1)u} \log \lambda_1, \end{split}$$

the quantile past entropy (7) for Davies Distribution.

Example 3.9: If *X* be a random variable following the Govindarajulu's distribution that do not have any closed form expressions for distribution and density function, then QFs and quantile density functions are given, respectively, by

$$Q(u) = a\{(b+1)u^b - bu^{b+1}\} \text{ and}$$
$$q(u) = ab(b+1)(1-u)u^{b-1}; \quad 0 \le u \le 1; a, b > 0.$$

Thus quantile-based interval entropy (8) for Govindarajulu's distribution is given as

$$\mathcal{H}(u_1, u_2) = \log(u_2 - u_1) + \frac{1}{(u_2 - u_1)} \\
 \int_{u_1}^{u_2} \log ab(b+1)(1-p)p^{b-1} dp, \\
 = \log(u_2 - u_1) + \log ab(b+1) \\
 + \frac{1}{(u_2 - u_1)} \int_{u_1}^{u_2} \log(1-p) dp \\
 + \frac{(b-1)}{(u_2 - u_1)} \int_{u_1}^{u_2} \log p dp$$
(24)

which gives

$$\begin{split} \dot{H}(u_1, u_2) &= \log(u_2 - u_1) + \log ab(b+1) \\ &- \left(\frac{1 - u_2}{u_2 - u_1}\right) \log(1 - u_2) \\ &+ \left(\frac{1 - u_1}{u_2 - u_1}\right) \log(1 - u_1) \\ &+ \frac{(b-1)}{u_2 - u_1} \left(u_2 \log u_2 - u_1 \log u_1\right). \end{split}$$

Table 3 provides the relationships between the quantile interval entropy $H(u_1, u_2)$, quantile conditional expectation

$$m(u_2, u_1) = E(X \mid u_1 < X < u_2)$$

= $\frac{1}{(u_1 - u_2)} \int_{u_1}^{u_2} Q(u) \, du,$ (25)

and generalised hazard quantile function $h_i(u_1, u_2)$; i = 1, 2 for some commonly used distribution.

4. Characterisation results

In the literature, the problem of characterising probability distributions has been investigated by many researchers. The standard practice in modelling 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 the situations, it would be of more use if we find characterisation theorems that explain the distribution using important measures of indices as. In this section, we discussed some characterisation theorems for lifetime distribution taking some important concepts like *GHQF*, *GVQF* and quantile-based *condition Shannon's measures* of uncertainty. **Theorem 4.1:** Let X be a random variable defined on $(0, \infty)$ having the quantile function Q(u), then the relationship

$$G(u_1, u_2) = \frac{1}{k} [(1 + CQ(u_1))h_1(u_1, u_2) \log Q(u_1) - (1 + CQ(u_2))h_2(u_1, u_2) \log Q(u_2) + R(u_1, u_2) + C]$$
(26)

where k, C are constants holds for all $(u_1, u_2) \in D$. If and only if for

- (i) C = 0, X has exponential distribution with quantile function $Q(u) = -(1/\lambda) \log(1-u)$,
- (ii) C > 0, X has Pareto distribution with quantile function $Q(u) = (1/a)(1-u)^{1/b}$, and
- (iii) C < 0, X has finite range distribution with quantile function $Q(u) = a(1 (1 u)^{1/b})$.

Proof: The if part is straightforward from the Table 1. To prove the converse, let us assume that (26) holds. Using (15), (9) and (16) in (26), we obtain

$$\int_{u_1}^{u_2} fQ(p)q(p)\log Q(p) dp$$

$$= \frac{1}{k} \left\{ \log Q(u_1) \left(1 + CQ(u_1) \right) fQ(u_1) - \left(1 + CQ(u_2) \right) fQ(u_2) \log Q(u_2) \right\} + \int_{u_1}^{u_2} \frac{1}{Q(p)} dp + C(u_2 - u_1).$$
(27)

Differentiating (27) with respect to u_i , i = 1,2 we get, after some algebraic calculations,

$$\frac{f'(Q(u_i))}{f(Q(u_i))} = \frac{(k+C)}{(1+CQ(u_i))}, \quad i = 1, 2$$

or f'(Q(u))/f(Q(u)) = (k + C)/(1 + CQ(u)), which gives the required result.

