Statistics and Probability Letters 82 (2012) 1049–1053

Contents lists available at SciVerse ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Quantile based entropy function

S.M. Sunoj *, P.G. Sankaran

Department of Statistics, Cochin University of Science and Technology, Cochin 682 022, Kerala, India

ARTICLE INFO

Article history:
Received 20 September 2011
Received in revised form 6 February 2012
Accepted 6 February 2012
Available online 3 March 2012

Keywords:
Shannon entropy
Residual lifetime
Quantile function
Reliability measures
Characterizations

ABSTRACT

Quantile functions are efficient and equivalent alternatives to distribution functions in modeling and analysis of statistical data (see Gilchrist, 2000; Nair and Sankaran, 2009). Motivated by this, in the present paper, we introduce a quantile based Shannon entropy function. We also introduce residual entropy function in the quantile setup and study its properties. Unlike the residual entropy function due to Ebrahimi (1996), the residual quantile entropy function determines the quantile density function uniquely through a simple relationship. The measure is used to define two nonparametric classes of distributions.

1. Introduction

In recent years, there has been a great interest in the measurement of uncertainty of probability distributions. Let X be a nonnegative absolutely continuous random variable (rv) representing the lifetime of a component with cumulative distribution function (CDF) $F(t) = P(X \le t)$ and survival function (SF) $\bar{F}(t) = P(X > t) = 1 - F(t)$. The measure of uncertainty defined by Shannon (1948) was

$$\xi(X) = \xi(f) = -\int_0^\infty (\log f(x)) f(x) dx = -E(\log f(X)), \tag{1}$$

where $f(t)$ is the probability density function (PDF) of X. Eq. (1) gives the expected uncertainty contained in $f(t)$ about the predictability of an outcome of X, which is known as Shannon entropy measure. Length of time during a study period has been considered as a prime variable of interest in many fields such as reliability, survival analysis, economics, business, etc. In such cases, the information measures are functions of time, thus they are dynamic. Based on this idea, Ebrahimi (1996) defined the residual Shannon entropy of X at time t as

$$\begin{aligned} \xi(X; t) = \xi(f; t) &= -\int_t^\infty \left(\frac{f(x)}{\bar{F}(t)}\right) \log\left(\frac{f(x)}{\bar{F}(t)}\right) dx, \\ &= \log \bar{F}(t) - \frac{1}{\bar{F}(t)} \int_t^\infty (\log f(x)) f(x) dx. \end{aligned} \tag{2}$$

Note that $\xi(X; t) = \xi(X_t)$, where $X_t = (X - t \mid X > t)$ is the residual time associated to X. By writing $h(t) = f(t)/\bar{F}(t)$, the failure rate function of X, (2) can equivalently be written as

$$\xi(X; t) = 1 - \frac{1}{\bar{F}(t)} \int_t^\infty (\log h(x)) f(x) dx. \tag{3}$$

* Corresponding author. Tel.: +91 484 2575893.
E-mail addresses: smsunoj@cusat.ac.in (S.M. Sunoj), sankaranpg@yahoo.com (P.G. Sankaran).

0167-7152/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2012.02.005

A similar function can be obtained in terms of the inactivity time $(t - X \mid X \le t)$ given in Di Crescenzo and Longobardi (2002). Interesting extensions and multivariate forms are also available in the literature. For additional information on these measures, we refer to Belzunce et al. (2004), Ebrahimi (1996), Ebrahimi and Kirmani (1996), Ebrahimi and Pellerey (1995), Nanda and Paul (2006) and Sunoj et al. (2009).

All these theoretical results and applications thereof are based on the distribution function. A probability distribution can be specified either in terms of the distribution function or by the quantile functions (QFs). Recently, it has been showed by many authors that QFs

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

are efficient and equivalent alternatives to the distribution function in modeling and analysis of statistical data (see Gilchrist, 2000; Nair and Sankaran, 2009). In many cases, QFs are more convenient as they are less influenced by extreme observations and thus provide a straightforward analysis with a limited amount of information. For a detailed and recent study on QF, its properties and its usefulness in the identification of models we refer to Lai and Xie (2006), Nair and Sankaran (2009), Nair et al. (2011), Sankaran and Nair (2009), Sankaran et al. (2010) and the references therein.

