

ON TRUNCATED MEASURES OF INCOME INEQUALITY FROM A FUZZY PERSPECTIVE

R. POURMOUSA

ABSTRACT. In most statistical analysis, inequality or extent of variation in income is represented in terms of certain summary measures. But some authors argued that the concept of inequality is vague and thus cannot be measured as an exact concept. Therefore, fuzzy set theory provides naturally a useful tool for such circumstances. In this paper we have introduced a real-valued fuzzy method of illustrating the measures of income inequality in truncated random variables based on the case where the conditional events are vague. To guarantee certain relevant properties of these measures, we first selected three main families of measures and obtained their closed formulas, then used two simulated and real data set to illustrate the usefulness of derived results.

1. Introduction

The problem of modeling income data as well as measurement of inequality on the income of members of a group or a society has about 200-year history and has been attractive to a lot of researchers in Economics, Statistics, Sociology, etc. As statistical analysis point of view, there are various indices to indicate these measures. Many of these measures have been used to summarize inequality in terms of a single number. Although there had been many attempts to provide measures of income inequality in the nineteenth century, the first major development in this area can be attributed to the work of Lorenz in 1905 [15]. In the study of inequality of the income distribution, in which the Lorenz curve has an outstanding role, one of the most relevant measures is the Gini index proposed by Gini in 1912, directly related to the Lorenz curve [8]. Subsequently, different measures of income inequality such as coefficient of variation, relative mean deviation, mean deviation, standard deviation of logarithms of incomes and some entropy indices has been suggested in the next years. For a detailed study on various measures of income inequality we refer to Arnold [1], Kakwani [10] and Yitzhaki [21]. Recently the measures of income inequality have also been used to measure inequality in health and in life expectancy among different groups, which are extensively used in survival and reliability analysis; see Bonetti et al. [4], Hanada [9] and Shkolnikov et al. [19]. In many practical situations, the complete data may not be observable due to various reasons. What is available is truncated data. The truncation of a distribution can be defined as a process which results in certain values being cut-off, thereby results

Received: March 2016; Revised: December 2016; Accepted: June 2107

Key words and phrases: Measures of income inequality, Gini index, Fuzzy event, Membership function, Truncated distribution.

in a truncated distribution. Not only we have left- or right-truncated distributions but also we have distributions that are truncated in the middle, such as the distribution of monthly income of laborers in a country where we can observe values of the random variable between zero and a . In these cases the measures of income inequality are computed only for incomes which are greater or smaller than a fixed value, or between two values, respectively. Hence the importance of studying inequality measures of truncated distributions is much of interest. For example see Bhattacharya [3], Moothathu [16] and Ord et al. [17]. Also we know the truncation of a population in eliminating high (richest population) or low (poorest population) values can attend by imprecision and uncertainty in real-world applications. Therefore, the fuzzy set theory provides naturally an appropriate tool for modeling the imprecise concepts. A number of authors have evoked the concepts of fuzzy sets in the analysis of inequality, for instance Chiappero-Martinetti (1994, 2000), Vero & Werquin (1997), Deutsch & Silber (2005) and Qizilbash (2006). Our motivation for writing this paper was evaluation of the effect of uncertainty of truncation on measures of inequality with fuzzy set. For this, we have introduced a real-valued fuzzy method to illustrating the measures of income inequality in truncated random variables with fuzzy threshold. The structure of this paper is as follows; after introductory section, which focuses on a brief review of the basic concepts on the probability of fuzzy events and the truncated random variables, in Section 3, we introduce a truncated random variable with fuzzy threshold. Section 4 produces the various measures used for summarizing income inequality in available articles. Section 5 defines certain measures of income inequality for the fuzzy truncated distributions. Characterization results according to some specific models such as exponential and Pareto based on the functional form of these measures are also discussed. Finally Section 6 presents an application of the proposed methodology and comparisons between other alternative methods.

2. Preliminaries

This section reviews commonly-used expressions of the measures of income inequality as well as the concepts of the probability measure of fuzzy events and the truncated random variables that will be used in this paper.

2.1. Probability of Fuzzy Events. Probability theory and fuzzy logic are the principal components of an array of methodologies for dealing with problems in which uncertainty and imprecision play important roles [18]. In this subsection we have collected together the basic ideas about fuzzy sets and probability of fuzzy events which are necessary in this paper.

The concept of fuzzy set was initiated by Zadeh in 1965 [22]. Let (Ω, A, P) be a probability space and $X \in \mathbb{R}$ be a random variable in Ω . A fuzzy event is a fuzzy set \tilde{A} in Ω whose membership function, $m_{\tilde{A}} : \Omega \rightarrow [0, 1]$, is Borel measurable function and the probability of the fuzzy event \tilde{A} is defined by

$$\Pr(X \in \tilde{A}) = \int_{x \in \Omega} m_{\tilde{A}}(x) dP(x). \quad (1)$$

So, the probability of a fuzzy event is the expectation of its membership function, [22].

Many of the basic notions in probability theory, such as those of the mean, variance, entropy, etc., are defined as functionals of probability distributions. The concept of a fuzzy event suggests that it may be of interest to define these notions in a more general way which relates them to both a fuzzy event and a probability measure. For example, the mean of a fuzzy event \tilde{A} relative to a probability measure P may be defined as follows:

$$E_P(\tilde{A}) = \frac{1}{\Pr(X \in \tilde{A})} \int_{x \in \Omega} x m_{\tilde{A}}(x) dP(x),$$

where $m_{\tilde{A}}$ is the membership function of \tilde{A} and $\Pr(X \in \tilde{A})$ serves as a normalizing factor, [23].

