Journal of Statistics & Management Systems
ISSN 0972-0510 (Print), ISSN 2169-0014 (Online)
Vol. 27 (2024), No. 5, pp. 985–995
DOI : 10.47974/JSMS-1274

Using fuzzy statistical models to estimate upper and lower bounds of Gini coefficient

Walaa Jabbar Al-Obaidi †
Ramin Imany-Nabiyyi *
Department of Statistics
Faculty of Mathematical Sciences
University of Tabriz
Tabriz
Iran

Abstract

In this article, several fuzzy flexible distributions are used to model fuzzy income data, and then we estimate the upper and lower bounds of the Gini coefficient. Estimates are obtained using the maximum likelihood method. We have conducted a simulation study to evaluate the performance of the proposed bounds. The simulation results showed that the introduced bounds have performed well and include the Gini coefficient. The data used in this article is related to Iranian household income data in both urban and rural areas, which was taken from the website of the Iran Statistics Center, https://www.amar.org.ir, for the years 2011,2014,2017,2020 and 2021.

Subject Classification: *03E72, 15B15.*

Keywords: *Gini coefficient, Fuzzy data, Income distribution, Dagum and Singh-Maddal distributions.*

1. Introduction

One of the social problems that humans are dealing with is income inequality and class differences in society. These problems are potential factors in the occurrence of abnormality in society. For this reason, politicians and economists pay attention to measuring income inequality and reducing its severity. Policy-making in this issue requires accurate and correct knowledge of the amount of income inequality in the population. Several indicators have been

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

† *E-mail:* **Wlajbar93@gmail.com**
* *E-mail:* **imany@tabrizu.ac.ir (Corresponding Author)**

defined in this field. One of the most common indicators of income inequality is the Gini coefficient, which was presented by Corridor Gini (1912). It is a number between zero and one. The closer the Gini coefficient is to zero, the more equality is observed in the income distribution, and conversely, the closer the Gini coefficient is to one, the greater the distribution of income inequality. If the coefficient is zero, it indicates that everyone in the population has the same income and wealth, which is complete equality. On the other hand, if the coefficient is one, it shows that complete inequality has taken place, where one person has all the wealth, and the rest have no income [1-5].

If X represents the household income of a population with the distribution function F(.), then the Gini coefficient is defined as follows:

$$G = 1 - 2\int_0^1 L(u)du\ ,\ u \in (0,1)\ . \tag{1}$$

where L(u) $\left(L(u) = \frac{1}{E(X)}\int_0^u F^{-1}(t)dt\right)$ is the Lorenz function.

Several calculation methods have been proposed to estimate the Gini coefficient, including the geometric method, differential average, covariance, matrix, etc. McDonald and Ransom (1981) used the gamma density, applied Monte Carlo methods, and presented an estimation of lower and upper bounds for Gini coefficient. Golden (2008) introduced a quick approximation of the Gini coefficient using numerical data in cumulative income quintiles. According to the authors' studies, all available sources are related to the calculation and estimation of the Gini coefficient based on point data. However, Ganjoei et al. (2021) introduced an estimate of upper and lower bounds for the Gini coefficient using fuzzy data. They proposed these estimates by considering the fuzzy regression model with fuzzy coefficients and several macroeconomic indicators as independent variables. In this paper, we present an estimate of the upper and lower bounds of the Gini coefficient using fuzzy statistical distributions.

2. Parametric Model of Income Distribution

The main advantage of the parametric method is that the information in the data can be concentrated in a small number of parameters. Moreover, useful information can be extracted directly from the estimated parameters, provided that the parameters have a clear economic interpretation. Therefore, one of the certain characteristics required for probabilistic models of income distribution is the economic interpretation of their parameters (Dagum, 1977). This advantage enables the use of parameter changes as reasonable indicators to explain changes in income distribution [6-9].