Although variety of research is available for various measures of uncertainty, a study of the same using QF does not appear to have been taken up. Also, many QFs used in applied works such as various forms of lambda distributions (Ramberg and Schmeiser, 1974; Freimer et al., 1998; Gilchrist, 2000; van Staden and Loots, 2009), the power-Pareto distribution (Gilchrist, 2000; Hankin and Lee, 2006), Govindarajulu distribution (Nair et al., 2011) etc. do not have tractable distribution functions. This makes the statistical study of the properties of $\xi(X)$ for these distributions by means of (1) difficult. Thus a formulation of the definition and properties of entropy function in terms of QFs is called for. Such a discussion has several advantages. Analytical properties of the entropy function obtained in this approach can be used as an alternative tool in modeling statistical data. Sometimes the quantile based approach is better in terms of tractability. New models and characterizations that are unresolvable in the distribution function approach can be resolved with the aid of quantile approach. Further, an explicit relationship between quantile entropy function and quantile density function in residual time can be derived.

The paper is organized as follows. In Section 2, we discuss some useful reliability measures in terms of quantile function. We introduce Shannon entropy function and residual entropy function in quantile setup and study their properties. Section 3 presents characterization results for certain lifetime quantile models based on the residual quantile entropy function.

2. Quantile based Shannon entropy

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 $fQ(u) = f(Q(u))$ (see Parzen, 1979) and quantile density function by $q(u) = Q'(u)$, where the prime denotes the differentiation, we have

$$q(u)fQ(u) = 1. \tag{5}$$

The hazard rate quantile function is defined by,

$$H(u) = hQ(u) = h(Q(u)) = (1-u)^{-1}fQ(u) = [(1-u)q(u)]^{-1}. \tag{6}$$

Following Nair and Sankaran (2009), $H(u)$ explains the conditional probability of failure in the next small interval of time given survival until $100(1-u)\%$ point of distribution. Like $h(t)$ that determines the CDF or SF uniquely, $H(u)$ also uniquely determines the QF by

$$Q(u) = \int_0^u \frac{dt}{(1-t)H(t)}.$$

From (5), the Shannon entropy defined in (1) can be written in terms of QF as

$$\xi(X) = \xi = \int_0^1 (\log q(p))\, dp. \tag{7}$$

Clearly, by knowing either $Q(u)$ or $q(u)$, the expression for $\xi(X)$ is quite simple to compute. An equivalent definition for the residual entropy (2) in terms of QF is given by

$$\xi Q(u) = \xi(X; Q(u)) = \log(1-u) + (1-u)^{-1}\int_u^1 (\log q(p))\, dp. \tag{8}$$

From (3), we can also write (8) as

$$\xi Q(u) = 1 - (1-u)^{-1}\int_u^1 \log H(p) dp. \tag{9}$$

$\xi Q(u)$ measures the expected uncertainty contained in the conditional density about the predictability of an outcome of X until $100(1-u)\%$ point of distribution. Further, differentiating Eq. (8) with respect to u, we get

$$\xi' Q(u) = -\frac{1}{(1-u)} + \frac{1}{(1-u)^2}\int_u^1 \log q(p) dp - \frac{1}{(1-u)}\log q(u),$$

Table 1
Quantile function and quantile residual entropy function for lifetime distributions.

