The Dagum Distribution and its Application to the Income Distribution in Egypt

Mervat Elgohary Nahed Helmy Batool Shaaban

Faculty of Commerce

Al – Azhar University, Girls' Branch

Abstract

The present study introduces Dagum distribution Type I with three parameters and its properties. Maximum likelihood estimation method is used to estimate the model parameters. Income data for urban and rural sectors in Egypt for two periods of time: 2008/2009 and 2012/2013 are used. The fit of the Dagum distribution against the Burr Type XII and the lognormal with three parameters (3P) distributions is compared by using one-sample Kolmogorov-Smirnov test, and three other criteria for model selection: Akaike information, Bayesian information criteria and log likelihood index. Gini coefficient is used as a measure of income inequality, its observed value is compared with the estimated value for the three models of income.

Key words: *Income, Dagum distribution, Burr distribution, lognormal distribution, Kolmogorov-Smirnov test, Akaike information, Bayesian information, log likelihood index, Gini coefficient.*

1. Introduction

There are many models that are presented to describe the distribution of income. The two models most often used are the Pareto and the lognormal distributions. The Pareto distribution fits high levels of income while the lognormal fits low levels of income, therefore other models were suggested to overcome this problem such as the generalized beta distribution of the second kind (GB2), Burr Type XII, Dagum distribution which is a sub-model of GB2 distribution. Mcdonald (1984) presented two generalized beta distributions. He applied the models to US family income data. Younan (1984) derived the probability distribution function of income with the application on urban and rural sectors in Egypt in the year 1978. Younan (1988) modeled statistical income data using various distribution functions to fit the data and selecting the one that best fits the data according to some goodness of fit criterion. Gertel *et al.* (2003) applied the Dagum distribution to perform a comparative analysis of income distribution in the Capital region of Argentina of individual income receivers. Lukasiewicz *et al.* (2010) made a comparison among four models with various numbers of parameters: exponential, Weibull, Dagum, and Singh-Maddala to know which model can represents the data that comes from the personal incomes in USA by using some of the important measures such as the sum of squared residuals, the sum of absolute values of the residuals. Tartaľová (2013) used the Dagum and Singh-Maddala distributions to describe the data which are taken from Slovak household's income data. She compared the two models with two other ones which are commonly used in income distributions such as the Pareto and the lognormal distributions. Candino and Domma (2013) proposed a new distribution called Beta Dagum distribution which includes some important distributions as special cases. They applied this model on the data, from survey of household’s income and wealth of Bank of Italy. Huang and Oluyede (2014) proposed a new family of distributions called exponentiated Kumaraswamy-Dagum distribution and applied it to income and lifetime data. Oluyede *et al.* (2014) presented a new class of generalized Dagum distribution called gamma Dagum distribution with applications to income and lifetime data.

The aim of the present paper is to apply Dagum distribution Type I with three parameters to the individual household income for the rural and urban sectors for two periods of time 2008/2009 and 2012/2013. Two other models of income Burr distribution and lognormal distribution (3p) are compared in order to study the best fitted model in presenting the data according to some goodness of fit criteria. Gini coefficient is used as a measure of income inequality. Its value is calculated from the income data and also derived as function in the estimated parameters of the three models. The accuracy of the Gini coefficient (the difference between the observed and the estimated Gini coefficient) is used as an additional goodness of fit criterion.

This paper is organized as follows. In Section 2 the Dagum distribution and its main properties are presented. Section 3 contains some important relationships between the Dagum model and other important models. The maximum likelihood estimation for the model parameters is presented in Section 4. Two other models for income are introduced in Section 5. The application of the three models on income data of Egypt is presented in Section 6. Section 7 contains Gini coefficient as a measure of inequality. Conclusions are discussed in Section 8.

2. The Dagum Distribution and its Main Properties

Dagum (1977) proposed the distribution which is referred to as Dagum distribution which is based on the log logistic distribution by adding another parameter. It is also called the generalized logistic-Burr distribution. There is both a three-parameter specification (Type I) and a four-parameter specification (Type II) of the Dagum distribution. The present paper applies Type I with three parameters.