In the parametric approach, it is assumed that the income distribution can be modelled with one of the continuous distribution functions. That is, by a member of

$$\Psi = \{F_\theta(x)\,, x > 0\} \tag{2}$$

Where $\theta = \left(\theta_1, \dots, \theta_p\right)^T \subset R^p$the parameter is vector and $f_\theta(x)$ represents the density function. In recent decades, many continuous distributions have been introduced to fit income data. Mandelbrot (1960) stated that the Pareto distribution is a one-parameter model that provides a good fit for the lower tail of the income distribution, while the fit is poor over the entire income range. Among two-parameter distributions, the log-normal distribution is one of the suitable distributions for modelling income distribution. This distribution provides closed forms for most inequality indices that depend only on the shape parameter (Kleiber and Kotz 2003). Salem and Mount (1974) showed that the gamma distribution fits the data better than the log-normal model, especially in the tails, but both overestimate the skewness. Singh and Maddala (1976) presented the three-parameter Singh-Maddala distribution that performs better than the gamma and log normal. Another three-parameter model, Dagum (1977) introduced the Dagum model and later realized that this model performs better than the Singh-Maddala model (Dagum, 1980). McDonald (1984) introduced a flexible four-parameter distribution called the generalized beta of the second type (GB2) and showed that it has a better fit than either of them. GB2 includes many frequently used income distribution models such as the Dagum, Singh Maddala, and log-normal, gamma, and Fisk distributions. However, the Singh-Maddala and Dagum distributions are widely considered as models for income distribution in empirical analysis. These distributions have some good properties and several estimation methods for the parameters of these distributions from various income data have been extensively discussed. Previous studies argue that the Dogma distribution provides a better fit than the Singh-Maddala distribution in empirical analyses. For more details, see Kleiber (1996), Kleiber and Kotz (2003), Dagum (1980), McDonald and Mantrala (1995), McDonald and Zhu (1995), Pena et al. (1996) and Victoria-Feser (1995, 2000) [5-14].

The density function of the Dagum (Da) and Singh-Maddala (SM) distributions are, respectively:

$$Da(x;\lambda.\beta.\delta) = \frac{\lambda\beta\delta x^{-\delta-1}}{\left(1+\lambda x^{-\delta}\right)^{\beta+1}}.x > 0, \lambda.\beta, \delta > 0 \tag{3}$$

$$SM(x;\lambda.\beta.\delta) = \frac{\lambda^{-\delta}\beta\delta x^{\delta-1}}{\left[1+\left(\frac{x}{\lambda}\right)^{\delta}\right]^{\beta+1}}.x > 0, \lambda, \beta, \delta > 0 \tag{4}$$

Some mathematical properties of the Dagum distribution such as moments, Gini coefficient and Lorenz curve are stated in Dagum (1977). Dagum (1977) obtained the Gini coefficient based on this distribution as follows:

$$G_{Da} = \frac{\Gamma(\beta)\Gamma\left(2\beta+\frac{1}{\delta}\right)}{\Gamma(2\beta)\Gamma\left(\beta+\frac{1}{\delta}\right)} - 1 \tag{5}$$

He also showed that the sign of the partial derivative of the Gini index is negative compared to δ and β, but their effects on income distribution are different from each other. Kumar (2017) explained several mathematical and statistical properties of SM distribution such as quintiles, moments, Lorenz curves, Gini coefficient etc. Based on this distribution, the Gini coefficient is obtained as follows:

$$G_{SM} = 1 - \frac{\Gamma(\beta)\Gamma\left(2\,\beta-\frac{1}{\delta}\right)}{\Gamma(2\beta)\Gamma\left(\beta-\frac{1}{\delta}\right)} \tag{6}$$

In the following sections, based on these two distributions in the fuzzy state, we try to introduce estimates of the upper and lower bounds of the Gini coefficient. To compare these estimates with the estimates based on a two-parameter distribution, we suggest the log-normal (LN) distribution for two reasons [15-22]. First, the good fit of this distribution to the income data, and second, the presentation of the closed form of the Gini coefficient based on this distribution. The Gini coefficient based on LN distribution is as follows:

$$G_{LN} = 2\Phi\left(\frac{\beta}{\sqrt{2}}\right) - 1 \tag{7}$$

Where β is the scale parameter and $\Phi(.)$ is standard normal distribution function.