Distribution	Quantile function	$\xi Q(u)$
Exponential	$\lambda^{-1}(-\log(1-u))$	$1-\log\lambda$
Pareto II	$\alpha[(1-u)^{-\frac{1}{c}}-1]$	$\ln\left(\frac{\alpha}{c}\right)+\left(\frac{c+1}{c}\right)-\frac{1}{c}\log(1-u)$
Rescaled beta	$R[1-(1-u)^{\frac{1}{c}}]$	$\log\left(\frac{R}{c}\right)+\left(\frac{c-1}{c}\right)+\frac{1}{c}\log(1-u)$
Half Logistic	$\sigma\log\left(\frac{1+u}{1-u}\right)$	$2+\log(2\sigma)-\frac{2\log 2}{(1-u)}+\frac{(1+u)}{(1-u)}\log(1+u)$
Power function	$\alpha u^{\frac{1}{\beta}}$	$\log\left(\frac{\alpha}{\beta}\right)+\left(\frac{\beta-1}{\beta}\right)+\log(1-u)+\left(\frac{\beta-1}{\beta}\right)\frac{u\log u}{(1-u)}$
Pareto I	$\sigma(1-u)^{-\frac{1}{\alpha}}$	$\log\left(\frac{\sigma}{\alpha}\right)+\left(\frac{\alpha+1}{\alpha}\right)-\frac{1}{\alpha}\log(1-u)$
Generalized Pareto	$\frac{b}{a}\left[(1-u)^{-\frac{a}{(a+1)}}-1\right]$	$\log\left(\frac{b}{a+1}\right)+\left(\frac{2a+1}{a+1}\right)-\left(\frac{a}{a+1}\right)\log(1-u)$
Log logistic	$\alpha^{-1}\left(\frac{u}{1-u}\right)^{\frac{1}{\beta}}$	$2-\log(\alpha\beta)+\left(\frac{\beta-1}{\beta}\right)\frac{u\log u}{(1-u)}-\frac{1}{\beta}\log(1-u)$
Exponential geometric	$\frac{1}{\lambda}\log\left(\frac{1-pu}{1-u}\right)$	$2+\log\left(\frac{1-p}{\lambda}\right)+p^{-1}(1-u)^{-1}[(1-p)\log(1-p)-(1-pu)\log(1-pu)]$
Linear hazard rate	$(a+b)^{-1}\log\left(\frac{a+bu}{a(1+u)}\right)$	$2+\log\left(\frac{b-a}{a+b}\right)+\log(1-u)-\frac{(a+b)}{b(1-u)}\log(a+b)+\frac{(a+bu)}{b(1-u)}\log(a+bu)-\frac{2\log 2}{(1-u)}+\frac{(1+u)}{(1-u)}\log(1+u)$
Davies	$Cu^{\lambda_1}(1-u)^{-\lambda_2}$	$\log C+\lambda_2-\lambda_1+1-(1-u)^{-1}(\lambda_2-\lambda_1)^{-1}\lambda_2\log\lambda_2-\lambda_2\log(1-u)-(1-u)^{-1}$ $(\lambda_1-1)u\log u-(1-u)^{-1}(\lambda_2-\lambda_1)^{-1}(\lambda_1(1-u)+\lambda_2 u)\log(\lambda_1(1-u)+\lambda_2 u)$

equivalently,

$$(1-u)\xi' Q(u)=-1+\xi Q(u)-\log(1-u)-\log q(u).$$

Thus,

$$q(u)=\exp\left\{\xi Q(u)-(1-u)\xi' Q(u)-\log(1-u)-1\right\}. \tag{10}$$

The relationship (10) determines the quantile density function from the quantile residual entropy $\xi Q(u)$. The relationship (10) is a unique characteristic of $\xi Q(u)$ unlike the residual entropy $\xi(X;t)$ in (3), where no such explicit relationship exists between $\xi(X;t)$ and $f(t)$. Table 1 provides the QFs and corresponding $\xi Q(u)$.

Now on the basis of residual quantile entropy (RQE) $\xi Q(u)$, we define the following nonparametric classes of life distributions.

Definition 1. X is said to have decreasing (increasing) residual quantile entropy (DRQE (IRQE)) if $\xi Q(u)$ is decreasing (increasing) in $u\geq 0$.

