ISSN 1973-2201 *STATISTICA, anno LXXXIV, n. 3, 2024*
DOI: https://doi.org/10.60923/issn.1973-2201/13446

INCOME INEQUALITY MEASURES - A QUANTILE APPROACH

P. G. Sankaran [1]
Department of Statistics, Cochin University of Science and Technology, Cochin, India
N. Unnikrishnan Nair
Department of Statistics, Cochin University of Science and Technology, Cochin, India

SUMMARY

Measures of income inequality are used for modelling and analysis of income data. In this paper, we present various income inequality measures in the quantile set up. We also introduce quantile version of well known dullness property. The interrelationships among these measures are investigated. The monotonic behaviour of income inequality measures are discussed. We also develop new quantile functions useful for income analysis. Various applications of the measures are discussed.

Keywords: Quantile functions; Income gap ratio; Lorenz curve; Dullness property; Mean proportional residual income; Harmonic mean proportional residual income.

1. INTRODUCTION

The study of poverty and affluence in a population has been and continues to be of considerable interest for researchers in economics and statistics. Consequently, various concepts for measuring income inequality have been developed in literature. The affluence and poverty are generally quantified in terms of the proportion of the rich and poor people and their incomes (Sen, 1976, 1986, 1988). Poverty and affluence indices based on the truncated means have been employed in such contexts by considering the duality between poverty and affluence. However, Sen (1988) mentioned the limitations of using truncated means in formulating affluence measures and proposed the harmonic mean as an alternative. Later, Belzunce *et al.* (1998) introduced a concept called mean proportional residual income which is closely related to the well known income gap ratio and used it to classify income distributions using the monotonic nature of the mean proportional residual income.

[1] Corresponding Author. E-mail:sankaran.p.g@gmail.com

It is an acknowledged fact that people in the higher income group under report their incomes to avoid higher amount of income tax. This phenomenon has been extensively investigated by income analysts to study its impact on the specification of the income distribution, for example, see Krishnaji (1970); Talwalker (1980) for some theoretical discussions. This property, referred to as dullness property, is an important tool for modelling and analysis of income data. For various applications of the dullness property we refer to Artikis *et al.* (1994); Slemrod (2007); Nataraja (2009).

In all the above mentioned aspects of income modelling, the income distribution is represented by its distribution function. A probability distribution can also be defined in terms of its quantile function and there are situations in which it enjoys more desirable properties compared to the distribution function. Those quantile functions which are highly flexible and do not possess a closed form distribution function, add to the richness of probability models. Sarabia (1997); Gilchrist (2000); Tarsitano (2004); Haritha *et al.* (2008); Beena and Kumaran (2010); Nair *et al.* (2013a) have used such quantile functions as income distributions. A review of quantile functions, income distributions, and income inequality measures is conducted by Varkey and Haridas (2023).

The primary objective of the present work is to demonstrate the role of the quantile function in the topics mentioned above so as to facilitate the analysis of incomes in a wider perspective. We propose quantile-based counterparts of inequality measures, dullness property, affluence index, etc., and their properties and provide examples as to how they can replace effectively the results in the distribution function approach.

The paper organized as follows. Section 2 presents quantile versions of income inequality measures. The interrelationships among various income inequality measures in the quantile set up are presented in Section 3. In Section 4, we present quantile version of dullness property and its complications to income inequality measures. The monotonic properties of income inequality measures are studied in Section 5. In Section 6, we present various applications of the study. Finally, Section 7 provides major conclusions of the study.

2. Basic Results

Let X be a non-negative random variable representing income of a population with absolutely continuous distribution function $F(x)$ and probability density function $f(x)$. Let $Q(u) = \inf_x\{F(x) \geq u\}$, $0 \leq u \leq 1$ be the quantile function X. Assume that $E(X) = \mu$ is finite. The mean of the distribution can be written as

$$\mu = \int_0^1 Q(p)dp = \int_0^1 (1-p)q(p)dp, \tag{1}$$

where $q(u)$ is the derivative of $Q(u)$, referred to as quantile density function. Instead of considering the entire incomes and the corresponding distributions, income analysts have found certain functions of income that characterize the distribution through the properties of the income distribution. Such properties can also help in finding suitable models of income. In this Section we present some such functions within the quantile framework so as to provide alternative methodology for the modelling and analysis of incomes. Individuals whose incomes are above a chosen level t^* are considered to be affluent. The income distribution of the affluent about the line $X = t^*$ is based on left truncated random variable $X(t^*) = \{X|X \geq t^*\}$. As a measure of affluence in population, the income gap ratio among the rich is defined with reference to the affluence line $X = t^*$ (Sen, 1986) as

$$\beta(t^*) = 1 - \frac{t^*}{E(X(t^*))}. \tag{2}$$

In the quantile function $Q(u)$ of X, we can modify Eq. (2) as

$$B^*(u) = \beta(Q(u)) = 1 - \frac{(1-u)Q(u)}{\int_u^1 Q(p)dp}. \tag{3}$$

When quantile functions which has no closed form distribution functions are used, formula in Eq. (2) is difficult to apply and Eq. (3) is needed in such situations. Definitions in Equations (2) and (3) have analogous measures used in the analysis of poverty and income data of the poor when the truncation is reversed. However, Sen (1988) pointed out the limitations of Eq. (2) when a few very large incomes or an outlier may adversely affect the estimate of $\beta(t^*)$ ($B^*(u)$) and to overcome this difficulty suggested the use of the harmonic mean instead of the mean in Eq. (2).