3. Elementary Fuzzy Arithmetic

In this section, we explain some basic principles. If $f\,(.)$ is a function from X to Y and A is a fuzzy set on X, the membership function of the fuzzy set image A under the mapping $f\,(.)$ can be represented as a fuzzy set B. On the other hand, the resolution identity theorem states that the membership function of a fuzzy set can be expressed by its α-cuts (Zimmerman 2001). These principles are explained below.

Suppose *Z* be the Cartesian product of $Z_1, \dots, Z_n$ and $A_1, \dots, A_n$ be *n* fuzzy sets in $Z_1, \dots, Z_n$, respectively. Let f is a function from Z to a , $y = f\,(z_1, \dots, z_n)$. Then, by extension principle, we can define a fuzzy set *B* in *Y* by (Zimmerman 2001);

$$B = \{(y, B(y)) | y = f\,(z_1, z_2, \dots, z_n), (z_1, z_2, \dots, z_n) \in Z\} \text{ where,} \forall y \in Y$$

$$B = \begin{cases} \sup_{(z_1,\dots,z_n)\in f^{-1}(y)} \min\{A_1(z_1), A_2(z_2), \dots, A_n\,(z_n)\} & f^{-1}(y) \neq \emptyset \\ 0, & o.w \end{cases} \tag{8}$$

Where f^{-1} is the inverse image of f, and in one dimension case?

$$B = \begin{cases} \sup_{z\in f^{-1}(y)} \min\{A(z)\} & f^{-1}(y) \neq 0 \\ 0, & o.w. \end{cases}$$

Triangular fuzzy numbers are widely used in expressing expert perception. They are used in many applied sciences to represent uncertainty (Al-Qudaimi et al. 2021). Suppose $A = (\alpha, m, \beta)$ be a triangular fuzzy number for which *m* is called the mode of *A*, and α and β are respectively called the left and right end points, and $\gamma \in R$. Then, based on the extension principle, the extended scalar multiplication is obtained as

$$\gamma A = \begin{cases} (\gamma\alpha, \gamma m, \gamma\beta), \gamma > 0, \\ I_{(0)}, \gamma = 0, \\ (\gamma\beta, \gamma m, \gamma\alpha), \gamma < 0. \end{cases}$$

Suppose $A = (\alpha, m, \beta)$ and $B = (\lambda, n, \delta)$ be two triangular fuzzy numbers. Then, using the extension principle, the extended summation, multiplication and division of A and B are given by $A + B = (\alpha, m, \beta) \oplus (\lambda, n, \delta) = (\alpha + \lambda, n + m, \beta + \delta), A \otimes B =$

$(\min(\alpha\lambda, \beta\delta), nm, \max(\alpha\lambda, \beta\delta))$ and $A \oslash B = (\min(\alpha/\lambda, \beta/\delta), n/m, \max(\alpha/\lambda, \beta/\delta))$,respectively (Zimmerman 2001).

Proposition 2.2: If $\tilde{A} = (a, b, c)$ is a triangular fuzzy number and $\tilde{A}[\alpha] = (a_\alpha, c_\alpha)$ is an α-cut of $\tilde{A}$ then for$0 < \alpha < 1$, based extension principle, we can write

$$\Gamma(\tilde{A}[\alpha]) = \int_0^\infty x^{\tilde{A}[\alpha]-1} e^{-x} dx = (\min[\Gamma(a_\alpha), \Gamma(c_\alpha)], \max[\Gamma(a_\alpha), \Gamma(c_\alpha)]),$$

Where $\Gamma(.)$ is gamma function? Note that if $a_\alpha \geq 1.5$ then we have $\Gamma(\tilde{A}[\alpha]) = (\Gamma(a_\alpha), \Gamma(c_\alpha))$.