Let X be a random variable with Dagum distribution. Its probability density function (pdf) takes the form:

$$f(x;a,b,p) = apbx^{-a-1}(1+bx^{-a})^{-p-1}, \qquad x>0,\ (a,p,b>0) \qquad (1)$$

has the following (cdf)The corresponding cumulative distribution function form:

$$F(x;a,b,p) = (1+bx^{-a})^{-p}, \qquad x>0,\ (a,p,b>0) \qquad (2)$$

is a scale parameter$b$ are shape parameters while a and $p$where

The quantile function (inverse cumulative distribution function) is:

$$F^{-1}(x;a,b,p) = \left(\frac{b}{q^{-\frac{1}{p}}-1}\right)^{\frac{1}{a}}, \qquad 0<q<1 \qquad (3)$$

If $q=0.25$ the first quartile is obtained, if $q=0.5$ the second quartile (median) is obtained and if $q=0.75$ the third quartile is obtained.

The r^{th} non-central moment of the Dagum distribution is given by:

$$E(x^r) = pb^{\frac{r}{a}}\beta\left(1-\frac{r}{a}\ ,p+\frac{r}{a}\right) \quad a>r \tag{4}$$

where $\beta(\ .,.\ )$ is the complete beta function.

can be written as: X The mean of the random variable

$$E(x) = b^{\frac{1}{a}}\frac{\Gamma\left(1-\frac{1}{a}\right)\Gamma\left(p+\frac{1}{a}\right)}{\Gamma(p)} \quad where\ \ a>1 \tag{5}$$

can be written as: X The variance for the random variable

$$V(x) = \frac{b^{\frac{2}{a}}}{\Gamma(p)}\left\{\Gamma(1-\tfrac{2}{a})\Gamma(p+\tfrac{2}{a}) - \frac{\left(\Gamma(1-\frac{1}{a})\Gamma(p+\frac{1}{a})\right)^2}{\Gamma(p)}\right\} \tag{6}$$

where $\Gamma(.)$ is the complete Gamma function.

is obtained as: $\propto_3$ The measure of the skewness,

$$\propto_3 = \frac{\mu_3}{\mu_2^{3/2}} = \left(\Gamma(p)\right)^{1/2}\left[\frac{\Gamma\left(1-\frac{3}{a}\right)\Gamma\left(p+\frac{3}{a}\right)-\frac{3\Gamma\left(1-\frac{2}{a}\right)\Gamma\left(p+\frac{2}{a}\right)\Gamma\left(1-\frac{1}{a}\right)\Gamma\left(p+\frac{1}{a}\right)}{\Gamma(p)}+\frac{2\left(\Gamma\left(1-\frac{1}{a}\right)\right)^3\left(\Gamma\left(p+\frac{1}{a}\right)\right)^3}{\left(\Gamma(p)\right)^2}}{\left(\Gamma\left(1-\frac{2}{a}\right)\Gamma\left(p+\frac{2}{a}\right)-\frac{\left(\Gamma\left(1-\frac{1}{a}\right)\Gamma\left(p+\frac{1}{a}\right)\right)^2}{\Gamma p}\right)^{3/2}}\right] \tag{7}$$

is obtained as: $\propto_4$ The measure of the kurtosis,

$$\propto_4 = \frac{\mu_4}{\mu_2^2} = \Gamma(p)\left[\frac{\Gamma\left(1-\frac{4}{a}\right)\Gamma\left(p+\frac{4}{a}\right)-\frac{4\Gamma\left(1-\frac{3}{a}\right)\Gamma\left(p+\frac{3}{a}\right)\Gamma\left(1-\frac{1}{a}\right)\Gamma\left(p+\frac{1}{a}\right)}{\Gamma(p)}+\frac{6\Gamma\left(1-\frac{2}{a}\right)\Gamma\left(p+\frac{2}{a}\right)\left(\Gamma\left(1-\frac{1}{a}\right)\Gamma\left(p+\frac{1}{a}\right)\right)^2}{\left(\Gamma(p)\right)^2}-3\frac{\left(\Gamma\left(1-\frac{1}{a}\right)\Gamma\left(p+\frac{1}{a}\right)\right)^4}{\left(\Gamma(p)\right)^3}}{\left(\Gamma\left(1-\frac{2}{a}\right)\Gamma\left(p+\frac{2}{a}\right)-\frac{\left(\Gamma\left(1-\frac{1}{a}\right)\Gamma\left(p+\frac{1}{a}\right)\right)^2}{\Gamma p}\right)^2}\right] \tag{8}$$