The harmonic mean of X is

$$H = \left(E\left(\frac{1}{X}\right)\right)^{-1} = \left[\int_0^1 \frac{1}{Q(p)} dp\right]^{-1}. \tag{4}$$

For the truncated income $X(t^*)$, harmonic mean is

$$h(t^*) = \left(\frac{1 - F(t^*)}{\int_{t^*}^{\infty} \frac{1}{x} f(x)dx}\right), \tag{5}$$

which is equivalent to

$$H^*(u) = h(Q(u)) = \left(\frac{1-u}{\int_u^1 \frac{1}{Q(p)} dp}\right). \tag{6}$$

Analogous to the income gap ratio, replacing the mean by harmonic mean, the harmonic mean proportional residual income (HMPRI) defined by

$$p(t^*) = \frac{h(t^*)}{t^*}, \tag{7}$$

has the quantile version

$$P^*(u)=p(Q(u))=\left(\frac{Q(u)}{1-u}\int_u^1 \frac{1}{Q(p)}dp\right)^{-1}=\frac{1-u}{Q(u)\int_u^1 \frac{dp}{Q(p)}}. \tag{8}$$

Now the income gap ratio for the affluent based on harmonic mean $h(t^*)$ given in Sen (1988),

$$r(t^*)=1-\frac{t^*}{h(t^*)}, \tag{9}$$

takes the quantile form

$$R^*(u)=r(Q(u))=1-\frac{Q(u)}{(1-u)}\int_u^1 \frac{dp}{Q(p)}. \tag{10}$$

From Eq. (8), it follows that

$$R^*(u)=1-\frac{1}{P^*(u)}. \tag{11}$$

A second measure proposed is the mean left proportional residual income (MPRI) defined in Belzunce *et al.* (1998) as

$$a(t^*)=E\left(\frac{X(t^*)}{t^*}\right), \tag{12}$$

which is related to $\beta(t^*)$ as

$$\beta(t^*)=1-\frac{1}{a(t^*)}. \tag{13}$$

The quantile version of MPRI is obtained as

$$A^*(u)=a(Q(u))=\frac{1}{Q(u)(1-u)}\int_u^1 Q(p)dp. \tag{14}$$

Now, it follows that

$$B^*(u)=1-\frac{1}{A^*(u)}. \tag{15}$$

In reliability, mean residual life $m(t^*)=E(X-t^*|X>t^*)=E(X(t^*))-t^*$ is an important concept to study failure patterns in lifetime data. Nair and Sankaran (2009) introduced quantile version of $m(t^*)$ as

$$M(u)=m(Q(u))=\frac{1}{(1-u)}\int_u^1 (Q(p)-Q(u))dp. \tag{16}$$

It is easy to see that

$$A^*(u) = \frac{M(u)}{Q(u)} + 1. \tag{17}$$

Eq. (17) helps to translate the properties of $M(u)$ to $A^*(u)$, and hence such results are not included here.

We now show that $P^*(u)$ uniquely determines the distribution $Q(u)$.

THEOREM 1. *Let X be a non negative continuous random variable with $E(\frac{1}{X}) < \infty$. Then $P^*(u)$ determines $Q(u)$, the quantile function of X uniquely.*

PROOF. From Eq. (8), we get

$$P^*(u)Q(u)\int_u^1 \frac{dp}{Q(p)} = (1-u). \tag{18}$$

Differentiating Eq. (18) with respect to u, we get

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

which leads to

$$\frac{P^{*'}(u)}{P^*(u)} + \frac{q(u)}{Q(u)} = \frac{P^*(u)}{(1-u)} - \frac{1}{(1-u)}. \tag{19}$$

Now integrating Eq. (19) over $(0, u)$, we get

$$\log\frac{P^*(u)}{P^*(0)} + \log Q(u) = \int_0^u \frac{P^*(p)}{(1-p)}dp + \log(1-u),$$

which gives

$$Q(u) = \frac{P^*(0)(1-u)}{P^*(u)} e^{\int_0^u \frac{P^*(p)}{(1-p)}dp}. \tag{20}$$

□

REMARK 2. *Theorem 1 is useful for deriving new models. Suppose that $P^*(u) = \frac{1}{au+b}$, a reciprocal linear function in u. Then*

$$Q(u) = \left(\frac{au+b}{b}\right)^{\frac{1-a-b}{a+b}} (1-u)^{\frac{a+b-1}{a+b}}, \quad 1-a-b > 0.$$

The quantile function $Q(u)$ is the product of two quantile functions; the first term represents a power distribution and the second term is the quantile function of a Pareto I distribution. This distribution has no closed form distribution function and hence cannot be obtained directly from the definition based on the distribution function. It also demonstrates the need for Theorem 1 to generate new useful models and also the advantage of the quantile approach.

In the following, we can show that $A^*(u)$ uniquely determines $Q(u)$.