4. Estimation of Upper and Lower Bounds of Gini Coefficient

In this section, we consider the fuzzy statistical model. To establish the fuzzy statistical model, suppose that the parameters of model (2) are fuzzy. Let, in model, $\tilde{\theta}_i = (\alpha_i, \theta_i, \beta_i); i \in \{1,2, \dots, p\}$ are triangular fuzzy numbers in which for$i \in \{1,2, \dots, p\}$, θ_i are called the mode of $\tilde{\theta}_i$ and α_i and β_i are called the left and right end points, respectively. In this section, we interest to estimate $\tilde{\theta} = (\tilde{\theta}_1, \dots, \tilde{\theta}_p)^T$ and use them to estimate the Gini coefficient.

If X has the distribution function$F_\theta(x)$, then according to equation (1) the Gini coefficient is obtained as a function ofθ $(g(\theta))$. Now, if we can estimate the parameters of the fuzzy model, by using the α-cuts of these estimates and the plug-in method, we can introduce an estimate of the upper and lower bounds of the Gini coefficient based onα.

Buckley (2006) states that $(\hat{\theta} \mp z[\alpha]\sigma_{\hat{\theta}})$ for $\alpha \in (0.01,1)$ is a good estimator for the α-cut of unknown triangular fuzzy parameter where $\hat{\theta}$ is maximum likelihood estimator (MLE) of θ and $z[\alpha]$ is the $(1 - \frac{\alpha}{2})$th quantile of the standard normal distribution. Therefore, we suggest the following estimator to estimate fuzzy model parameters:

$$\hat{\tilde{\theta}}[\alpha] = \left(\hat{\theta}_{mle} \mp z_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{nI(\theta)}} \right);$$ where $I(\theta)$ represents Fisher's information.

For an observation of a random sample$(x_1, x_2, \dots, x_n)$, the maximum likelihood estimator, $\hat{\theta}$, corresponds to the solution $\hat{\theta} = \arg\max_{\theta \in \Theta} \prod_{i=1}^n f(x_i; \theta)$, where $\theta = (\theta_1, \dots, \theta_p)^T$ is the unknown parameters vector and n is the sample size. Usually, the maximum of the logarithm is obtained. Thus, taking into account that the probability distribution function is differentiable, the maximum likelihood estimator corresponds to the solution of the following nonlinear system of equations:

$$\sum_{i=1}^n \frac{\partial}{\partial \theta} \log(f(x_i; \theta))|_{\theta = \hat{\theta}} = 0 \qquad (9)$$

Using equation (9), for Dagum model in equation (3), we have

$$\begin{cases} \frac{n}{\beta} - \sum_{i=1}^{n} \log\left(1 + \lambda x_i^{-\delta}\right) = 0, \\ \frac{n}{\lambda} - (\beta + 1) \sum_{i=1}^{n} \frac{x_i^{-\delta}}{1+\lambda x_i^{-\delta}} = 0, \\ \frac{n}{\delta} - \sum_{i=1}^{n} \log x_i + (\beta + 1)\lambda \sum_{i=1}^{n} \frac{x_i^{-\delta} \log x_i}{1+\lambda x_i^{-\delta}} = 0, \end{cases} \tag{10}$$

And for Singh-Maddala model in equation (4), we can write

$$\begin{cases} \frac{n}{\beta} - (\delta + 1) \sum_{i=1}^{n} \log\left(1 + \left(\frac{x_i}{\lambda}\right)^{\delta}\right) = 0 \\ -\frac{n\delta}{\lambda} + \delta(\beta + 1) \sum_{i=1}^{n} \frac{x_i^{\delta}}{\lambda^{\delta+1}\left[1+\left(\frac{x_i}{\lambda}\right)^{\delta}\right]} = 0, \\ \frac{n}{\delta} - n\log\lambda + \sum_{i=1}^{n} \log x_i + (\beta + 1)\lambda \sum_{i=1}^{n} \frac{\left(\frac{x_i}{\lambda}\right)^{\delta} \log\frac{x_i}{\lambda}}{1+\left(\frac{x_i}{\lambda}\right)^{\delta}} = 0. \end{cases} \tag{11}$$