The mode of this distribution is obtained as:

$$x_{mode} = b^{\frac{1}{a}}\left(\frac{ap-1}{a+1}\right)^{\frac{1}{a}}, \qquad ap>1 \tag{9}$$

The reliability, R(x) function is obtained as:

$$R(x) = 1-(1+bx^{-a})^{-p}, \qquad x>0 \tag{10}$$

The hazard rate, H(x) is given by:

$$H(x) = \frac{apbx^{-a-1}(1+bx^{-a})^{-p-1}}{1-(1+bx^{-a})^{-p}}, \quad x>0$$

(11)

3. Some Important Relationships

There are some relationships between the Dagum distribution and other important distributions such as: Burr Type XII, Beta Type I, II, Pareto, exponential.

- If we add a scale parameter to Burr Type III distribution, this distribution will be transferred to the Dagum distribution.

Table (1) summarizes the transformations from Dagum to other distributions.

Table (1) Summary of transformations applied to the Dagum and resulting distributions

transformation	distribution
$\frac{1}{x}$	Burr Type XII $(a, \frac{1}{b}, p)$
bx^{-a}	Beta Type II $(1, p)$
$\frac{1}{(1+bx^{-a})}$	Power function $(p, 1)$
$(1+bx^{-a})$	Pareto $(1, p)$
$\ln(1+bx^{-a})$	Exponential (p)
$-\ln bx^{-a}$	Generalized logistic Type $I(p)$

4. Parameters Estimation

In this section, the maximum likelihood method is used to estimate the parameters of the Dagum distribution.

The likelihood function of the Dagum distribution for the parameters a, p, b is given by:

$$L(\underline{x}; a, p, b) = (apb)^n \prod_{i=1}^{n} x^{-a-1} \prod_{i=1}^{n} (1+bx^{-a})^{-p-1} \qquad (12)$$

Taking the natural logarithm for the likelihood function of n observations to be as follows:

$$\text{Log } L(\underline{x}; a, p, b) = n\log a + n\log p + n\log b - (a+1)\Sigma \log x - (p+1)\Sigma \log(1+bx^{-a}) \qquad (13)$$

Partially differentiating (13) with respect to the parameters a, p, b respectively, and equating to zero we get the following equations:
denote $L(\underline{x}; a, p, b) = l$

$$\frac{\partial \log l}{\partial a} = \frac{n}{\hat{a}} + \Sigma \log x + (\hat{p} + 1)\Sigma \frac{1}{(1 + \hat{b}x^{-\hat{a}})} \hat{b}x^{-\hat{a}} \log x = 0 \tag{14}$$

$$\frac{\partial \log l}{\partial p} = \frac{n}{\hat{p}} - \Sigma \log(1 + \hat{b}x^{-\hat{a}}) = 0 \tag{15}$$

$$\frac{\partial \log l}{\partial b} = \frac{n}{\hat{b}} - (\hat{p} + 1)\Sigma \frac{1}{(1 + \hat{b}x)} x^{-\hat{a}} = 0 \tag{16}$$

The maximum likelihood estimates (MLEs) of a, p, b can be obtained by solving (14) to (16) simultaneously. Equation (15) can be solved and we obtain $\hat{p}$ as:

$$\hat{p} = \frac{n}{\Sigma \log(1 + \hat{b}x^{-\hat{a}})} \tag{17}$$

The estimators of the parameters *a* and *b*, can't be obtained in closed form, therefore numerical method is used.