Let \tilde{A} be a fuzzy event with membership function of $m_{\tilde{A}}$ and X be an absolutely continuous random variable with probability density function (pdf) $f(\cdot; \theta)$, where θ represents the vector of parameters. The above equation can be rewritten as

$$E_P(\tilde{A}) = \int_x x \frac{m_{\tilde{A}}(x)f(x; \theta)}{\Pr(X \in \tilde{A})} dx.$$

Using the definition of expectation of a random variable and comparing it with above equation, we can show the function $f_{\tilde{A}}$, defined by

$$f_{\tilde{A}}(x; \theta) = \frac{m_{\tilde{A}}(x)f(x; \theta)}{\Pr(X \in \tilde{A})}, \quad x \in \Omega. \tag{2}$$

is a formula for the distribution X on Ω obtained by updating the distribution P on the basis of fuzzy information. Indeed, equation (2) can be viewed as a version of Bayes theorem provided we interpret the membership function $m_{\tilde{A}}(x)$ as the likelihood of vague concept θ given value x [14].

We know that a truncated distribution is a conditional distribution that results from restricting the domain of a random variable. So equation (2) can be corresponded to conditional distribution of X given that it is contained in the extension of θ [14]. So the defined function $f_{\tilde{A}}$ can be viewed as a version of truncated distribution obtained from the fuzzy constraint for a random variable that more details will be given in next sections. For this purpose the next subsection studies truncated distribution and its properties.

Example 2.1. [14], Let $\Omega = [0, 1]$ and the distribution of X be uniform. If \tilde{A} be the fuzzy set of approximately 0.5 with the following membership function

$$m_{\tilde{A}}(x) = \begin{cases} 2x, & x \leq 0.5 \\ 2(x - 1), & x > 0.5 \end{cases}$$

according to equation (1) and equation (2), we have

$$f_{\tilde{A}}(x) = \begin{cases} 4x, & x \leq 0.5 \\ 4(x - 1), & x > 0.5 \end{cases} .$$

2.2. Truncated Random Variables. Truncated distributions arise in practical statistics in cases where the ability to record, or even to know about, occurrences is limited to values which lie above or below a given threshold or within a specified range such as income, expenditure, blood pressure, humidity, etc. So truncated distributions are widely applied in different fields such as finance, physics, hydrology, geology, astronomy and biomedical science [6]. For more details suppose policyholders are subject to a policy limit, u , then any losses that are actually above u , are reported to the insurance company as being exactly u because u is the maximum amount that insurance companies pay. The insurance company knows that the actual loss is greater than u but they don't know what it is. On the other hand, left truncation occurs when policyholders are subject to a deductible. If policyholders are subject to a deductible d , any loss amount that is less than d will not even be reported to the insurance company. If there is a claim on a policy limit of u and a deductible of d , any loss amount that is greater than u will be reported to the insurance company as a loss of $u-d$ because that is the amount the insurance company has to pay. Therefore insurance loss data is left-truncated because the insurance company doesn't know if there are values below the deductible d because policyholders won't make a claim. The insurance loss is also right censored if the loss is greater than u because u is the most the insurance company will pay, so it only knows that your claim is greater than u , not what the claim amount is exactly. In these situations, an underlying random variable X is observed only when X belongs to a particular set A [6].

A truncated distribution is a conditional distribution that results from restricting the domain of a probability distribution. For this, suppose X be a random variable and A be a fixed subset of values of the random variable X . The truncated random variable X_A with truncation points in A can be represented as

$$X_A \stackrel{d}{=} X | (X \in A),$$

with the pdf

$$f_A(x, \theta) = \frac{f(x; \theta)}{\Pr(X \in A)}, \quad \forall x \in A, \quad (3)$$

where $\stackrel{d}{=}$ means equal in distribution. In the particular case where X be an absolutely continuous random variable and $A = [c, +\infty)$, the truncated random variable X_A has a left-truncated distribution with the pdf

$$f_A(x; \theta) = \frac{f(x; \theta)}{1 - F(c; \theta)}, \quad (4)$$

for $x \geq c$ where $F(\cdot; \theta)$ the cdf of X . In this case the expectation X_A reduces to the vitality function introduced by Kupka and Loo [13], defined as

$$\mu_c = E(X | X \geq c) = \frac{1}{1 - F(c; \theta)} \int_c^{+\infty} xf(x; \theta) dx. \quad (5)$$

Using equality $E(X | X \leq c) = -E(X^* | X^* \geq c^*)$, where $X^* \stackrel{d}{=} -X$ and $c^* = -c$, we can relate the results for left truncated variables with the results for right truncated variables and vice versa. So

$$v_c = E(X | X \leq c) = \frac{1}{F(c; \theta)} \int_{-\infty}^c xf(x; \theta) dx. \quad (6)$$

In the following, we consider some examples of truncated distributions:

Example 2.2. (Left-truncated exponential distribution). A random variable X_A is said to have a left-truncated exponential distribution with parameters $\lambda, c \in \mathbb{R}^+$; denoted by $X_A \sim EXP_{LT}(\lambda, c)$, if $X_A \stackrel{d}{=} X \mid (X \in A)$, where X is the exponential random variable with parameter λ and $A = [c, +\infty)$. The corresponding pdf and expectation of X_A are, respectively

$$f_A(x; \lambda, c) = \lambda e^{-\lambda(x-c)}, \quad x > c, \tag{7}$$

and

$$\mu_c = c + \frac{1}{\lambda}. \tag{8}$$

Also, we can show simply that

$$E(X_A^n) = n! \sum_{j=0}^n \frac{c^j}{j! \lambda^{n-j}}.$$

Example 2.3. (Pareto distribution). The random variable X_A is said to have a Pareto distribution with parameters $\lambda, c \in \mathbb{R}^+$, denoted by $X_A \sim Pareto(\lambda, c)$, if its pdf be

$$f_A(x; \lambda, c) = \frac{\lambda c^\lambda}{x^{\lambda+1}}, \quad x > c.$$

The corresponding expectation of X_A is

$$\mu_c = E(X_A) = \frac{\lambda}{\lambda - 1} c, \quad \lambda > 1.$$

It is clear that $X_A \stackrel{d}{=} X \mid (X \in A)$, $X \sim Pareto(\lambda, 1)$ and $A = [c, +\infty)$.

It should also be noted that the Pareto distribution is closely connected to the truncated exponential distribution which this connection is: if $X_A \sim Pareto(\lambda, c)$, then $Y = \ln(X_A) \sim EXP_{LT}(\frac{1}{\lambda}, \ln(c))$. Thus the obtained results in the following examples will be readily extended for the Pareto distribution.

3. Truncating of Random Variables by Fuzzy Thresholds

From the discussion on previous section and to compare equation (1) and equation (2), we see that the defined function $f_{\tilde{A}}$ can be viewed as a version of truncated distribution obtained from the fuzzy constraint. For this, let X be a random variable and \tilde{A} is a fuzzy event defined over Ω . We introduce the random variable $X_{\tilde{A}}$ and from now call it "fuzzy truncated random variable" if it is truncated by fuzzy event \tilde{A} as follows

$$X_{\tilde{A}} \stackrel{d}{=} X \mid (X \in \tilde{A}), \tag{9}$$

where \tilde{A} is a fixed fuzzy subset of values of the random variable X with the membership function $m_{\tilde{A}}$. Note, the pdf of $X_{\tilde{A}}$ can be easily derived from equation (2). In the special case consider $\tilde{A} = \{x \in \mathbb{R}^+ \mid x \succsim c\}$, where the symbol \succsim denotes the fuzzified version of \geq with the linguistic interpretation "approximately greater than or equal to". So $X_{\tilde{A}}$ is a fuzzy left truncated random variable with the expectation $\tilde{\mu}_c$. In the following two examples, we obtain the closed form of the pdf and the expectation of the fuzzy left truncated exponential random variable.

Example 3.1. Let $X \sim EXP_{LT}(\lambda, c)$ and $\tilde{A} = \{x \in \mathbb{R}^+ \mid x \succsim c\}$ be a fuzzy event with the membership function

$$m_{\tilde{A}}(x) = \begin{cases} 0 & x < c \\ 1 - e^{-d(x-c)} & x \geq c \end{cases}, \quad c, d > 0. \quad (10)$$

Based on this membership function and equation (2), the fuzzy random variable $X_{\tilde{A}}$ is said to have a fuzzy left truncated exponential distribution with parameters $\lambda, c, d > 0$, denoted by $FEXP_{LT}(\lambda, c, d)$ if its pdf is given by

$$f_{\tilde{A}}(x; \lambda, c, d) = \frac{\lambda(\lambda+d)}{d} e^{-\lambda(x-c)} (1 - e^{-d(x-c)}), \quad x \geq c. \quad (11)$$

We can easily show that the cdf of $X_{\tilde{A}}$ and its corresponding expectation are

$$F_{\tilde{A}}(x; \lambda, c, d) = \begin{cases} 0 & x < c \\ 1 - \frac{\lambda+d}{d} e^{-\lambda(x-c)} + \frac{\lambda}{d} e^{-(\lambda+d)(x-c)} & x \geq c, \end{cases} \quad (12)$$

$$\tilde{\mu}_c = c + \frac{2\lambda + d}{\lambda(\lambda + d)}. \quad (13)$$

Since the concept of fuzzy truncated random variable introduced in this paper, is an extension of the classical random variables, it can be reduced into a classical random variable in cases where either truncation is not vague. Because, the limit of $m_{\tilde{A}}(x)$ and $\tilde{\mu}_c$ as d approaches infinity are indicator function and μ_c , respectively.

Similarly, based on the membership function

$$m_{\tilde{A}}(x) = \begin{cases} 0 & x < c - a \\ \frac{x+a-c}{a+b} & c - a \leq x < c + b \\ 1 & x \geq c + b \end{cases}, \quad (14)$$

the pdf of the fuzzy left truncated exponential distribution with parameters $\lambda, a, b, c > 0$, denoted by $FEXP_{LT}(\lambda, a, b, c)$, are obtained as follows

$$f_{\tilde{A}}(x; \lambda, a, b, c) = \frac{\lambda^2(a+b)}{e^{-(c-a)} - e^{-(c+b)}} \begin{cases} 0 & x < c - a \\ \frac{x+a-c}{a+b} e^{-\lambda x} & c - a \leq x < c + b \\ e^{-\lambda x} & x \geq c + b \end{cases}. \quad (15)$$