We should use a numerical method to solve (10) and (11). For LN distribution, the MLEs of parameters are as follows:

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} \log x_i \,, \hat{\beta}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\log x_i - \hat{\lambda}\right)^2.$$

Remark 2.3: Suppose that $\hat{\tilde{\theta}}[\alpha]$ is estimator of$\tilde{\theta}[\alpha]$. Using $\hat{\tilde{\theta}}[\alpha]$ and plug-in method, an estimator of upper and lower bounds for Gini coefficient is given as follows($\alpha \in [0.01,1]$):

$$\hat{\tilde{G}}[\alpha] = (g_1[\alpha], g_2[\alpha]) \tag{12}$$

Where $g_1[\alpha] = \min\left\{g(\hat{\theta})\middle|\hat{\theta} \in \hat{\tilde{\theta}}[\alpha]\right\}$, $g_2[\alpha] = \max\left\{g(\hat{\theta})\middle|\hat{\theta} \in \hat{\tilde{\theta}}[\alpha]\right\}$.

5. Simulation Study

In this section, we perform a simulation study to evaluate the performance of the proposed estimates and also compare the estimates obtained from three distributions: Dagum, Singh Maddala and Log Normal distributions. The simulation process has been done for $n = 30,50,100,250$ and each time with 1000 repetitions, so that each time a random sample with size n is generated with different parameter values from the desired distribution and then equation (12) based on $\alpha = 0.01, 0.1, 0.25,$ are calculated and at the end we average for each value of α. It means that

$$LGC = \frac{1}{1000}\sum_{i=1}^{1000} g_1^{(i)}[\alpha]\,, UGC = \frac{1}{1000}\sum_{i=1}^{1000} g_2^{(i)}[\alpha],$$

Where LGC and UGC are lower and upper bounds of Gini coefficient and for $i = 1,2,\dots,1000$ $g_j^{(i)}[\alpha]$ is $g_j[\alpha]$ value in the i-Th step (j=1, 2). We consider the parameter values of Dagum, Singh-Maddala and Log-normal distributions as follows,respectively:$\{(\lambda,\beta,\delta) = (20.12,5.12,3.26), (15.73,4.78,2.41), (5.28,8.23,2.10), (2.78,15.20,1.88)\}$, $\{(\lambda,\beta,\delta) = (2.72,3.14,2.84), (10.51,15.02,1.67), (22.17,9.28,1.45), (8.67,5.12,1.32)\}$

And$\{(\lambda,\beta)=(1.54,0.451),(2.98,0.642),(15.43,0.742),(10.98,0.845)\}$.

We have considered the values of the parameters in such a way that the Gini coefficients are equal to $GC=\{0.25,0.35,0.4,0.45\}$ according to these values, respectively. Details are reported in Tables 1 to 3.

Table 1
Lower and Upper estimate of Gini coefficient for membership degrees of 0.01

		GC=0.25		GC=0.35		GC=0.40		GC=0.45	
n	distribution	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
30	LN	0.142	0.311	0.200	0.428	0.232	0.488	0.263	0.544
	Da	0.191	0.363	0.265	0.514	0.316	0.548	0.370	0.578
	SM	0.183	0.414	0.301	0.425	0.334	0.510	0.358	0.636
50	LN	0.171	0.298	0.243	0.416	0.279	0.472	0.316	0.528
	Da	0.201	0.329	0.280	0.465	0.332	0.506	0.386	0.544
	SM	0.195	0.357	0.311	0.404	0.347	0.478	0.375	0.576
100	LN	0.200	0.289	0.280	0.401	0.322	0.457	0.364	0.512
	Da	0.213	0.301	0.297	0.424	0.350	0.470	0.403	0.513
	SM	0.208	0.315	0.321	0.386	0.361	0.451	0.394	0.530
250	LN	0.219	0.275	0.308	0.384	0.353	0.438	0.398	0.491
	Da	0.225	0.280	0.315	0.394	0.367	0.442	0.419	0.488
	SM	0.221	0.287	0.331	0.372	0.374	0.430	0.413	0.497