THEOREM 3. *Let X be a non negative continuous random variable with $E(X) < \infty$. Then $A^*(u)$ uniquely determines $Q(u)$.*

PROOF. From the definition of $A^*(u)$, we get

$$A^*(u)Q(u)(1-u) = \int_u^1 Q(p)dp. \tag{21}$$

Differentiating Eq. (21) with respect to u, we obtain

$$\frac{-q(u)}{Q(u)} = \frac{1 + A^{*'}(u)(1-u) - A^*(u)}{A^*(u)(1-u)}, \tag{22}$$

where $A^{*'}(u)$ is the derivative of $A^*(u)$.

Integrating Eq. (22) over $(0, u)$ we get

$$Q(u) = \frac{A^*(0)}{A^*(u)(1-u)} e^{-\int_0^u \frac{dp}{A^*(p)(1-p)}}. \tag{23}$$

□

COROLLARY 4. *The constancy of $A^*(u)$ ($A^*(u) = k, k > 1$) characterizes Pareto I with $Q(u) = (1-u)^{-\frac{(k-1)}{k}}; k > 1$.*

REMARK 5. *Theorem 3 is useful to characterize quantile functions. For example, the distribution of X is specified by*

$$Q(u) = -(c+\mu)\log(1-u) - 2cu, \quad \mu > 0, -\mu \le c < \mu$$

if and only if

$$A^*(u) = \frac{cu + (c+\mu)\log(1-u) - \mu}{2cu + (c+\mu)\log(1-u)}.$$

The properties of the above distribution are discussed in Midhu et al. (2013). The latter part of Remark 2 applies to this case, as well.

REMARK 6. *Theorem 3 is useful for deriving new quantile function models for income based on empirical properties. For example, when $A^*(u) = cu + d$, then*

$$Q(u) = d^{\frac{c+d+1}{c+d}}(1-u)^{\frac{1}{c+d}-1}(cu+d)^{-\frac{c+d+1}{c+d}}, \quad c, d > 0.$$

EXAMPLE 7. *Consider the quantile function given by*

$$Q(u)=\left(\frac{au+b}{b}\right)^{\frac{1-a-b}{a+b}}(1-u)^{\frac{a+b-1}{a+b}},\quad 1-a-b>0.$$

Then, for this quantile function, we can obtain

$$B^*(u)=1-\frac{(au+b)^{\frac{1-a-b}{a+b}}(1-u)^{\frac{2(a+b)-1}{a+b}}}{(a+b)^{\frac{1-2(a+b)}{a+b}}Beta\left(\frac{a(1-u)}{a+b},2-\frac{1}{a+b},\frac{1}{a+b}\right)},$$

$$H^*(u)=\frac{1}{(au+b)}\left(\frac{au+b}{b}\right)^{\frac{1-a-b}{a+b}}(1-u)^{\frac{a+b-1}{a+b}},$$

$$P^*(u)=\frac{1}{au+b},$$

$$R^*(u)=1-(au+b)$$

and

$$A^*(u)=\frac{(a+b)^{\frac{1-2(a+b)}{a+b}}Beta\left(\frac{a(1-u)}{a+b},2-\frac{1}{a+b},\frac{1}{a+b}\right)}{(au+b)^{\frac{1-a-b}{a+b}}(1-u)^{\frac{2(a+b)-1}{a+b}}}.$$

3. RELATIONSHIPS TO OTHER INCOME INEQUALITY MEASURES

A well known classical measure of income inequality is the Lorenz curve given by

$$L(u)=\frac{1}{\mu}\int_0^u Q(p)dp. \tag{24}$$

The curve $L(u)$ is increasing and convex with $L(0)=0$, $L(1)=1$ and $L(u)\le u$. Various functions given in the last Section can be expressed in terms of $L(u)$. For example, the $A^*(u)$ curve is related to the Lorenz curve by the identity

$$\begin{aligned}L(u)&=1-\frac{1}{\mu}\int_u^1 Q(p)dp\\&=1-\frac{A^*(u)(1-u)Q(u)}{\mu}.\end{aligned}$$

This gives two unique representations

$$A^*(u)=\frac{1-L(u)}{(1-u)L'(u)}$$

and

$$L(u) = 1 - \exp[-\int_0^u [(1-p)A^*(p)]^{-1} dp]. \quad (25)$$

Apart from the non-parametric version of income inequality presented by the Lorenz curve, there have been attempts to derive parametric versions of inequality measures. One approach in the latter case is to assume a family of income distributions and then to obtain its Lorenz curve analytically. See the literature from Kakwani and Podder (1973) to Gómez-Déniz *et al.* (2021). Such a method was adopted by Sarabia (1997) by considering the quantile function of the generalized lambda distribution of Ramberg and Schmeiser (1974) and pointed out that it can generate a hierarchy of Lorenz curves. A similar work is done in Haritha *et al.* (2008) using the modified lambda model of Freimer *et al.* (1988). With more general and flexible models in the form of quantile functions are available in literature, this approach can be further strengthened to find more realistic Lorenz curves which are easier to estimate. For example, consider a parametric Lorenz curve given in Sarabia *et al.* (2017) with

$$L(u) = 1 - (1-u^a)^{\frac{1}{a}}, \quad a \geq 1. \quad (26)$$

The quantile function for Eq. (26) is

$$Q(u) = \mu(1-u^{-a})^{-\frac{(a-1)}{a}},$$

where $\mu = EX$.