5. Two Other Models of Income

The two other models of income which are selected to compare with Dagum distribution Type I are Burr Type XII distribution and lognormal distribution with three parameters.

- **Burr Type XII**

The pdf of the Burr Type XII distribution is given by:

$$f(x; \alpha, k, \beta) = \frac{\alpha k (\frac{x}{\beta})^{\alpha - 1}}{\beta (1 + (\frac{x}{\beta})^{\alpha})^{k+1}}, \qquad x > 0, \quad (\alpha, k, \beta > 0) \tag{18}$$

where α and k are the shape parameters while β is a scale parameter.
has the following form:cdf The corresponding

$$F(x; \alpha, k, \beta) = 1 - \frac{1}{(1 + (\frac{x}{\beta})^{\alpha})^{k}}, \qquad x > 0, \quad (\alpha, k, \beta > 0) \tag{19}$$

- **The lognormal distribution with three parameters**

The pdf of the lognormal (3P) distribution is given by:

$$f(x;\gamma,\mu,\sigma)=\left(\frac{1}{(x-\gamma)\sigma\sqrt{2\pi}}\right)e^{-\frac{1}{2}\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)^{2}},\ \gamma<x<\infty,\quad -\infty<\mu<\infty,\ \sigma>0 \tag{20}$$

.

where μ is a shape parameter, σ is a scale parameter and γ is the threshold or location parameter.

has the following form:cdf The corresponding

$$\Phi(x;\gamma,\mu,\sigma)=\Phi\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right),\qquad \gamma<x<\infty,-\infty<\mu<\infty,\ \sigma>0 \tag{21}$$

where Φ is the cdf of the standard normal distribution.

6. Application to Egypt income data

6.1 Description of income data

The income data are obtained from Income, Consumption and Expenditure survey for the urban and rural sectors for Egypt for two periods of time the first in 2008/2009 (before January 25th revolution) and the second in 2012/2013 (after January 25th revolution). The descriptive statistics for the distribution of income are shown in Table (2).

Table (2) Descriptive statistics for the distribution of income for urban and rural sectors in the two periods 2008/2009 and 2012/2013

Sector / Measure	2008/2009		2012/2013	
	Urban	Rural	Urban	Rural
	Sample size $n=3298$	Sample size $n=3271$	Sample size $n=3294$	Sample size $n=4124$
Min	17.59	12.1	21.34	12.34
Max	5435	2696	5499.5	6127
Range	5417	2684	5478.2	6114.7
Mean	239.2	177.44	340.03	269.94
Median	187.2	150.5	278.4	241.11
Var.	54269	18418	75912	31924
S.D.	232.9	135.71	275.52	178.67
C.V.	0.97	0.76	0.81	0.66
Skewness	8.18	7.26	7.09	11.23
Kurtosis	122.5	99.51	97.39	307.45

Table (4.1) illustrates a higher gap in urban sector than the rural sector in the year 2008/2009 but the opposite is happened in the year 2012/2013. Also the values of the variance (Var.) and standard deviation (S.D) for urban sector are higher than for rural sector. This means that the data points of urban are more spread out over a large range of values than the rural. The coefficient of variation (C.V) refers to individuals in rural are more homogeneous than in urban in the two periods. The mean and median income for urban sector is higher than its value for rural sector and there is an increase in mean and median income for both sectors in 2012/2013. The curve of the income distribution is skewed to the right for both urban and rural sectors and very peaked.

6.2 Parameters Estimates for Dagum, Burr Type XII and Lognormal distributions

In this section, the estimates of the parameters for the three distributions are obtained by using the maximum likelihood method. The standard error of the estimates for the three models: Dagum, Burr Type XII and lognormal (3P) distributions are obtained. Program R I386 3.1.1 software is used to satisfy this purpose. The estimates and their standard error for the Dagum, Burr Type XII and lognormal (3P) distributions for urban and rural sectors in the two periods 2008/2009 and 2012/2013 are shown in Table (3).