Now it is easy to show from the relationship (8) that if X is DRQE (IRQE), then $\xi Q(u)\leq(\geq)1+\log(q(u)(1-u))$. From the relationship (9) it follows that if X is DRQE (IRQE), then $\xi Q(u)\leq(\geq)1-\log H(u)$. Note that for the exponential distribution, $q(u)=\frac{1}{\lambda(1-u)}$ and $H(u)=\lambda$ so that $\xi Q(u)=1+\log(q(u)(1-u))$ and $\xi Q(u)=1-\log H(u)=1-\log\lambda$. Thus exponential distribution is the boundary of IRQE and DRQE classes.

Theorem 1. (a) *If X is IRQE and if ϕ is nonnegative, increasing and convex, then $\phi(X)$ is also IRQE.* (b) *If X is DRQE and if ϕ is nonnegative, increasing and concave, then $\phi(X)$ is also DRQE.*

Proof. If $g(y)$ is the pdf of $Y=\phi(X)$, then $g(y)=\frac{f(\phi^{-1}(y))}{\phi'(\phi^{-1}(y))}=\frac{fQ(u)}{\phi' Q(u)}=\frac{1}{q_X(u)\phi' Q(u)}$. We have

$$\begin{aligned}\xi_Y Q(u)&=\log(1-u)+\frac{1}{(1-u)}\int_u^1\log q_Y(p)dp,\\&=\log(1-u)+\frac{1}{(1-u)}\int_u^1\log\left(q_X(p)\phi' Q(p)\right)dp,\\&=\log(1-u)+\frac{1}{(1-u)}\int_u^1\log q_X(p)dp+\frac{1}{(1-u)}\int_u^1\log\phi' Q(p)dp\\&=\xi_X Q(u)+E[\log\phi'(X)\mid X>\phi^{-1}(u)],\end{aligned} \tag{11}$$

where $\xi_X Q(u)$ and $\xi_Y Q(u)$ are the RQEs of X and Y, respectively. Now if X is IRQE and if ϕ is nonnegative, increasing and convex, then $\phi(X)$ is also IRQE. Proof of (b) is similar. □

Example 1. Let X have the exponential distribution with failure rate λ and let $Y=X^{1/\alpha},\alpha>0$. Then Y has the Weibull distribution with $Q(u)=\lambda^{-1/\alpha}(-\log(1-u))^{1/\alpha}$. The nonnegative increasing function $\phi(x)=x^{1/\alpha}, x>0, \alpha>0$, is convex (concave) if $0<\alpha<1(\alpha>1)$. Hence due to Theorem 1, the Weibull distribution is IRQE (DRQE) if $0<\alpha<1(\alpha>1)$.

The concept of weighted distributions is usually considered in connection with modeling statistical data, where the usual practice of employing standard distributions is not found appropriate in some cases. In recent years, this concept has been applied in many areas of statistics, such as analysis of family size, human heredity, wildlife population study, renewal theory, biomedical, statistical ecology, reliability modeling, etc. Associated with an rv X having PDF $f(t)$, we can define the weighted rv X_w with density function $f_w(t) = \frac{w(t)f(t)}{Ew(X)}$, where $w(t)$ is a weight function with $0 < Ew(X) < \infty$. When $w(t) = t$, X_w is called length (size) biased rv. For recent works on weighted distributions, we refer the reader to Bartoszewicz (2009), Navarro et al. (2006) and Navarro et al. (2011). Using $f_w(t)$, the corresponding density quantile function is given by

$$f_wQ(u) = wQ(u)fQ(u)/\mu,$$

where $\mu = \int_0^1 wQ(p)fQ(p)dQ(p) = \int_0^1 wQ(p)dp$, or equivalently in the quantile density form $\frac{1}{q_w(u)} = \frac{wQ(u)}{\mu q(u)}$. Therefore, the RQE of X_w denoted by $\xi_wQ(u)$ is of the form

$$\xi_wQ(u) = \xi_XQ(u) + \log\mu - \frac{1}{(1-u)}\int_u^1 \log wQ(p)dp.$$

Then the following preservation theorem is immediate.