Then we can get,

$$A^*(u) = \frac{\mu(1-L(u))}{(1-u)Q(u)} = \frac{(1-u^a)^{\frac{1}{a}}(1-u^{-a})^{\frac{a-1}{a}}}{(1-u)}$$

and

$$B^*(u) = 1 - \frac{(1-u)}{(1-u^a)^{\frac{1}{a}}(1-u^{-a})^{\frac{a-1}{a}}}.$$

Another concept employed to measure income inequality is the Bonferroni curve. From Nair *et al.* (2013b), Bonferroni curve is defined by

$$B(u) = \frac{\int_0^u Q(p)dp}{u\mu}, \quad (27)$$

which is closely related to $L(u)$ by

$$uB(u) = L(u).$$

Thus $B(u)$ is related to $A^*(u)$ by the relation

$$A^*(u)(1-u)Q(u) = \mu(1-uB(u)) \quad (28)$$

and hence we have

$$A^*(u) = \frac{(1-uB(u))}{(1-u)(B(u)+B'(u))} \quad \text{and} \quad B(u) = u[1-\exp[-\int_0^u (1-p)A^*(p)dp]].$$

A third important measure used in informatics as well as in income analysis is the Leimkuhler curve $K(u)$ (Sarabia, 2008). It represents the cumulative proportion of productivity against cumulative proportion of sources. We define $K(u)$ as

$$K(u) = \frac{1}{\mu}\int_{1-u}^{1} Q(p)dp. \tag{29}$$

Thus

$$K(1-u) = \frac{1}{\mu}\int_{u}^{1} Q(p)dp.$$

Then $A^*(u)$ is related to $K(u)$ by

$$A^*(u)(1-u)Q(u) = \mu K(1-u). \tag{30}$$

More recently, Zenga (2007) proposed a new measure to compare the mean income of the poorest income earners with the mean income of the remaining richest part of the population. The Zenga curve $c(t)$ is defined by

$$c(t) = 1 - \frac{E(X|X \le t)}{E(X|X > t)}.$$

In terms of quantiles, $c(t)$ becomes (Nair *et al.*, 2012)

$$\begin{aligned} C^*(u) = c(Q(u)) &= 1 - \frac{(1-u)}{u}\frac{\int_0^u Q(p)dp}{\int_u^1 Q(p)dp} \\ &= 1 - \frac{(1-u)(\mu - \int_u^1 Q(p)dp)}{u\int_u^1 Q(p)dp}. \end{aligned} \tag{31}$$

Now $A^*(u)$ and $C^*(u)$ are related by

$$C^*(u) = \frac{1}{u}\left[1 - \frac{\mu}{A^*(u)Q(u)}\right]. \tag{32}$$

which can be written in terms of $A^*(u)$ only using Eq. (23).

Now we investigate relationship between $A^*(u)$ and various income inequality indices. One of the popular indices is Gini index G which is defined by

$$
\begin{aligned}
G &= 2\int_0^1 (p - L(p))dp \\
&= 2\int_0^1 \left(\frac{A^*(p)(1-p)Q(p)}{\mu} - (1-p)\right)dp \\
&= 2\int_0^1 (1-p)\left(\frac{A^*(p)Q(p)}{\mu} - 1\right)dp.
\end{aligned}
\tag{33}
$$

Similarly, one can write Bonferroni index B given by

$$B = 1 - \int_0^1 B(p)dp,$$

in terms of $A^*(u)$.

From Eq. (28), we get

$$B = 1 - \int_0^1 \left(1 - \frac{A^*(p)(1-p)Q(p)}{\mu}\right)\frac{dp}{p}. \tag{34}$$

The Zenga index Z given in Arnold (2014) is obtained as

$$
\begin{aligned}
Z &= \int_0^1 C^*(u)du \\
&= \log u - \mu \int_0^1 \frac{du}{uA^*(u)Q(u)}.
\end{aligned}
$$

Thus the quantile functional form of $Q(u)$ along with $A^*(u)$ from Eq. (14) is enough to find most of the income measures.

4. Dullness Property in Quantile Set up

Let X be a non-negative continuous random variable as mentioned in Section 2. In this Section we present the quantile-based definition of the dullness property of income distribution and derive some new results.

Definition 8. *The distribution of X is said to have dullness property if*

$$P(X \geq st | X \geq t) = P(X \geq s);\ s, t > 1. \tag{35}$$

This means that the conditional probability that the true income X is at least s times the reported value t is same as the unconditional probability that X has at least income s. As mentioned in the introduction, under reporting of income occur among the affluent and therefore to have realistic assessment of the income distribution one should know the under reported income or reporting error. It has been observed that the distribution of reporting error in income is independent of the reported value. We now present quantile version of Eq. (35). To derive this, we can write Eq. (35) as

$$F(st) = 1-(1-F(s))(1-F(t)). \tag{36}$$

Setting $F(s) = u$ and $F(t) = v$, we obtain $s = Q(u)$ and $t = Q(v)$ so that Eq. (36) leads to

$$st = F^{-1}(1-(1-F(s))(1-F(t))). \tag{37}$$

Since $F(Q(u)) = u$ and $F(Q(v)) = v$, we get

$$Q(u)Q(v) = Q(1-(1-u)(1-v)); \; 0 < u, v < 1. \tag{38}$$

The identity in Eq. (38) is the quantile version of dullness property. The property Eq. (38) is the multiplicative version of the well known lack of memory property given in Nair *et al.* (2013a) in the quantile set up.