Theorem 4.2: For a non-negative random variable *X*, the relation

$$M(u_1, u_2) - \lambda m(u_1, u_2) = k, \quad (k > 0, a \text{ constant})$$
(28)

holds for all $(u_1, u_2) \in D$ if and only if X follows exponential distribution with quantile function $Q(u) = -(1/\lambda) \log(1-u)$.

Proof: The if part is straightforward from the Table 2. To prove the converse, let us assume that (28) holds.

Distribution	Quantile Functions	𝔅(u₁, u₂)
Exponential	$-\lambda^{-1}\log(1-u)$	$\lambda m(u_1, u_2) - \lambda Q(u_2)) - \log \mathfrak{h}_2(u_1, u_2)$
Finite Range	$\frac{1}{a}(1-(1-u)^{1/b})$	$(-b+1)E[\log(1-aQ(u) \mid u_1 < X < u_1] - \log \frac{ab}{((1-aQ(u_1)^b - ((1-aQ(u_2)^b - U_1)^b - (U_1)^b $
Pareto-II	$\frac{1}{p}((1-u)^{1/q}-1)$	$(q+1)E[\log(1+pQ(u)) \mid u_1 < X < u_1] - \log \frac{pq}{((1+pQ(u_1)^{-q} - ((1+pQ(u_2)^{-q} - (1+pQ(u_2)^{-q} -$
Power	au ^{1/b}	$1 + G(u_1, u_2) + Q(u_1)\mathfrak{h}_1(u_1, u_2)\log\left(\frac{Q(u_1)}{a}\right) - Q(u_2)\mathfrak{h}_2(u_1, u_2)\log\left(\frac{Q(u_2)}{a}\right) - \log\left[\frac{b}{\left(\frac{Q(u_1)}{a}\right)^b - \left(\frac{Q(u_1)}{a}\right)^b}\right]$
Pareto-1	$a(1-u)^{-1/b}$	$1 + G(u_1, u_2) - Q(u_1) \dot{h}_1(u_1, u_2) \log\left(\frac{a}{Q(u_1)}\right) + Q(u_2) \dot{h}_2(u_1, u_2) \log\left(\frac{a}{Q(u_2)}\right) - \log\left[\frac{b}{\left(\frac{a}{Q(u_1)}\right)^b - \left(\frac{a}{Q(u_1)}\right)^b}\right]$

Table 3. Relationships between $\oint_i (u_1, u_2)$; i = 1, 2 and $\oint_i (u_1, u_2)$.

Then using (16) and (25), (28) becomes

$$\int_{u_1}^{u_2} \log q(p) \, \mathrm{d}p - \lambda \int_{u_1}^{u_2} Q(p) \, \mathrm{d}p = k(u_2 - u_1).$$
(29)

Differentiating (29) with respect to u_i , i = 1, 2, we get, after some algebraic calculations,

 $f(Q(u_i)) = Ke^{-\lambda Q(u_i)}, \quad i = 1, 2 \text{ and } K > 0 \text{ (constant)}$

or $f(Q(u)) = Ke^{-\lambda Q(u)}$, which characterise the exponential distribution.

Theorem 4.3: If X be a non-negative random variable with quantile function Q(u), and constants k > 0 and c > 0. A relationship of the form

$$M(u_1, u_2) - (c+1)G(u_1, u_2) = k,$$
 (30)

holds for $a < u_1 < u_2$ with $Q(u_2) < Q(u_2)$ if and only if X follows Pareto-1 distribution with quantile function $Q(u) = (a/(1-u)^{1/c}); a > 0.$

Proof: The if part is straightforward. To prove the converse, let us assume that (30) holds. Then using (15) and (16), we have

$$\int_{u_1}^{u_2} \log q(p) \, \mathrm{d}p - (c+1)$$
$$\int_{u_1}^{u_2} \log Q(p) \, \mathrm{d}p = k(u_2 - u_1). \tag{31}$$

Differentiating (31) with respect to u_i , i = 1, 2, we get, after some algebraic calculations

$$f(Q(u_i)) = k_1(Q(u_i))^{-(c+1)},$$

$$i = 1, 2 \text{ and } k_1 \text{ (constant)}$$

or $f(Q(u)) = k_1 (Q(u))^{-(c+1)}$, which gives the required result.

Next, we state the characterisation of power distribution. The proof is similar to that of Theorem 4.3 and hence omitted.