Table (3) The estimates and their standard error for the Dagum, Burr Type XII and lognormal (3P) distributions for urban and rural sectors in the two periods 2008/2009 and 2012/2013

distribution	year	sector	$\hat{a}$	Std. Error	$\hat{b}$	Std. Error	$\hat{p}$	Std. Error
Dagum distribution	2008/2009	Urban	2.68	0.06	150.04	6.351	1.57	0.119
		Rural	3.28	0.084	146.97	4.411	1.06	0.06
	2012/2013	Urban	3.028	0.075	258.4	8.56	1.201	0.078
		Rural	3.99	0.096	266.11	5.31	0.762	0.037
	year	sector	$\hat{\mu}$	Std. Error	$\hat{\sigma}$	Std. Error	$\hat{\gamma}$	Std. Error
Lognormal(3P) distribution	2008/2009	Urban	5.20	0.016	0.638	0.011	10.51	1.79
		Rural	5.047	0.018	0.537	0.011	-3.55	2.22
	2012/2013	Urban	5.66	0.0178	0.565	0.011	-2.15	3.66
		Rural	5.476	0.012	0.508	0.008	-2.33	2.102

distribution	year	sector	$\hat{a}$	Std. Error	$\hat{b}$	Std. Error	$\hat{p}$	Std. Error
	year	sector	$\hat{\alpha}$	Std. Error	$\hat{\beta}$	Std. Error	$\hat{k}$	Std. Error
Burr Type XII distribution	2008/2009	Urban	3.627	0.103	157.93	3.943	0.665	0.0371
		Rural	3.483	0.092	143.93	3.878	0.897	0.053
	2012/2013	Urban	3.491	0.096	255.982	6.852	0.806	0.047
		Rural	3.344	0.073	263.424	7.101	1.254	0.073

6.3 Goodness of fit

In order to evaluate relative performance of the three distributions we evaluate one-sample Kolmogorov-Smirnov test, Akaike information criterion and Bayesian information criterion. The Akaike information criterion (AIC) is a measure of the relative quality of a statistical model for a given set of data. It deals with the trade-off between the goodness of fit of the model and the complexity of the model. AIC is given by:

$$AIC = -2\log L(\hat{\theta}|y) + 2\boldsymbol{k}.$$

The Bayesian information criterion (BIC) or Schwarz criterion is a criterion for model_selection among a finite set of models. It is based, in part, on the likelihood function and it is closely related to the AIC; the model with the lowest AIC and BIC is preferred. BIC is given by:

$$BIC = -2\ln\hat{L} + \boldsymbol{k}\ln n,$$

where L: is the value of the likelihood function evaluated at the parameter estimates,

n: is the number of observations and

k: is the number of estimated parameters

Results of goodness of fit are shown in Table (4).

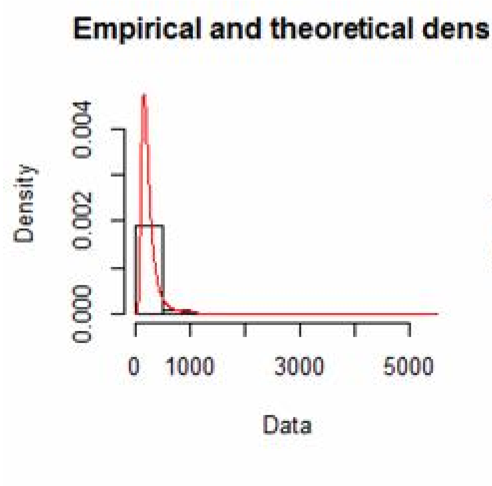
Table (4) Goodness of fit criteria for Dagum distribution, Burr Type XII and lognormal (3P) distributions