Table 2
Lower and Upper estimate of Gini coefficient for membership degrees of 0.1

		GC=0.25		GC=0.35		GC=0.40		GC=0.45	
n	distribution	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
30	LN	0.186	0.288	0.261	0.399	0.301	0.456	0.339	0.510
	Da	0.208	0.311	0.290	0.440	0.342	0.484	0.396	0.525
	SM	0.202	0.330	0.317	0.393	0.355	0.461	0.386	0.548
50	LN	0.201	0.280	0.284	0.392	0.326	0.445	0.368	0.499
	Da	0.216	0.295	0.302	0.416	0.354	0.463	0.407	0.506
	SM	0.211	0.307	0.324	0.382	0.364	0.445	0.399	0.521
100	LN	0.219	0.275	0.306	0.382	0.352	0.437	0.397	0.490
	Da	0.225	0.280	0.315	0.394	0.367	0.443	0.419	0.489
	SM	0.221	0.287	0.331	0.372	0.374	0.431	0.412	0.498
250	LN	0.231	0.266	0.323	0.371	0.371	0.424	0.417	0.476
	Da	0.234	0.268	0.327	0.377	0.379	0.426	0.430	0.474
	SM	0.231	0.272	0.338	0.364	0.383	0.419	0.425	0.479

Table 3
Lower and Upper estimate of Gini coefficient for membership degrees of 0.25

		GC=0.25		GC=0.35		GC=0.40		GC=0.45	
n	distribution	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
30	LN	0.205	0.276	0.287	0.382	0.330	0.437	0.373	0.490
	Da	0.219	0.290	0.305	0.408	0.359	0.456	0.415	0.500
	SM	0.199	0.300	0.332	0.379	0.368	0.440	0.390	0.513
	LN	0.215	0.270	0.304	0.378	0.347	0.430	0.392	0.483

50	Da	0.225	0.279	0.313	0.394	0.367	0.442	0.419	0.488
	SM	0.214	0.287	0.337	0.372	0.374	0.430	0.413	0.497
100	LN	0.228	0.268	0.319	0.372	0.376	0.426	0.413	0.478
	Da	0.231	0.270	0.322	0.380	0.333	0.429	0.428	0.477
	SM	0.210	0.275	0.342	0.365	0.381	0.421	0.421	0.482
250	LN	0.236	0.261	0.331	0.365	0.379	0.417	0.427	0.468
	Da	0.237	0.262	0.331	0.368	0.384	0.419	0.436	0.467
	SM	0.215	0.265	0.346	0.359	0.386	0.413	0.429	0.470

According to Tables 1 to 3 and, it can be said that the bounds become sharper as the sample size increases. Therefore, it seems that the bounds converge to the Gini coefficient when→∞, Also, it seems that the bounds obtained from the log-normal distribution (LNGC) are underestimated, but nevertheless, for each n=30,50,100,250 and a=0.01,0.1,0.25, all the bounds include their corresponding Gini coefficient.

6. Real Data Analysis

In this section, Iranian household income data is used. Iranian household income data is collected every year by the experts of Iran Statistics Center. These data can be obtained in both urban and rural areas and in the form of income and cost in the form of raw data in an Excel file from the website of the Iranian Statistics Center, https://www.amar.org.ir. The source of information in this section is household income data in 2011,2014,2017,2020,2021 in both urban and rural areas. Table 4 shows descriptive statistics related to Iranian household income (in millions of Rials) for these years in both rural and urban areas.