Theorem 4.4: If X is a non-negative random variable with quantile function Q(u), and K > 0, C > 1 be constants. A relationship of the form

$$M(u_1, u_2) + (c - 1)G(u_1, u_2) = k$$

is holds for $0 < u_1 < u_2 < b$ with $Q(u_1) < Q(u_1)$ if and only if X follows power distribution with quantile function $Q(u) = au^{1/b}$.

Theorem 4.5: Let X be a random variable defined on $(0, \infty)$ with quantile function Q(u). Then X follows oneparameter log exponential distribution if and only if

$$M(u_1, u_2) = \log A(\theta) - \theta G(u_1, u_2) - m_c(u_1, u_2),$$
(32)
where $m_c(u_1, u_2) = E[\log C(X) | u_1 < X < u_2]$ for all
 $(u_1, u_1) \in D.$

Proof: The if part is straightforward from the Table 1. To prove the converse, let us assume that (32) holds. Using (15) and (16) in (32), we have

$$\int_{u_1}^{u_2} \log f(Q(p)) \, \mathrm{d}p = (u_2 - u_1) \log A(\theta)$$
$$-\theta \int_{u_1}^{u_2} \log Q(p) \, \mathrm{d}p - \int_{u_1}^{u_2} \log C(Q(p)) \, \mathrm{d}p. \tag{33}$$

Differentiating (33) with respect to u_i , i = 1, 2 we get, after some algebraic calculations,

$$f(Q(u_i)) = \frac{(Q(u_i))^{\theta} C(Q(u_i))}{A(\theta)}, \quad i = 1, 2$$

or $f(Q(u)) = (Q(u))^{\theta} C(Q(u)) / A(\theta)$, which gives the required result.

We conclude this section by characterising exponential distribution. The proof is similar to that of Theorem 4.5 and hence omitted.

Theorem 4.6: Let $X \in (0, \infty)$ be a random variable having absolutely continuous quantile function Q(u). Then the relationship of the form

$$M(u_1, u_2) = \log b(\theta) - m(u_1, u_2) \log \theta - m_a(u_1, u_2),$$

where $m_a(u_1, u_2) = E[\log a(X) | u_1 < X < u_2]$ holds for all $(u_1, u_2) \in D$ if and only if X follows one-parameter exponential distribution.

4.1. Exploratory data analysis using Q-Q Plot

Quantile-quantile (Q-Q) plot is a diagnostic tool, which is widely used to assess the distributional similarities and differences between two independent univariate samples. It is also a popular device for checking the appropriateness of a specified probability distribution for a given univariate data. The advantages of the Q-Q plot are: The sample sizes do not need to be equal. Many distributional aspects can be simultaneously tested. For example, shifts in location, shifts in scale, changes in symmetry, and the presence of outliers can all be detected from this plot. The Q-Q plot is similar to a probability plot. For a probability plot, the quantiles for one of the data samples are replaced with the quantiles of a theoretical distribution.

Example 4.1: Consider the Rainfall data from seeded clouds and non-seeded clouds are given below: (numbers in paratheses indicate the number of repetitions of the values)

Rainfall from control clouds: 1, 4.9(2), 11.5, 17.3, 21.7, 24.4, 26.1, 26.3, 28.6, 29,36.6,41.1, 47.3, 68.5, 81.2, 87.0, 95, 147.8, 163, 243.3, 321.2, 345.5,372.4, 830.1, 1202.6.

Rainfall from seeded clouds: 4.1, 7.7, 17.5, 31.4, 32.7, 40.6, 92.4, 115.3, 118.3, 119, 129.6, 198.6, 200.7, 242.5, 255.0, 274.7(2),302.8,334.1, 430.0, 489.1, 703.4, 978, 1656, 1697.8, 2745.6. (Source: Simpson, Olsen, & Eden, 1975)

The conclusions of our data analysis are as follows: *Compare the location of the data sample (mean, median)*: Seeded rainfall has greater location parameter than non-seeded rainfall.

Compare scale: The interquartile range indicates the variability of seeded rainfall is greater than non-seeded rainfall.

 Table 4. Numerical summary of rainfall for seeded clouds and non- seeded clouds.