Therefore, it is easy to show that

$$\tilde{\mu}_c = c + \frac{(2 - a\lambda)e^a - (2 + b\lambda)e^{-b}}{\lambda(e^a - e^{-b})}. \quad (16)$$

4. Measuring Inequality

There are various measures of income inequality such as relative mean deviation, mean variation of logarithms of incomes and some entropy indices. The choice of measure of inequality and related distribution has been discussed elsewhere and we do not consider this issue further in the present paper. Elteto et al. [7] proposed three measures which have direct economic interpretations and are given by

$$U_c = \frac{\mu}{v_c}, \quad V_c = \frac{\mu}{\mu_c}, \quad W_c = \frac{\mu_c}{v_c},$$

where μ_c and v_c defined in equation (5) & equation (6) and μ is the expected value of the un-truncated random variable X .

The measure V_c may be regarded as a measure of inequality for the entire income distribution, while U_c and W_c indicate the inequalities of the two respective parts of the distribution below and above the mean.

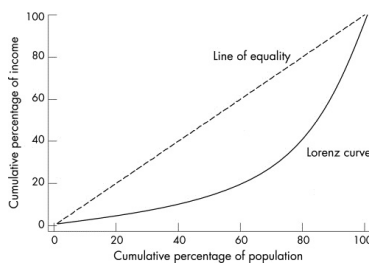


FIGURE 1. The Lorenz Curve Framework (Hypothetical Data)

The most widely used measure of inequality is the Lorenz curve, defined for finite populations. For this, let X be an absolutely continuous nonnegative random variable with cdf F , pdf f and finite mean μ , which describes income. Now, if $F^{-1}(p) = \sup\{x | F(x) \leq p\}$ is the quantile function of F , then the Lorenz curve of F is defined to be

$$L_F(p) = \frac{1}{\mu} \int_0^p F^{-1}(t) dt, \quad (17)$$

for $p \in [0, 1]$. $L_F(p)$ gives the fraction of total income that the holders of the lowest p^{th} fraction of income possesses. The Lorenz curve shows the percentage of total income earned by cumulative percentage of the population [15]. Most of the measures of income inequality are derived from the Lorenz curve framework illustrated in Figure 1.

The significance of the Lorenz curve lies in the Gini index defined as

$$G = 1 - 2 \int_0^1 L_F(p) dp. \quad (18)$$

The Gini index is a measure of statistical dispersion intended to represent the income or wealth distribution of a population, and is the most commonly used measure of inequality [20]. The Gini index is equivalent to the size of the area between the Lorenz curve and the 45° line of equality divided by the total area under the 45° line of equality. In Figure 1, if the area between the line of equality and Lorenz curve is A , and the area under the Lorenz curve is B , then the Gini index is $A/(A+B)$. This ratio is expressed as a percentage or as the numerical equivalent of that percentage, which is always a number between 0 and 1. A value of 0 indicates a perfectly equal society where all income is equally shared, while a value of 1 represents a perfectly unequal society where all income is earned by one individual, [1], [10] and [21].

In many practical situations, the complete data may not be observable due to various reasons. What is available is truncated data. In this situation, we will obtain measures of income inequality for truncated data. The truncated Gini index would be useful in these cases. The properties and applications of the income inequality measures for truncated data have been discussed by Ord et al. [17].

Proposition 4.1. *Let X_A be a non-negative and absolutely continuous random variable with truncation points in $A = [c, +\infty)$. An alternative representation of*

the introduced Gini index into equation (18) for X_A is equivalent to

$$G = \frac{2}{\mu_c} \int_c^{+\infty} x f(x, c) F(x, c) dx - 1, \quad (19)$$

where μ_c has defined in equation (5).

Proof. Using equation (17) and equation (18), we can write

$$G = 1 - 2 \int_0^1 L_F(p) dp = 1 - \frac{2}{\mu_c} \int_0^1 \int_0^p F^{-1}(t, c) dt dp.$$

It arises by replacing the integrals in the following expression

$$G = 1 - \frac{2}{\mu_c} \int_0^1 \int_t^1 F^{-1}(t, c) dp dt.$$

Since random variable $X \geq c$ then $F(x, c) = 0$ for $x < c$, using the substitution $y = F^{-1}(t, c)$ we have

$$G = 1 - \frac{1}{\mu_c} \int_c^{+\infty} 2y f(y, c)(1 - F(y, c)) dy,$$

and at last

$$G = \frac{2}{\mu_c} \int_c^{+\infty} x f(x, c) F(x, c) dx - 1. \quad \square$$

Applications of the income inequality measures in the context of income analysis are well known, see [11]. For this, two other important income inequality measures for the left truncated random variable X_A with $A = [c, +\infty)$ are

i) Truncated γ -entropy measure,

$$E_\gamma = \frac{1}{\gamma} \int_c^{+\infty} f(x, c) \{1 - f^\gamma(x, c)\} dx, \quad \gamma > -1. \quad (20)$$

By using the definition of the derivative and the Hopital's rule, we can show that

$$E_0 = - \int_c^{+\infty} \ln(f(x, c)) f(x, c) dx. \quad (21)$$

ii) As can be seen in Taillie (1981) corresponding to any convex function Ψ we may define a measure of inequality by $H_\Psi(X) = E(\Psi(X))$ and these measures preserve the Lorenz ordering. In particular corresponding to the family of convex functions

$$\Psi(X) = \frac{x^{\gamma+1} - 1}{\gamma(\gamma + 1)} \quad -\infty < \gamma < +\infty;$$

we obtain the family of truncated measures derived from the Mellin transform as

$$H_\gamma = \frac{1}{\gamma(\gamma + 1)} \int_c^{+\infty} \left(\left(\frac{x}{\mu_c} \right)^{\gamma+1} - 1 \right) f(x, c) dx, \quad (22)$$

with

$$H_0 = \int_c^{+\infty} \left(\frac{x}{\mu_c} \right) \ln \left(\frac{x}{\mu_c} \right) f(x, c) dx. \quad (23)$$

These family includes some standard indices, as special cases we have that H_1 is the Hirschman index, H_0 is the Theil index [2].