The relationship among dullness property, $A^*(u)$ and Pareto I distribution is given below.

THEOREM 9. *Let X be a non-negative random variable as stated in Section 2 with quantile function $Q(u)$. Assume that $\mu = E(X) < \infty$. Then following results are equivalent:*

(i) *The distribution of X satisfies dullness property in Eq. (38).*

(ii) $A^*(u) = \mu$.

(iii) *The distribution of X is Pareto I.*

PROOF. We prove (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (i). To prove (i) $\Longrightarrow$ (ii), we note that

$$Q(u)Q(v) = Q(1-(1-u)(1-v)); \; 0 < u, v < 1.$$

Integrating above over $v \in [0,1]$, we get

$$Q(u)\int_0^1 Q(v)dv = \int_0^1 Q(1-(1-u)(1-v))dv,$$

which leads to

$$Q(u)\mu = \frac{\int_u^1 Q(p)dp}{(1-u)}.$$

Thus

$$A^*(u) = \mu.$$

To prove (ii) $\implies$ (iii), we note that

$$Q(u)\mu(1-u) = \int_u^1 Q(p)dp. \tag{39}$$

Differentiating Eq. (39) with respect to u, we get

$$[q(u)(1-u) - Q(u)]\mu = -Q(u),$$

which implies

$$\frac{q(u)}{Q(u)} = \frac{\mu - 1}{\mu(1-u)}.$$

We now integrate over $(0, u)$, we obtain

$$Q(u) = (1-u)^{-\frac{(\mu-1)}{\mu}}, \tag{40}$$

which is the quantile function of Pareto I.

The proof for (iii) $\implies$ (i) follows by substituting Eq. (40) in the identity of Eq. (38). □

Since it is more advantageous to use the harmonic mean instead of the arithmetic mean when measurement of affluence and related concepts are studied, a similar result based on harmonic mean $P^*(u)$ is worth investigating.

THEOREM 10. *Let X be a non-negative random variable as stated in Section 2. Assume that $\left(E\left(\frac{1}{X}\right)\right)^{-1} = \mu^* < \infty$. The following statements are equivalent.*

(i) *The distribution of X satisfies dullness property (38).*

(ii) $P^*(u) = \mu^*$.

(iii) *X follows the Pareto I distribution.*

PROOF. The proof is similar to that of Theorem 9. □

We now examine the behaviour of income inequality indices when dullness property prevails. When $A^*(u) = \mu$, we get Eq. (25) as

$$L(u) = 1 - (1-u)Q(u)$$

$$\text{or} \quad Q(u) = \frac{1 - L(u)}{1-u}. \tag{41}$$

On similar lines, $A^*(u) = \mu$ leads to

$$B(u) = \frac{1-(1-u)Q(u)}{u}, \tag{42}$$

$$K(1-u) = (1-u)Q(u) \tag{43}$$

and

$$\begin{aligned} C^*(u) &= \frac{1}{u}\left[1 - \frac{1}{Q(u)}\right] \\ &= \frac{Q(u)-1}{uQ(u)}. \end{aligned} \tag{44}$$

5. MONOTONICITY OF RESIDUAL INCOMES

In this Section, we discuss monotonic properties of $A^*(u)$ and $P^*(u)$. Throughout this Section, prime denotes derivative with respect to u.

DEFINITION 11. *The income distribution* $Q(u)$ *is said to have increasing (decreasing)* $Q(u)$ *(IQMPRI/DQMPRI) if* $A^*(u)$ *is non-decreasing (non-increasing) in* u.

In other words, $Q(u)$ has IQMPRI (DQMPRI) if $A^{*'}(u) > (<) 0$.

For Pareto I distribution, $A^*(u)$ is constant which means that $Q(u)$ has both IQMPRI and DQMPRI.

EXAMPLE 12. *When X follows exponential distribution with* $Q(u) = -\frac{1}{\lambda}\log(1-u)$, *we obtain*

$$A^*(u) = 1 - \frac{1}{\log(1-u)}.$$

Thus $A^*(u)$ *is non-decreasing in u and* $Q(u)$ *has DQMPRI.*

EXAMPLE 13. *Let X has a power distribution with* $Q(u) = \frac{1}{\theta}\log(u)$. *Then*

$$A^*(u) = -\frac{u}{1-u} - \frac{1}{\log u},$$

has derivative

$$A^{*'}(u) = \frac{(1-u)^2 - u(\log u)^2}{u(1-u)^2(\log u)^2}.$$

Since $(1-u)^2 > u(\log u)^2$, *we get* $A^{*'}(u) > 0$ *so that* $Q(u)$ *has IQMPRI.*

EXAMPLE 14. *When X has lambda distribution with $Q(u) = u^{\beta} - (1-u)^{\beta}; \beta > 0$, then*

$$A^*(u) = \frac{1 - u^{\beta+1} - (1-u)^{\beta+1}}{(\beta+1)(1-u)(u^{\beta} - (1-u)^{\beta})}.$$

Then $A^{*'}(u) > 0$ so that X has IQMPRI.