The Distribution	Sector	Year	K.S test		AIC	BIC	Log likelihood
			Test statistics	P value			
The Dagum distribution	Urban	2008/2009	0.013	0.599	40535.5	40553.8	-20264.7
		2012/2013	0.0144	0.5057	42742.01	42760.31	-21368
	Rural	2008/2009	0.009	0.954	38069.43	38087.71	-19031.7
		2012/2013	0.0117	0.6296	51054.86	51073.83	-25524.43
The log-normal distribution	Urban	2008/2009	0.0416	0.000	40709.15	40727.45	-20351.57
		2012/2013	0.0382	0.000	42890.15	42908.45	-21442.07
	Rural	2008/2009	0.0349	0.000	38264.89	38283.17	-19129.44
		2012/2013	0.1508	0.000	51304.12	51323.1	-25649.06
The Burr distribution	Urban	2008/2009	0.0106	0.849	40529.81	40548.1	-20261.9
		2012/2013	0.0115	0.779	42737.51	42755.81	-21365.76
	Rural	2008/2009	0.0077	0.990	38067.07	38085.35	-19030.53
		2012/2013	0.0142	0.3723	51066.92	51085.9	-25530.46

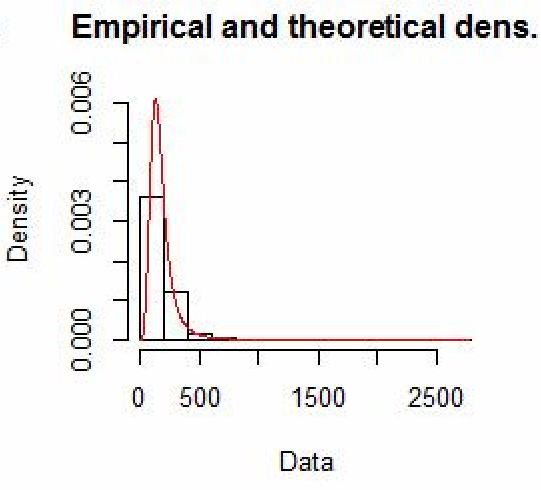
The AIC and BIC indicators determine which distribution is the best fitting for the empirical distribution by the lowest value of these indicators, and the last column shows the log likelihood value and the preference will be at its highest value. For both sectors and both years we observe that the Dagum and Burr Type XII (3p) distributions fit the empirical distribution using one−sample Kolmogorov−Smirnov test since the p −value for them is greater than the level of significance $\alpha = 0.05$. But the lognormal distribution doesn't fit the empirical distribution. For both years and both sectors except rural 2012/2013, the AIC, BIC and the log likelihood indicators showed that the Burr distribution is slightly better than the Dagum distribution, then the lognormal distribution came last in preference. For Rural sector in the year 2012/2013 it is found that the Dagum distribution is better than the Burr distribution. The AIC, BIC and the log likelihood indicators showed that both Dagum and Burr Type XII distributions have almost the same preference in the representation of the actual data.

The following figures provide a visual comparison between the theoretical distribution and the empirical distribution. The Q-Q plots are used to give more insight into the nature of the difference between the theoretical and the empirical distribution. Figure (a) represents urban sector while Figure (b) represents rural sector, Figure (c) is a Q-Q plot for urban sector while Figure (d) is a Q-Q plot for rural sector. The red curve is the theoretical distribution but the histogram is the empirical one.

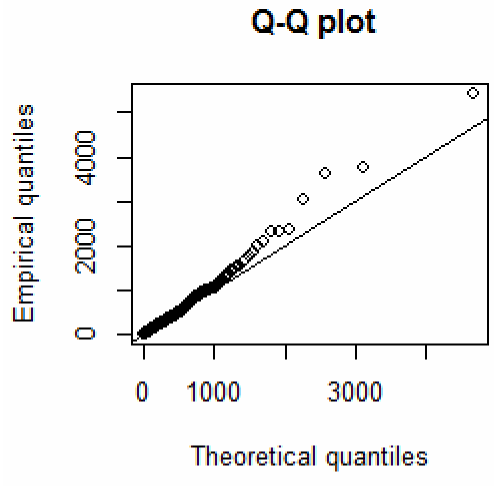
(a)



(b)



(c)



(d)