Theorem 2. (a) *If X is IRQE and if $w(X)$ is nonnegative, increasing and concave, then X_w is also IRQE.* (b) *X is DRQE and if $w(X)$ is nonnegative, increasing and convex, then X_w is also DRQE.*

The proof is similar to Theorem 1.

Definition 2. X is said to have less quantile entropy than Y if $\xi_XQ(u) \le \xi_YQ(u)$ for all $u \ge 0$. We write $X \overset{LQE}{\le} Y$.

It is interesting to note that if X and Y are two exponential rv's with failure rates λ_1 and λ_2, respectively, and if $\lambda_1 \le \lambda_2$, then $X \overset{LQE}{\le} Y$.

Theorem 3. *If $X \overset{QFR}{\le} Y$, (i.e., $H_F(u) \ge H_G(u)$), then $X \overset{LQE}{\le} Y$.*

Theorem 4. *If $X \overset{QFR}{\le} Y$, and if ϕ is nonnegative, increasing and convex, then $\phi(X) \overset{LQE}{\le} \phi(Y)$.*

Proof. Let $q_X(u)$, $q_X^*(u)$, $q_Y(u)$ and $q_Y^*(u)$ denote the quantile density function of X, $\phi(X)$, Y and $\phi(Y)$, respectively. Then by Eq. (11), for all $u \ge 0$,

$$\xi_{\phi(X)}Q(u) - \xi_{\phi(Y)}Q(u) = \xi_XQ(u) - \xi_YQ(u) + E[\log\phi'(X) \mid X > \phi^{-1}(u)] - E[\log\phi'(Y) \mid X > \phi^{-1}(u)]. \quad \square$$

Now, when $X \overset{QFR}{\le} Y$ using Theorem 3 we have $X \overset{LQE}{\le} Y$, and since ϕ is nonnegative, increasing and convex, we obtain $\phi(X) \overset{LQE}{\le} \phi(Y)$.

Remark 1. For equilibrium distribution with density function $f_E(t) = \bar{F}(t)/\mu$, the RQE is given by $\xi_EQ(u) = 1 + \log\mu$.

3. Characterizations

Since $\xi Q(u)$ uniquely determines the quantile density function $q(u)$ using (10), the characterizations of $\xi Q(u)$ for various distributions can be easily obtained from Table 1. For instance, generalized Pareto family is characterized by the relationship $\xi Q(u) = \xi + c\log(1-u)$, where c is a constant, for which $c = 0$ gives an exponential distribution and $c < (>)\ 0$ results Pareto II (rescaled beta) distributions, respectively. Characterizations of $\xi Q(u)$ for other distributions can be constructed in a similar fashion. Among various QFs given in Table 1, an important one is the Davies distribution proposed by Hankin and Lee (2006). It is a flexible family for right-skewed nonnegative data that provides good approximation to the exponential, gamma, lognormal and Weibull distributions and when $\lambda_1 = \lambda_2 = \lambda$, it becomes the log logistic distribution. Table 1 provides $\xi Q(u)$s for QFs that has closed form expressions, however, in some cases only $q(u)$ that has closed form expression. Accordingly, we prove a characterization theorem using $\xi Q(u)$, for a family of distributions represented by $q(u)$.

Theorem 5. *An rv X is distributed with quantile density function*

$$q(u) = Ku^{\alpha}(1-u)^{-(A+\alpha)} \tag{12}$$

for all u if and only if it satisfies the relationship

$$\xi Q(u) = \xi + [1-(A+\alpha)]\log(1-u) - \frac{\alpha u\log u}{(1-u)},$$

where α and A are real constants.