	Rainfall from non-seeded clouds	Rainfall from seeded clouds
Sample size	26	26
Number of different values	25	25
Sample mean	601.8	1374.85
Sample standard deviation	849.66	1938.533
S/\sqrt{n}	166.63	380.18
sample min	1	4.1
Sample max	1202.6	2745.6
Next to min	4.9	7.7
Next to max	830.1	1697.8
<i>Q</i> ₁	24.4	92.4
Q ₂	44.2	221.25
Q ₃	163	430
MQ	93.7	261.2
IQR2	277.2	675.2
$\frac{Q_2 - MQ}{IQR2}$	-0.18	-0.06
Upper fence	370.9	936.4
Upper outliers	830.1,1202.6	1656,1697.8,2745.6
MIN-MQ IQR2	-0.33	-0.38
Conclusion	Skew	symmetric

Table 5. Quantile	plot	table of	seeded	rainfall.
-------------------	------	----------	--------	-----------

Rank	$Q(u_i)$	Ui	Zi
1	4.1	.01923	-2.070
2	7.7	.076923	-1.425
3	17.5	.096154	-1.304
•			
24	1656	.90385	1.304
25	1697	.94231	1.425
26	2745.6	.98077	2.069



Side by side box plots: These plot reader can draw using Table 4; non-seeded rainfall has a skew distribution, whereas seeded rainfall is symmetric.

Identification of probability laws: Seeded rainfall and non-seeded data both indicates fit by exponential distribution.

The Q-Q plot is a graphical technique for determining if two data sets come from populations with a common distribution. A Q-Q plot is a plot of the quantiles of the first data set against the quantiles of the second data set. In general for computing a normal probability plot, the standard normal table could be used to approximate the normal quantile. We look up the *Z* value corresponding to the p_i for $i = 1, 2, 3 \cdots n$ and then plot the ordered data against the corresponding *Z* value. Table 5 displays the quantiles for rainfall

 Table 6. Quantile plot table of non-seeded rainfall.

Rank	$Q(u_i)$	ui	Zi
1	1.0	.01923	-2.070
2	4.9	.076923	-1.425
3	4.9	.096154	-1.304
4	11.5	.13462	-1.108
24	372.4	.90385	1.304
25	830.1	.94231	1.425
26	1202.6	.98077	2.070



Quantile Plot for Non-Seeded Rainfall

data from seeded clouds as well as the corresponding normal quantile approximated from a standard normal table (Table 6).

5. Conclusion

Recently, there has been a great interest in the study of information measures based on quantile functions, namely quantile entropy. When a system has lifetime between two time points (t_1, t_2) , the *interval entropy* plays an important role, in the field of reliability theory and survival analysis. The present work introduced an alternative approach to interval entropy measure using quantile functions. The proposed measures may help information theorists and reliability analysts to study the various characteristics of a system when it fails between two time instants. The results presented here generalise the related existing results in context with quantile entropy for residual and past lifetime random variables.

Acknowledgments

The authors would like to express their gratitude to the reviewers and the editor-in-chief for their valuable comments, which have considerably improved the earlier version of the article.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The first author wishes to acknowledge the Science and Engineering Research Board (SERB), Government of India, for the financial assistance (Ref. No. ECR/2017/001987) for carrying out this research work.

Notes on contributors

Vikas Kumar has obtained his M.Sc and M.Phil degree in Applied Mathematics from IIT Roorkee and ISM University, Dhanbad in 2005 and 2007 respectively. He received the Ph.D. degree in Mathematics from University of Delhi, Delhi. Currently, he is a Assistant Professor in Mathematics, UIET, M. D. University, Rohtak, India. His research interests are information theory and its applications and mathematical modeling. He has published research articles in reputed international journals of Mathematics and Statistical sciences.

Dr. Gulshan Taneja is working as Professor in Mathematics at M. D. University, Rohtak. He has vast experience of about 25 years of teaching in mathematics and statistics both at UG and PG levels. Dr. Taneja has published more than sixty research papers in the field of Information Theory & Reliability Theory in journals of international repute and is a member of various national and international societies.

Samsher Chokker has obtained his M.Sc and M.Phil degree in Mathematics from M. D. University, Rohtak in 2015 and 2017 respectively. He pursuing his Ph.D. degree in Mathematics from M. D. University, Rohtak . Currently, he is a Assistant Professor in Mathematics, Government PG Nehru College Jhajjar, India. His research interests are information theory and and mathematical modeling.