DEFINITION 15. *The distribution of X has decreasing (increasing) mean residual life (DMRL/IMRL) if $M(u)$ is non-increasing (non-decreasing) in u (Nair et al., 2013a).*

THEOREM 16. *Let X be a non-negative continuous random variable with mean residual life $M(u)$ and QMPRI $A^*(u)$. Then following results hold.*

(i) *DMRL $\Longrightarrow$ DQMPRI*

(ii) *IQMPRI $\Longrightarrow$ IMRL*

PROOF. From $A^*(u) = \frac{M(u)}{Q(u)} + 1$, we get the derivative of $A^*(u)$ as

$$A^{*'}(u) = \frac{Q(u)M'(u) - q(u)M(u)}{(Q(u))^2}. \tag{45}$$

When X has DMRL, we get $M'(u) < 0$, so that Eq. (45) leads to $A^{*'}(u) < 0$. Thus X has DQMPRI.
To prove (ii), note that

$$A^*(u)Q(u) = M(u) + Q(u).$$

Differentiating above with respect to u, we obtain

$$A^{*'}(u)Q(u) + q(u)[A^*(u) - 1] = M'(u),$$

which implies

$$M'(u) = A^{*'}(u)Q(u) + \frac{q(u)M(u)}{Q(u)}.$$

When $A^{*'}(u) > 0$, then $M'(u) > 0$. Thus X has IMRL. □

The DMRL class is discussed extensively in reliability analysis and it is a subclass of DQMPRI and IMRL class is a super class of IQMPRI. Note that monotonic behaviour of $A^*(u)$ and quantile income gap ratio $B^*(u)$ are identical.

DEFINITION 17. *The income distribution $Q(u)$ is said to have increasing (decreasing) income gap ratio (IIGR/DIGR) property if $B^*(u)$ is non-decreasing (non-increasing) in u.*

The Pareto I distribution has both IIGR and DIGR property in the sense that its IGR is a constant.

6. Applications

In income studies, identification of an income model based on physical characteristics of the data is an interesting problem. The dullness property in terms of $A^*(u)$ ($P^*(u)$) can be employed in such contexts. We define

$$D^*(u) = \frac{A^*(u)}{\mu}.$$

Then from Theorem 9, it follows that $D^*(u) = 1$ characterizes Pareto I distribution. A rough estimate of $D^*(u)$ may provides adequate information on the form of $D^*(u)$. Non-parametric estimator of $A^*(u)$, using empirical quantile function is employed to get an estimator of $D^*(u)$. Further, income models can be classified on the basis of the values of $D^*(u)$, as explained in the following discussions.

DEFINITION 18. *The distribution of X is said to be positively (negatively) dull if $D^*(u) > (<) 1$ for all u, $0 < u < 1$. It is completely dull when $D^*(u) = 1$.*

DEFINITION 19. *The distribution of X is said to have mixed type of dullness (positive at certain regions and negative at other regions) if $D^*(u)$ crosses the line $u = 1$ at least once.*

There is a useful interpretation for the above classification. In the class of distributions, for which $D^*(u) > 1$, the behaviour is such that as the residual income beyond the tax level increases, the under reporting error tends to become larger. On the other hand, for lesser incomes, the under reporting error assumes a smaller percentage in excess of the amount at the tax exemption level.

EXAMPLE 20. *When X has exponential distribution with $Q(u) = -\frac{1}{\lambda}\log(1-u)$, we obtain*

$$D^*(u) = \lambda\Big(1 - \frac{1}{\log(1-u)}\Big).$$

The $D^*(u)$ depends on the values of λ. Suppose $0 < \lambda < 1$, then $D^*(u) > 1$ if $0 < u < 1 - e^{-\frac{\lambda}{1-\lambda}}$ and $D^*(u) < 1$ if $1 - e^{-\frac{\lambda}{1-\lambda}} < u < 1$. For $\lambda > 1$, $D^*(u) > 1$. Thus $D^*(u)$ has mixed dull property if $0 < \lambda < 1$.

EXAMPLE 21. *Let X has generalized Pareto distribution with $Q(u) = \frac{b}{a}[(1-u)^{-\frac{a}{a+1}} - 1]$; $a, b > 0$. Then*

$$A^*(u) = 1 + \frac{a(1-u)^{-\frac{a}{a+1}}}{(1-u)^{-\frac{a}{a+1}} - 1} \text{ and } \mu = \frac{b}{a}.$$

Thus

$$D^*(u) = \frac{(a+1)(1-u)^{-\frac{a}{a+1}}}{b[(1-u)^{-\frac{a}{a+1}} - 1]} = \frac{(a+1)}{b[1-(1-u)^{\frac{a}{a+1}}]}.$$

If $b \le a+1$, $D^*(u) > 1$. Thus $Q(u)$ has positive dullness. However, if $b > a+1$, $D^*(u)$ crosses the line $u = 1$ so that $Q(u)$ has mixed dullness.

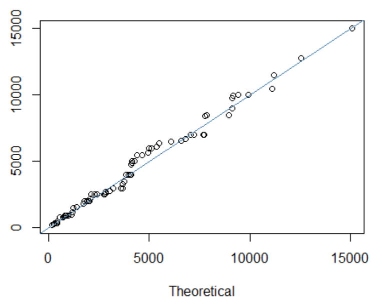

Figure 1 – Q-Q plot for the dataset.