Example 4.2. According to Proposition 4.1, equation (20) and equation (22), we can readily obtain the income inequality measures based on the left truncated exponential random variable $X_A \sim EXP_{LT}(\lambda, c)$ with $\lambda, c \in \mathbb{R}^+$ as:

$$G = \frac{1}{2\lambda\mu_c} = \frac{1}{2(\lambda c + 1)}, \quad (24)$$

$$E_\gamma = \frac{1}{\gamma} \left(1 - \frac{\lambda^\gamma}{\gamma + 1} \right), \quad \gamma > -1, \quad (25)$$

and

$$H_\gamma = \frac{1}{\gamma(\gamma + 1)} \left(\frac{e^{\lambda c} \Gamma(\gamma + 2, \lambda c)}{(\lambda c + 1)^{\gamma+1}} - 1 \right), \quad (26)$$

where Γ is the upper incomplete gamma function as

$$\Gamma(\alpha, c) = \int_c^{+\infty} x^{\alpha-1} e^{-x} dx = (\alpha - 1)! e^{-c} \sum_{j=0}^{\alpha-1} \frac{c^j}{j!}.$$

Using the definition of Γ and equation (26), we can get H_γ for $\gamma = 1, 2$ as

$$H_1 = \frac{1}{2(\lambda c + 1)^2}, \quad H_2 = \frac{5 + 3\lambda c}{6(\lambda c + 1)^3}.$$

Example 4.3. If $X_A \sim Pareto(\lambda, c)$, then with similar analyzes as in Example 4.2, we can determine that the measures G , E_γ and H_γ for the Pareto distribution, as

$$G = \frac{1}{2\lambda - 1}, \quad \lambda > \frac{1}{2}, \quad (27)$$

$$E_\gamma = \frac{1}{\gamma} \left(1 - \frac{\lambda^{\gamma+1}}{(\gamma + 1)\lambda + \gamma} c^{-\gamma} \right), \quad \gamma > -1, \quad (28)$$

and

$$H_\gamma = \frac{1}{\gamma(\gamma + 1)(\lambda - \gamma - 1)} c^{-\lambda + \gamma + 1}, \quad \gamma \in (0, \lambda - 1). \quad (29)$$

5. Fuzzy Measures of Income Inequality

In this section, we extend the income inequality measures recalled in equation (19), equation (20) and equation (22) by the fuzzy truncated distributions and obtain the explicit expressions for them which call the fuzzy measures of income inequality. In the previous section we discussed various forms of measures that could represent data on incomes. An alternative way of describing a measure of income inequality is the using of fuzzy set theory. The fuzzy set theory provides naturally an appropriate tool for modeling the imprecise concepts.

The three measures given in the previous section can be translated to fuzzy forms \tilde{G} , \tilde{E}_γ and \tilde{H}_γ if the random variable X_A will be replaced by the fuzzy truncated random variable $X_{\tilde{A}}$ where $\tilde{A} = \{X \in \mathbb{R}^+ \mid X \succsim c\}$ is a fuzzy event.

In this paper, the measures \tilde{G} , \tilde{E}_γ and \tilde{H}_γ called the fuzzy measures of income inequality. This section will be finished with a series of examples which obtain the measures of income inequality for the exponential distribution via fuzzy events. The proofs follow directly from the equation (2) and equation (11).

Example 5.1. Consider a fuzzy event $\tilde{A} = \{X \in \mathbb{R}^+ \mid X \succsim c\}$ with the membership function given in equation (10), where $X \sim EXP_{LT}(\lambda, c)$. Using equation (11), we simply obtain the measures of income inequality as follows:

(i) The fuzzy Gini index is defined as

$$\tilde{G} = \frac{2}{\tilde{\mu}_c} \int_c^{+\infty} x f_{\tilde{A}}(x; \lambda, c, d) F_{\tilde{A}}(x; \lambda, c, d) dx - 1,$$

where $f_{\tilde{A}}$, $F_{\tilde{A}}$ and $\tilde{\mu}_c$ have been given in equation (11), equation (12) and equation (13), respectively. It can be easily calculated as follows,

$$\tilde{G} = \frac{A}{\tilde{\mu}_c} - 1, \quad (30)$$

where

$$A = c + \frac{11\lambda^2 + 11\lambda d + 3d^2}{2\lambda(\lambda + d)(2\lambda + d)},$$

(ii) The fuzzy γ -entropy measure is independent of c and has the general form

$$\tilde{E}_\gamma = \frac{1}{\gamma} \int_c^{+\infty} f_{\tilde{A}}(x; \lambda, c, d) \left(1 - f_{\tilde{A}}^\gamma(x; \lambda, c, d)\right) dx, \quad \gamma > -1.$$