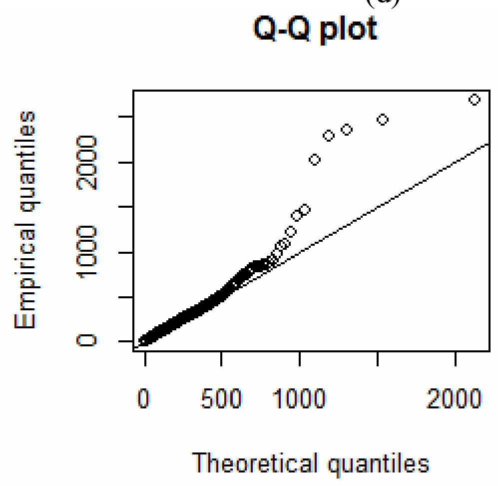
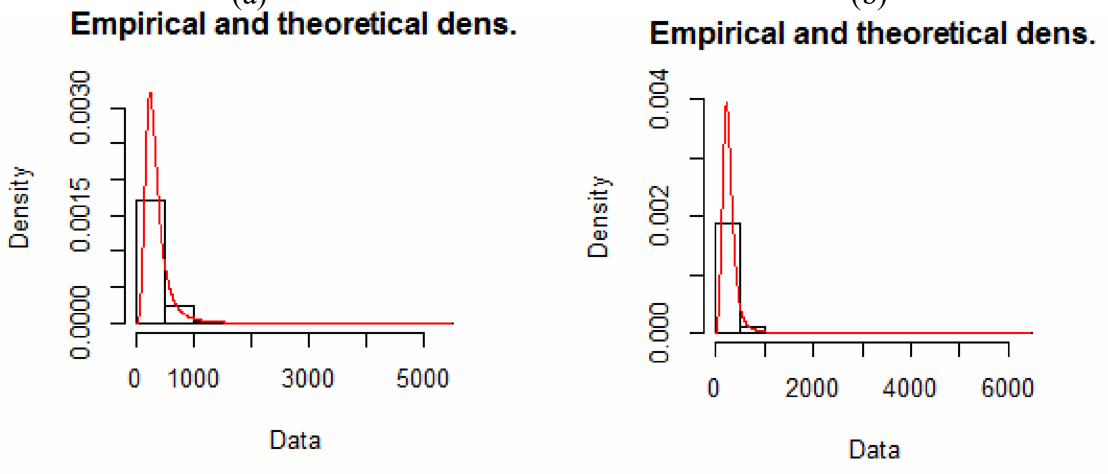


Figure (1) Fitting of the Dagum distribution to the income data for urban and rural sectors in year 2008/2009.

Figure (1) shows that the Dagum distribution fits well the low and middle income levels, but the fit is not good for upper levels of income.

(a) (b)



(c) (d)