References

- Baratpour, S., & Khammar, A. H. (2018). A quantile-based generalized dynamic cumulative measure of entropy. *Communications in Statistics – Theory and Methods*, 47(13), 3104–3117.
- Di Crescenzo, A., & Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. *Journal of Applied Probability*, *39*, 434–440.
- Di Crescenzo, A., & Longobardi, M. (2004). A measure of discrimination between past lifetime distributions. *Statistics* & *Probability Letters*, 67, 173–182.
- Ebrahimi, N. (1996). How to measure uncertainty in the residual life distributions. *Sankhya Series A*, *58*, 48–57.
- Gilchrist, W. (2000). *Statistical modelling with quantile functions*. Boca Raton, FL: Chapman and Hall/CRC.
- Gong, W., Yang, D., Gupta, H. V., & Nearing, G. (2014). Estimating information entropy for hydrological data: One-dimensional case. *Water Resources Research*, 50, 5003–5018. doi:10.1002/2014WR015874
- Hankin, R. K. S., & Lee, A. (2006). A new family of nonnegative distributions. Australian and New Zealand Journal of Statistics, 48, 67–78.
- Kayal, S., & Moharana, R. (2016). Some Results on a doubly truncated generalized discrimination measure. *Applications of Mathematics*, 61, 585–605.
- Khorashadizadeh, M., Rezaei Roknabadi, A. H., & Mohtashami Borzadaran, G. R. (2013). Mohtashami Borzadaran Doubly truncated (interval) cumulative residual andpast entropy. *Statistics & Probability Letters*, 83, 1464–1471.
- Kumar, V. R. (2018). A quantile approach of Tsallis entropy for order statistics. *Physica A: Statistical Mechanics and its Applications*, 503, 916–928.
- Kundu, C. (2017). On weighted measure of inaccuracy for doubly truncated random variables. *Communications in Statistics – Theory and Methods*, 46, 3135–3147.
- Misagh, F., & Yari, G. (2011). On weighted interval entropy. *Statistics & Probability Letters*, 29, 167–176.
- Misagh, F., & Yari, G. H. (2012). Interval entropy and Informative Distance. *Entropy*, 14, 480–490.

- Nair, K. R. M., & Rajesh, G. (2000). Geometric vitality function and its applications to reliability. *IAPQR Transactions*, 25, 1–8.
- Nair, N. U., Sankaran, P. G., & Balkrishanan, N. (2013). Quantile based reliability analysis. Statistics for industry and technology. New York, NY: Springer Science+Business Media.
- Nair, N. U., Sankaran, P. G., & Vinesh Kumar, B. (2012). Modeling lifetimes by quantile functions using Parzen's score function. *Statistics-A Journal of Theoretical and Applied Statistics*, 46(6), 799–811.
- Nanda, A. K., Sankaran, P. G., & Sunoj, S. M. (2014). Renyi's residual entropy: A quantile approach. *Statistics & Probability Letters*, 85, 114–121.
- Qiu, G. (2018). Further results on the residual quantile entropy. *Communications in Statistics – Theory and Methods*, 47(13), 3092–3103.
- Ruiz, J. M., & Navarro, J. (1996). Characterizations based on conditional expectations of the doubled truncated distribution. *Annals of the Institute of Statistical Mathematics*, 48(3), 563–572.
- Sankaran, P. G., & Gupta, R. P. (1999). Characterization of lifetime distributions using measure of uncertainty. *Calcutta Statistical Association Bulletin*, 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. *Statistical Papers*, 45, 97–109.
- Sankaran, P. G., & Sunoj, S. M. (2017). Quantile based cumulative entropies. *Communications in Statistics – Theory and Methods*, 46(2), 805–814.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423.
- Simpson, J., Olsen, A., & Eden, J. (1975). A Baysian analysis of a multiplicative treatment effect in weather modification. *Technometrics*, *17*, 161–166.
- Sunoj, S. M., & Sankaran, P. G. (2012). Quantile based entropy function. *Statistics & Probability Letters*, 82, 1049–1053.
- Sunoj, S. M., Sankaran, P. G., & Maya, S. S. (2009). Chararacterization of life distributions using conditional expectations of doubly (Intervel)truncated random variables. *Communications in Statistics – Theory and Methods*, 38(9), 1441–1452.
- Sunoj, S. M., Sankaran, P. G., & Nanda, A. K. (2013). Quantile based entropy function in past lifetime. *Statistics & Probability Letters*, 83, 366–372.
- van Staden, P. J., & Loots, M. T. (2009). L-moment estimation for the generalized lambda distribution. Third annual ASEARC conference, New Castle.