A second application of the study is to employ $A^*(u)$ for comparing two quantile function models by a test of hypothesis. To test $H_0 : Q_1(u) = Q_2(u)$ for all u against the alternative $H_1 : Q_1(u) < (>) Q_2(u)$ for at least one u, we can formulate hypothesis based on $A^*(u)$ as $H_0^{'} : A_1^*(u) = A_2^*(u)$ for all u against $H_1^{'} : A_1^*(u) < (>) A_2^*(u)$ for at least one u, where $A_1^*(u)$ and $A_2^*(u)$ are QMPRI corresponding to $Q_1(u)$ and $Q_2(u)$ respectively. We then define

$$T = \sup_u (\hat{A}_1^*(u) - \hat{A}_2^*(u)),$$

where $\hat{A}_i^*(u)$ is the non-parametric estimator of $A_i^*(u)$, $i = 1,2$. We then reject H_0 if T is large. The cut off values of the test T can be calculated using the asymptotic distribution of $\hat{A}_i^*(u)$, $i = 1,2$, which is normal for fixed u.

To illustrate the importance of the proposed quantile based concepts, we consider a data on household income of people in Malaysia. Sulaiman *et al.* (2020) carried out a detailed analysis of the data and identified various causes of the rising cost of living. We considered an extract of the data obtained from randomly selected 73 respondents of the Johor state of Malaysia. The quantile function given by Midhu *et al.* (2013),

$$Q(u) = -(c+\mu)\log(1-u) - 2cu, \quad \mu > 0, -\mu \le c < \mu,$$

is employed. The method of L-moments is used to estimate the unknown parameters μ and c. The estimates are $\hat{\mu} = 4550.89$ and $\hat{c} = -1590.95$. The Q-Q plot is presented in Figure 1 and it is concluded that the model is a good fit to the given data.

We also plot the quantities $B^*(u)$, $P^*(u)$, $R^*(u)$ and $A^*(u)$ in Figure 2 to 5 respectively. It is easy to see from figures that $B^*(u)$, $P^*(u)$, $R^*(u)$ and $A^*(u)$ are decreasing. The $D^*(u)$ is also plotted in Figure 6 and it follows that the distribution has a mixed dull property.

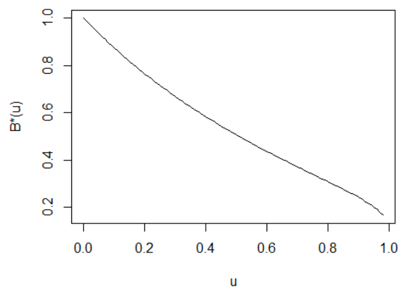

Figure 2 – $B^{*}(u)$ for the dataset.

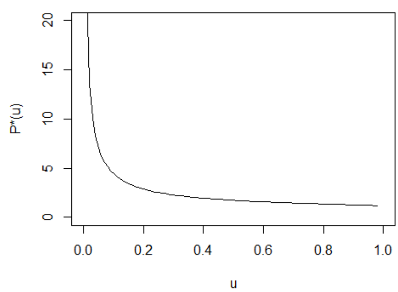

Figure 3 – $P^{*}(u)$ for the dataset.

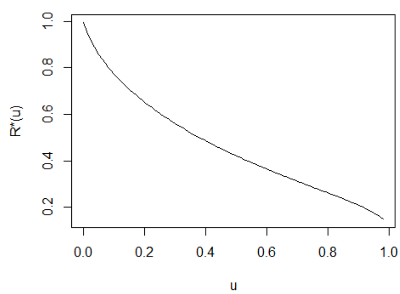

Figure 4 – $R^{*}(u)$ for the dataset.

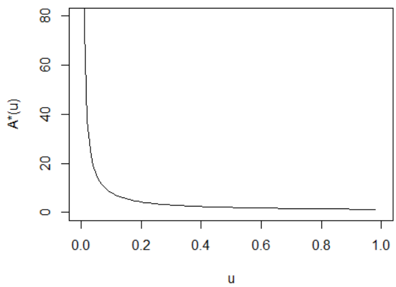

Figure 5 – $A^*(u)$ for the dataset.

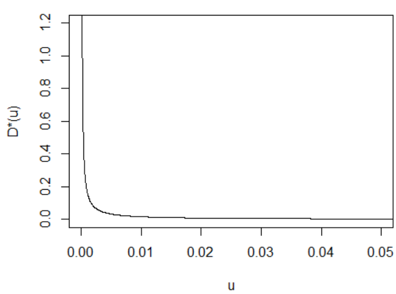

Figure 6 – $D^*(u)$ for the dataset.

7. Conclusion

Modelling and analysis of income data are extensively discussed in literature. In this paper, we discussed the quantile-based income analysis. We have introduced quantile versions of various income inequality measures. Properties of the income distributions using mean proportional residual income were studied. Various applications of the results were discussed.

Two income models can be compared using $A^*(u)$ ($P^*(u)$). Stochastic ordering of the models using $A^*(u)$ ($P^*(u)$) is an area of study to be explored. The non-parametric estimator of $A^*(u)$ can be derived from the non-parametric estimator of $M(u)$, which was studied by Sankaran and Midhu (2016). The work in these directions will be reported in a separate paper.