With solving this integral, for $\gamma > 0$, we obtain

$$\tilde{E}_\gamma = \frac{1}{\gamma} \left(1 + \frac{\lambda^{\gamma+1} (\lambda + d)^{\gamma+1}}{d^\gamma} \sum_{j=0}^{\gamma} \frac{(-1)^{j+1} \binom{\gamma}{j}}{((\gamma + 1)\lambda + jd)((\gamma + 1)\lambda + (j + 1)d)} \right),$$

Also, by mathematical induction on γ , we can show that

$$\tilde{E}_\gamma = \frac{1}{\gamma} - \frac{k_\gamma}{\gamma + 1} \frac{\lambda^\gamma (\lambda + d)^\gamma}{\prod_{n=1}^{\gamma} ((\gamma + 1)\lambda + nd)}, \quad \gamma > 0, \quad (31)$$

where k_γ is a constant and determined based on γ value. Table 1 shows some values k_γ for some constant γ .

In two special cases, when $\gamma = 1, 2$ we have

$$\tilde{E}_1 = 1 - \frac{1}{2} \frac{\lambda(\lambda + d)}{2\lambda + d},$$

and

$$\tilde{E}_2 = \frac{1}{2} - \frac{1}{3} \frac{\lambda^2 (\lambda + d)^2}{(3\lambda + d)(3\lambda + 2d)}.$$

(iii) The fuzzy measures derived from the Mellin transform have the general form

$$\tilde{H}_\gamma = \frac{1}{\gamma(\gamma + 1)} \int_c^{+\infty} \left(\left(\frac{x}{\tilde{\mu}_c} \right)^{\gamma+1} - 1 \right) f_{\tilde{A}}(x; \lambda, c, d) dx.$$

In a special case, when $\gamma > 0$, we have

$$\tilde{H}_\gamma = \frac{1}{\gamma(\gamma+1)} \left[\frac{\lambda(\lambda+d)}{d\tilde{\mu}_c^{\gamma+1}} \sum_{j=0}^{\gamma+1} \frac{c^j(\gamma+1)!}{j!} \left(\frac{1}{\lambda^{\gamma+2-j}} - \frac{1}{(\lambda+d)^{\gamma+2-j}} \right) - 1 \right],$$

It can simply be shown that \tilde{H}_γ has the general form

$$\tilde{H}_\gamma = \frac{1}{\tilde{\mu}_c} P_{\gamma+1}(c), \quad \gamma \in \mathbb{N} \tag{32}$$

where $P_{\gamma+1}(c)$ is an algebraic polynomial of degrees $\gamma+1$. This form for $\gamma=1$ is

$$\tilde{H}_1 = \frac{\lambda(\lambda+d)}{2d\tilde{\mu}_c^2} \left[\left(\frac{1}{\lambda} - \frac{1}{\lambda+d} \right) c^2 + 2 \left(\frac{1}{\lambda^2} - \frac{1}{(\lambda+d)^2} \right) c + 2 \left(\frac{1}{\lambda^3} - \frac{1}{(\lambda+d)^3} \right) \right] - \frac{1}{2}.$$

γ	1	2	3	4	5	6
k_γ	1	1	$\frac{1}{2}$	6	$\frac{1}{12}$	120

TABLE 1. Some Values of k_γ in Equation (31) to Determine \tilde{E}_γ

Example 5.2. Similarly by the membership function given in equation (14) and using equation (15), we can show that

(i) The corresponding fuzzy Gini index is given by

$$\tilde{G} = \frac{A_2 + B_2c}{\tilde{\mu}_c}, \tag{33}$$

where

$$A_2 = \frac{(11 - 4(a + \tilde{\mu}_c)\lambda)e^{2\lambda(a+b)} - 8((b - \tilde{\mu}_c)\lambda + 2)e^{\lambda(a+b)} - 2(a - b + 2\tilde{\mu}_c)\lambda + 5}{4\lambda(2e^{\lambda(a+b)} - e^{2\lambda(a+b)} - 1)},$$

and

$$B_2 = \frac{2e^{\lambda(a+b)} - e^{\lambda(a+2b)} - 1}{2e^{\lambda(a+b)} - e^{2\lambda(a+b)} - 1},$$

and $\tilde{\mu}_c$ has been given in equation (16).

(ii) The fuzzy γ -entropy measure has no closed form in this situation, therefore it must be determined from numerical methods.

(iii) The fuzzy measures derived from the Mellin transform is given by

$$\tilde{H}_\gamma = \frac{e^{-(c-a)}}{e^{-(c-a)} - e^{-(c+b)}} V_{\gamma+1}(c) + \frac{e^{-(c+b)}}{e^{-(c-a)} - e^{-(c+b)}} W_{\gamma+1}(c), \tag{34}$$

where $V_{\gamma+1}(c)$ and $W_{\gamma+1}(c)$ are two algebraic polynomials of degree $\gamma+1$. For example if $\gamma=1$, then we can obtain

$$V_2(c) = \frac{(6 - 4a\lambda + (a^2 - \tilde{\mu}_c^2)\lambda^2) + 2\lambda(2 - a\lambda)c + \lambda^2c^2}{2\lambda^2\tilde{\mu}_c^2},$$

and

$$W_2(c) = \frac{-(6 + 4b\lambda + (b^2 - \tilde{\mu}_c^2)\lambda^2) + 2\lambda(2 + b\lambda)c - \lambda^2c^2}{2\lambda^2\tilde{\mu}_c^2}.$$

6. Numerical Illustration

This section first considers a simulation study to illustrate the application of the measures of income inequality and then uses real data to illustrate the usefulness of the results derived in the previous section.