Remark 2. In family of distributions (12) some of its members have properties of nonmonotone hazard quantile functions and some have monotone hazard quantile functions. Further, it contains several well-known distributions which include the exponential ($\alpha = 0, A = 1$), Pareto ($\alpha = 0, A < 1$), rescaled beta ($\alpha = 0, A > 1$), the loglogistic distribution

($\alpha = \lambda - 1, A = 2$) and the life distribution proposed by Govindarajulu (1977) ($\alpha = \beta - 1, A = -\beta$), with QF $Q(u) = \theta + \sigma((\beta + 1)u^{\beta} - \beta u^{\beta+1})$. In terms of distribution function (12) has the form

$$f(x) = K[F(x)]^{-\alpha}[1 - F(x)]^{A+\alpha},$$

belongs to the class of distributions defined by the relationship between their density and distribution functions of Jones (2007) (for more details, see Nair et al., 2011).

Acknowledgments

We are grateful to the editor and referees for their constructive comments.

References

Bartoszewicz, J., 2009. On a representation of weighted distributions. Statist. Probab. Lett. 79, 1690–1694.
Belzunce, F., Navarro, J., Ruiz, J.M., del Aguila, Y., 2004. Some results on residual entropy function. Metrika 59, 47–161.
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., Kirmani, S.N.U.A., 1996. Some results on ordering of survival functions through uncertainty. Statist. Probab. Lett. 29, 167–176.
Ebrahimi, N., Pellerey, F., 1995. New partial ordering of survival functions based on notion of uncertainty. J. Appl. Probab. 32, 202–211.
Freimer, M., Mudholkar, S., Kollia, G., Lin, C.T., 1998. A study of the generalized lambda family. Comm. Statist. Theory Methods 17, 3547–3567.
Gilchrist, W., 2000. Statistical Modelling with Quantile Functions. Chapman and Hall/CRC, Boca Raton, FL.
Govindarajulu, A., 1977. A class of distributions useful in lifetesting and reliability with applications to nonparametric testing. In: Tsokos, C.P., Shimi, I.N. (Eds.), Theory and Applications of Reliability, Vol. 1. Academic Press, New York, pp. 109–130.
Hankin, R.K.S., Lee, A., 2006. A new family of non-negative distributions. Aust. NZ J. Stat. 48, 67–78.
Jones, M.C., 2007. On a class of distributions defined by the relationship between their density and distribution functions. Comm. Statist. Theory Methods 36, 1835–1843.
Lai, C.D., Xie, M., 2006. Stochastic Ageing and Dependence for Reliability. Springer, New York.
Nair, N.U., Sankaran, P.G., 2009. Quantile based reliability analysis. Comm. Statist. Theory Methods 38, 222–232.
Nair, N.U., Sankaran, P.G., Vinesh Kumar, B., 2011. Modelling lifetimes by quantile functions using Parzen's score function. Statistics 1, 1–13.
Nanda, A.K., Paul, P., 2006. Some properties of past entropy and their applications. Metrika 64, 47–61.
Navarro, J., Ruiz, J.M., del Aguila, Y., 2006. Multivariate weighted distributions. Statistics 40, 51–54.
Navarro, J., Sunoj, S.M., Linu, M.N., 2011. Characterizations of bivariate models using dynamic Kullback–Leibler discrimination measures. Statist. Probab. Lett. 81, 1594–1598.
Parzen, E., 1979. Nonparametric statistical data modeling. J. Amer. Statist. Assoc. 74, 105–122.
Ramberg, J.S., Schmeiser, B.W., 1974. An approximate method for generating asymmetric random variables. Comm. ACM 17, 78–82.
Sankaran, P.G., Nair, N.U., 2009. Nonparametric estimation of the hazard quantile function. J. Nonparametr. Stat. 21, 757–767.
Sankaran, P.G., Nair, N.U., Sreedevi, E.P., 2010. A quantile based test for comparing cumulative incidence functions of competing risks models. Statist. Probab. Lett. 80, 886–891.
Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423.
Sunoj, S.M., Sankaran, P.G., Maya, S.S., 2009. Characterization of life distributions using conditional expectations of doubly (interval) truncated random variables. Comm. Statist. Theory Methods 38 (9), 1441–1452.
van Staden, P.J., Loots, M.R., L-moment estimation for the generalized lambda distribution. Third Annual ASEARC Conference, New Castle, Australia, 2009.