Acknowledgements

We thank the editor and the referees for their constructive comments.

References

B. C. Arnold (2014). *On Zenga and Bonferroni curves*. Metron, 73, no. 1, pp. 25–30.

T. Artikis, A. Voudouri, M. Malliaris (1994). *Certain selecting and underreporting processes*. Mathematical and Computer Modelling, 20, no. 1, pp. 103–106.

V. T. Beena, M. Kumaran (2010). *Measuring inequality and social welfare from any arbitrary distribution*. Brazilian Journal of Probability and Statistics, 24, no. 1, pp. 78–90.

F. Belzunce, J. Candel, J. M. Ruiz (1998). *Ordering and asymptotic properties of residual income distributions*. Sankhyā: The Indian Journal of Statistics, Series B, pp. 331–348.

M. Freimer, G. Kollia, G. S. Mudholkar, C. T. Lin (1988). *A study of the generalized Tukey lambda family*. Communications in Statistics, Theory and Methods, 17, no. 10, pp. 3547–3567.

W. Gilchrist (2000). *Statistical Modelling with Quantile Functions*. Chapman and Hall/CRC, New York.

E. Gómez-Déniz, J. M. Sarabia, V. Jordá (2021). *Parametric Lorenz curves based on the beta system of distributions*. Communications in Statistics, Theory and Methods, pp. 1–20.

N. H. Haritha, N. U. Nair, K. R. M. Nair (2008). *Modelling income using the generalised lambda distribution*. Journal of Income Distribution, 17, no. 2, pp. 37–51.

N. C. KAKWANI, N. PODDER (1973). *On the estimation of Lorenz curves from grouped observations*. International Economic Review, pp. 278–292.

N. KRISHNAJI (1970). *A characteristic property of the Yule distribution*. Sankhyā: The Indian Journal of Statistics, Series A, pp. 343–346.

N. N. MIDHU, P. G. SANKARAN, N. U. NAIR (2013). *A class of distributions with the linear mean residual quantile function and it's generalizations*. Statistical Methodology, 15, pp. 1–24.

N. U. NAIR, K. R. M. NAIR, N. SREELAKSHMI (2012). *Some properties of the new Zenga curve*. Statistica and Applicazioni X, pp. 43–52.

N. U. NAIR, P. G. SANKARAN (2009). *Quantile-based reliability analysis*. Communications in Statistics, Theory and Methods, 38, no. 2, pp. 222–232.

N. U. NAIR, P. G. SANKARAN, N. BALAKRISHNAN (2013a). *Quantile-based Reliability Analysis*. Springer, New York.

N. U. NAIR, P. G. SANKARAN, S. M. SUNOJ (2013b). *Quantile based stop-loss transform and its applications*. Statistical Methods & Applications, 22, no. 2, pp. 167–182.

S. NATARAJA (2009). *Models of over reported incomes*. Applied Economic Letters, 16, pp. 699–703.

J. S. RAMBERG, B. W. SCHMEISER (1974). *An approximate method for generating asymmetric random variables*. Communications of the ACM, 17, no. 2, pp. 78–82.

P. G. SANKARAN, N. N. MIDHU (2016). *Nonparametric estimation of mean residual quantile function under right censoring*. Journal of Applied Statistics, 44, no. 10, pp. 1856–1874.

J. M. SARABIA (1997). *A hierarchy of Lorenz curves based on the generalized Tukey's lambda distribution*. Econometric Reviews, 16, no. 3, pp. 305–320.

J. M. SARABIA (2008). *A general definition of the Leimkuhler curve*. Journal of Informetrics, 2, no. 2, pp. 156–163.

J. M. SARABIA, V. JORDÁ, C. TRUEBA (2017). *The Lamé class of Lorenz curves*. Communications in Statistics, Theory and Methods, 46, no. 11, pp. 5311–5326.

A. SEN (1976). *Poverty: an ordinal approach to measurement*. Econometrica: Journal of the Econometric Society, pp. 219–231.

A. SEN (1986). *The measurement of poverty: An axiomatic approach*. Econometrica, 44, pp. 219–232.

P. K. SEN (1988). *The harmonic Gini coefficient and affluence indexes*. Mathematical Social Sciences, 16, no. 1, pp. 65–76.

J. SLEMROD (2007). *Cheating ourselves: The economics of tax evasion*. Journal of Economic Perspectives, 21, no. 1, pp. 25–48.

N. F. C. SULAIMAN, N. A. SANUSI, S. MUHAMAD (2020). *Survey dataset of Malaysian perception on rising cost of living*. Data in Brief, 28, p. 104910.

S. TALWALKER (1980). *On the property of dullness of Pareto distribution*. Metrika, 27, no. 1, pp. 115–119.

A. TARSITANO (2004). *Fitting the generalized lambda distribution to income data*. In *COMPSTAT'2004 Symposium*. Physica-Verlag/Springer, pp. 1861–1867.

A. VARKEY, H. N. HARIDAS (2023). *A review on quantile functions, income distributions, and income inequality measures*. Reliability: Theory & Applications, 18, no. 2 (73), pp. 50–62.

M. ZENGA (2007). *Inequality curve and inequality index based on the ratios between lower and upper arithmetic means*. Statistica & Applicazioni, 5, no. 1, pp. 3–27.