6.1. Simulation Study. We first consider a simulation study to illustrate the application of the measures of income inequality in which we generate a random sample that considers the annual wages of 50 production-line workers in an industrial firm. For this purpose, we use $EXP_{LT}(0.012, 100)$ to analyze the distribution of the annual wages. As an extra fuzzy information we have that minimum wages have been defined approximately 100\$. In this situation, we use $FEXP_{LT}(0.012, 100, 2)$ and $FEXP_{LT}(0.012, 80, 80, 100)$ as the distributions of the annual wages. Some values of the simulated measures of income inequality which have been calculated for these data, displayed in Table 2.

Distribution of the annual wages	μ_c	G	E_1	H_1
$EXP_{LT}(0.012, 100)$	183.33	0.2272	0.994	0.1030
$FEXP_{LT}(0.012, 100, 2)$	183.83	0.2267	0.994	0.1027
$FEXP_{LT}(0.012, 84, 84, 100)$	182.67	0.3306	—	0.2080

TABLE 2. Income Inequality Measures for the Truncated and Fuzzy Truncated Exponential Distribution Fitted on the Annual Wages Data

6.2. Application of the Measures of Income Inequality in Survival Time.

The area of fitting distributions to data has seen explosive growth in recent years. Distribution fitting is the procedure of selecting a statistical distribution that best fit to a data set generated by some random process. Statistical analysis, in traditional form, is based on crispness of data, random variables, decision rules, parameters, and so on. However, imprecision and uncertainty information exist in real-world applications that can be caused by human errors in collecting data or some unexpected situations. Therefore, the fuzzy set theory provides naturally an appropriate tool in modeling the imprecise concepts. This can be a good reason for using the fuzzy random variable in fitting distributions to data. We use real data to illustrate the usefulness of the results derived in the previous section. The data used in this subsection are observations that represent the survival times of the 72 guinea pigs injected with different doses of tubercle bacilli, denoted by X . This data set has been considered by several authors in the literature. For example, Kundu et al. [12] considered an inverse Weibull (IW) distribution for these data set. Also, Cordeiro et al. [5], compared it with the half-normal (HN) and generalized half-normal models. In this situation, we want to estimate the measures of inequality of X by using fuzzy truncated random variables in these data given that the minimum survival times already have been defined approximately 9.26. For this, we consider $\tilde{A} = \{x \in \mathbb{R}^+ \mid x \lesssim 9.26\}$ with the membership function

$$m_{\tilde{A}}(x) = \begin{cases} 0 & x < 9.26 \\ 1 - e^{-d(x-9.26)} & x \geq 9.26 \end{cases}, \quad d > 0. \quad (35)$$

Using equation (11), it is clear $X \sim FEXP_{LT}(\lambda, 9.26, d)$ with the following pdf

$$f_{\tilde{A}}(x; \lambda, 9.26, d) = \frac{\lambda(\lambda+d)}{d} e^{-\lambda(x-9.26)} (1 - e^{-d(x-9.26)}), \quad x \geq 9.26.$$

We have fitted the truncated exponential $EXP_{LT}(\lambda, 9.26)$ and the fuzzy truncated exponential distribution $FEXP_{LT}(\lambda, 9.26, d)$ to this data set. They are also compared with the Weibull, IW and generalized inverse Weibull (GIW) distributions, while their corresponding pdfs are respectively given as follows:

$$f_1(x; \lambda, d) = \lambda d x^{-(d+1)} e^{-\lambda x^{-d}}, \quad x, \lambda, d > 0,$$

$$f_2(x; \lambda, d) = \lambda d x^{d-1} e^{-\lambda x^d}, \quad x, \lambda, d > 0,$$

$$f_3(x; \lambda, c, d) = c \lambda^d d x^{-(d+1)} e^{-c \lambda x^{-d}}, \quad x, \lambda, d, c > 0.$$

For each distribution, we derive the maximum likelihood estimates, the maximized log-likelihood (log-L) and the Akaike information criterion (AIC). The results for all the mentioned models are presented in Table 3. Notice that the proposed $FEXP_{LT}$

Models	λ	c	d	log-L	AIC
IW	283.84	1.415	—	-395.64	795.30
Weibull	0.0014	1.393	—	-397.15	798.30
GIW	61.008	1.415	0.146	-395.65	797.29
EXP_{LT}	0.0114	12	—	-394.22	792.44
$FEXP_{LT}$	0.0137	9.26	0.442	-391.146	788.29

TABLE 3. Parameter Estimates, Log-likelihood and AIC Obtained by Fitting Each of the Distributions on the Survival Times Data

model is fitted to the data set better than the other models. This conclusion is also supported by the values from the log-likelihood and AIC given in Table 3. In summary, the new $FEXP_{LT}$ distribution might be an interesting alternative to the other available models in the literature for modeling positive real data. Figure 2 depicts a histogram of the data and the fitted $FEXP_{LT}$ probability density function. We observed that the $FEXP_{LT}$ distribution fits the the survival times data better than EXP_{LT} distribution.