Table 4
Descriptive statistics related to Iranian household income

Areas	Year	Mean	Median	Max	Min	Sample size
Rural	2011	101.281	95.123	296.461	53.178	18521
	2014	183.168	142.369	351.120	72.123	18261
	2017	201.842	161.870	456.126	96.452	19261
	2020	340.679	246.189	542.136	109.104	18251
	2021	637.133	498.341	964.258	189.462	18632
Urban	2011	167.716	154.123	481.263	71.480	17623
	2014	241.318	189.462	734.782	97.630	18121
	2017	366.846	301.456	1043.238	128.132	18701
	2020	764.746	613.892	2410.453	198.462	19306
	2021	1124.210	964.789	4356.561	389.145	19112

First, the parameters of the models were estimated using maximum likelihood and then using equation (12), the upper and lower bounds of the Gini coefficient for α=0.01,0.2 were calculated based on the three desired distributions. Details are listed in Table 5. We are interested to visually compare the bounds of Gini coefficients obtained by Dagum model (DGC), Singh-Maddala model (SMGC) and Log-normal model (LNGC).

Table 5
The lower and upper estimate of Gini coefficient for membership degrees of 0.01 and 0.2

Areas	Year	distribution	$\alpha = 0.01$		$\alpha = 0.2$	
Rural			Lower	Upper	Lower	Upper
	2011	Log-Normal	0.2574	0.4181	0.2978	0.3777
		Dagum	0.3481	0.3844	0.3456	0.3704
		Singh-Maddala	0.3292	0.3947	0.3434	0.3757
	2014	Log-Normal	0.2605	0.4222	0.3011	0.3816
		Dagum	0.3490	0.3892	0.3462	0.3735
		Singh-Maddala	0.3182	0.4878	0.3373	0.3988
	2017	Log-Normal	0.2690	0.4341	0.3105	0.3926
		Dagum	0.3588	0.4012	0.3558	0.3848
		Singh-Maddala	0.3316	0.4014	0.3467	0.3812
	2020	Log-Normal	0.2613	0.4223	0.3020	0.3827
		Dagum	0.3525	0.3781	0.3507	0.3683
		Singh-Maddala	0.3224	0.4006	0.3378	0.3754
	2021	Log-Normal	0.2584	0.4193	0.2988	0.3789
		Sagum	0.3536	0.3803	0.3517	0.3701
		Singh-Maddala	0.3158	0.4081	0.3317	0.3742
Urban	2011	Log-Normal	0.2736	0.4407	0.3156	0.3987
		Dagum	0.3363	0.4021	0.3508	0.3833
		Singh-Maddala	0.3338	0.4060	0.3494	0.3850
	2014	Log-Normal	0.2761	0.4442	0.3183	0.4020
		Dagum	0.3496	0.3843	0.3577	0.3749
		Singh-Maddala	0.3529	0.3783	0.3610	0.3697
	2017	Log-Normal	0.2891	0.4624	0.3326	0.4189
		Dagum	0.3531	0.4129	0.3664	0.3960
		Singh-Maddala	0.3467	0.4209	0.3628	0.3994
	2020	Log-Normal	0.2941	0.4601	0.3308	0.4167
		Dagum	0.3631	0.4178	0.3754	0.4025
		Singh-Maddala	0.3549	0.4309	0.3714	0.4089
	2021	Log-Normal	0.2945	0.4700	0.3386	0.4259
		Dagum	0.3669	0.4091	0.3767	0.3976
		Singh-Maddala	0.3571	0.4281	0.3726	0.4077

The values of the Gini coefficients reported by the Iranian Statistics Centre for rural and urban areas show that the bound of the Gini coefficient obtained from the Dagum model is sharper than other bounds, so this bound is recommended. Second, the bound of the Gini coefficient obtained from the Singh-Maddala model is sharper than the bound obtained from the log-normal model, except for 2014 in rural areas. In addition, it can be said that the bound obtained from the log-normal model seems smoother than other bounds, which means that this bound is less sensitive to the amount of income difference.