Table 4 indicates the measures of inequality estimated according to the two last distributions in Table 3 which by using them we can compare the effect of fuzzy truncation in these measures.

Distribution	μ_c	G	E_1	H_1
EXP_{LT}	84.4471	0.23379	0.99431	0.07287
$FEXP_{LT}$	99.7193	0.20002	0.99778	—

TABLE 4. Income Inequality Measures for the Truncated and Fuzzy Truncated Exponential Fitted on the Annual Wages Data

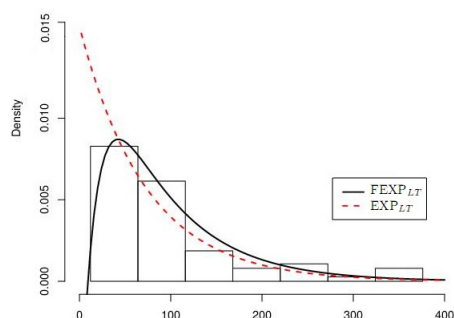


FIGURE 2. Fitted EXP_{LT} and $FEXP_{LT}$ Distributions to the Survival Times Data

7. Conclusions

In this paper we have introduced a real-valued fuzzy method of illustrating the measures of income inequality in truncated random variables based on the case where the conditional events are vague. It focused entirely on the three families of the income inequality indices and obtained closed formulas for these. Characterization results according to some specific models such as exponential and Pareto based on the functional form of these measures are also discussed. We have used two simulated and real data set to illustrate the usefulness of derived results. Notice that the proposed model is fitted to the data set better than the other models. The method proposed in this paper can be extended to higher dimensions.

REFERENCES

- [1] B. C. Arnold, *Majorization and the Lorenz order*, Lecture notes in statistics 43, Springer, Berlin and New York, 1987.
- [2] F. Belzunce, J. Candel and J. M. Ruiz, *Ordering of truncated distributions through concentration curves*, Sankhya A, **57** (1995), 375-383.
- [3] N. Bhattacharya, *A property of the Pareto distribution*, Sankhya B, **25** (1963), 195-196.
- [4] M. Bonetti, C. Gigliarano and P. Muliere, *The Gini concentration test for survival data*, Lifetime Data Analysis, **15** (2009), 493-518.
- [5] G. M. Cordeiro, *The Kumaraswamy generalized half-normal distribution for skewed positive data*, Journal of Data Science, **10** (2012), 195-224.
- [6] Y. Dodge, *The Oxford Dictionary of Statistical Terms*, OUP, 2003.
- [7] O. Elteto, E. Frigyes, *New income inequality measures as efficient tools for causal analysis and planning*, Econometrica, **36** (1968), 383-396.
- [8] C. Gini, *On the measure of concentration with special reference to income and statistics*, Colorado College Publication, General Series, **208** (1936), 73-79.
- [9] K. Hanada, *A Formula of Gini's concentration ratio and its applications to life tables*, Journal of Japanese Statistical Society, **19** (1983), 293-325.
- [10] N. Kakwani, *On a class of poverty measures*, Econometrica, **48** (1980), 437-446.
- [11] C. Kleiber and S. Kotz, *Statistical Size Distributions in Economics and Actuarial Sciences*, Wiley, Hoboken, 2003.

- [12] D. Kundu and H. Howlader, *Bayesian Inference and prediction of the inverse Weibull distribution for Type-II censored data*, Computational Statistics and Data Analysis, **54** (2010), 1547-1558.
- [13] J. Kupka and S. Loo, *The hazard and vitality measures of aging*, Journal of Applied Probability, **26** (1989), 532-542.
- [14] J. Lawry, *Modelling and Reasoning with Vague Concepts*, Springer, 2006.
- [15] M.O. Lorenz, *Methods of measuring the moncentration of wealth*, Journal of the american statistical association, **70(9)** (1905), 209-217.
- [16] T. S. K. Moothathu, *A characterization of power function distribution through a property of the Lorenz curve*. *Sankhya*, The Indian Journal of Statistics, Series B, **48** (1986), 262-265.
- [17] J. K. Ord, G. P. Patil and C. Taillie, *Truncated distributions and measures of income inequality*. *Sankhya*, The Indian Journal of Statistics, Series B, **45** (1983), 413-430.
- [18] T. J. Ross, J. M. Booker and W. J. Parkinson, *Fuzzy Logic and Probability Applications*, SIAM, Philadelphia, 2002.
- [19] V. M. Shkolnikov, E. E. Andreev and A. Z. Begun, *Gini coefficient as a life table function: Computation from discrete data, decomposition of differences and empirical examples*, Demographic Research, **8** (2003), 305-358.
- [20] L. M. Surhone, M. T. Timpledon and S. F. Marseken, *Truncated Distribution*, Betascript Publishing, 2010.
- [21] S. Yitzhaki, *On an extension of the Gini inequality index*, International economic review, **24** (1983), 617-628.
- [22] L. A. Zadeh, *Fuzzy Sets*, Information and Control, **8** (1965), 338-353.
- [23] L.A. Zadeh, *Probability measures of fuzzy events*, Journal of Mathematical Analysis and Applications, **23**(1968), 421-427

R. POURMOUSA, DEPARTMENT OF STATISTICS, FACULTY OF MATHEMATICS AND COMPUTER,
SHAHID BAHONAR UNIVERSITY OF KERMAN, KERMAN, IRAN

E-mail address: pourm@uk.ac.ir