7. Conclusion

In this article, we introduced the upper and lower bounds for the Gini coefficient by considering the fuzzy Dagum, Singh-Maddala and Log-normal distributions. Then, using the data of Iranian household income in the period of 2011,2014,2017,2020 and 2021, we obtained the values of these bounds according to the different degrees of membership. The simulation study and analysis of real data showed that the bounds obtained by Singh-Maddala and Dagum distributions are almost equivalent. Still, they perform better than the bounds obtained from the log-normal distribution.

References

[1] Al-Qudaimi, A. Kaur, K. and Bhat, SH. Triangular fuzzy numbers multiplication: QKB method. **Fuzzy Optimization and Modelling** 2(2), 34-40 (2021).

[2] Buckley, J. Fuzzy Probability and Statistics. Springer-Verlag Berlin Heidelberg (2006).

[3] Dagum, C. A New Model of Income Distribution, the Lorenz Curve and the Gini Concentration Ratio, forth comming paper (1977).

[4] Chotikapanich, D.William Griffiths,Hajargasht. G. Using the GB2 Income Distribution. econometrics, 21, 1-24 (2018).

[5] Dagum, C. The generation and distribution of income, the Lorenz curve and the Gini ratio, Economic Applique's, 33, 327–67 (1980).

[6] Dagum, C. A New Model of Personal Income Distribution: Specification and Estimation. Springer (1980).

[7] Ganjoei, R.A. Akbarifard, H. Mashinchi, M. and Jalaee, A. M. Estimation of upper and lower bounds of Gini coefficient by fuzzy data. Data in brief, 29, 212-224 (2020).

[8] Golden, J. A simple geometric approach to approximating the Gini coefficient. *J. Economic Education*, 39(1), 68-77 (2008).

[9] Fellman, J. Estimation of Gini coefficients using Lorenz curves. *Journal of Statistical and Econometric Methods*. 1(2), 31-38 (2012).

[10] Kleiber, C. and Kotz, S. Statistical Size Distributions in Economics and Actuarial Sciences, John Wiley and Sons, New York (2003).

[11] Kakamu, K. Simulation s tudies comparing Dagum and Singh-Maddala income distributions. *Computational Economics,* 48(4), 593-605 (2016).

[12] iber, C. Dagum versus Singh–Maddala income distributions, *Economics Letters,* 53, 265–8 (1996).

[13] McDonald J.B. and M. R. Ransom, An analysis of the bounds for the Gini coefficient, *Journal of Econometrics*, 17, 177-188 (1981).

[14] McDonald, J. B. Some generalized functions for the size distribution of income, Econometrica, 52 (1984),

[15] McDonald, J. B. and Mantrala, A. The distribution of income: revisited, *Journal of Applied Econometrics*, 10, 201–4 (1995).

[16] McDonald, J. B. and Xu, Y. J. A generalization for the beta distribution with applications, *Journal of Econometrics*, 66, 133–52 (1995).

[17] perez, G.C and Alaiz, P. Using Dagum model to explain changes in personal income distribution. *Applied Economic*, (43), 77-86 (2011).

[18] Pena, J. B., Callealta, J. Casas, J. M., Merediz, A. and Nunez, J. Distribution Personal de la Renta en Espana, Piramide, Madrid (1996).

[19] Salem, S. K. and Mount, G. S. A convenient descriptive model of income distribution: the gamma distribution, *Econometrica,* 42, 115–27 (1974).

[20] Victoria-Feser, M. P. Robust Methods For Personal Income Distribution Models with Applications to Dagum's models, in Reasearch on Economic Inequality, Vol. 1 (Eds) C. Dagum and A. Lemmi, JAI Press, Greenwich, CT, pp. 123–57 (1995).

[21] Victoria-Feser, M. P. Robust methods for the analysis of income distribution, inequality and poverty, *International Statistical Review,* 68, 277–93 (2000).

[22] Zimmerman H. J. Fuzzy set theory and its applications. 4th ed. Boston: Kluwer Nih off (2001).

Received January, 2024