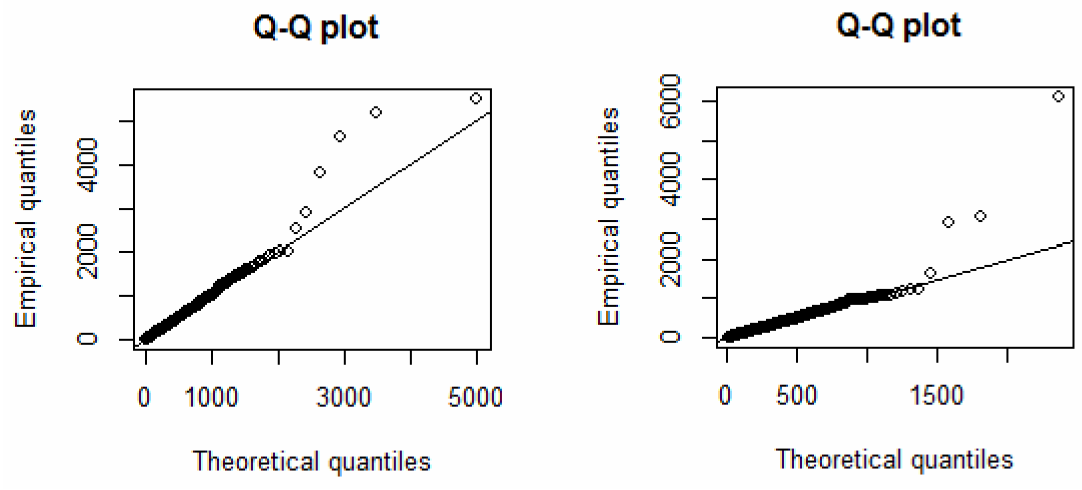
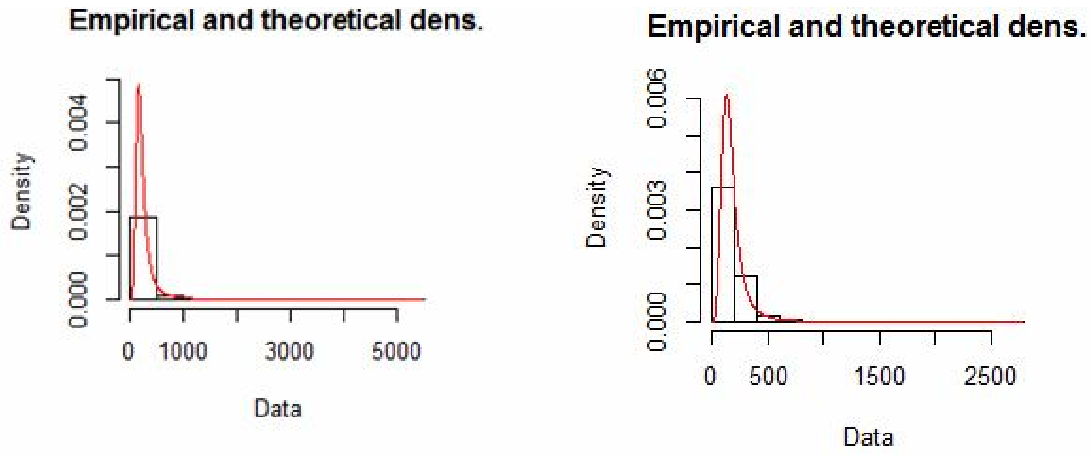


Figure (2) Fitting the Dagum distribution to the income data for urban and rural sectors in 2012/2013

From this figure we notice that the Dagum distribution fits well the low and middle levels of income.

(a) (b)



(c) (d)

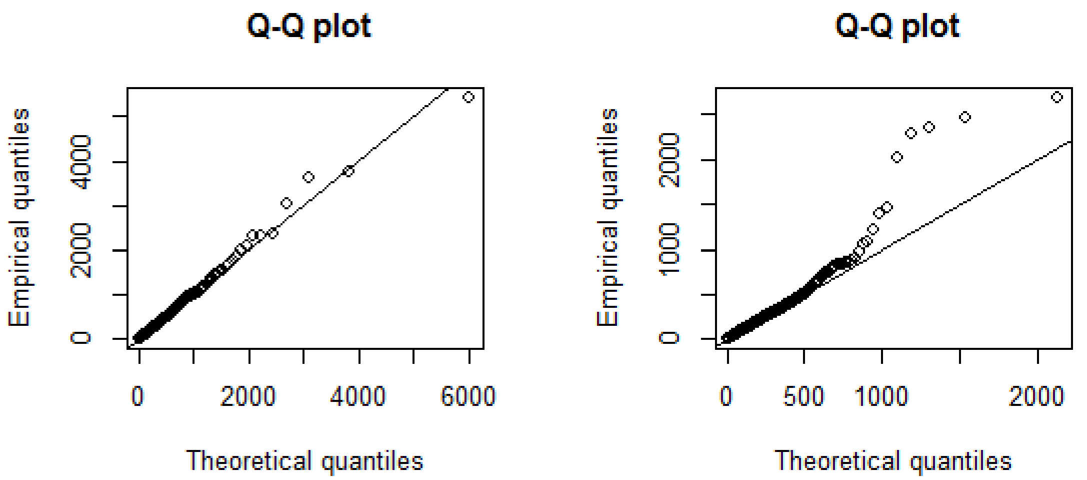
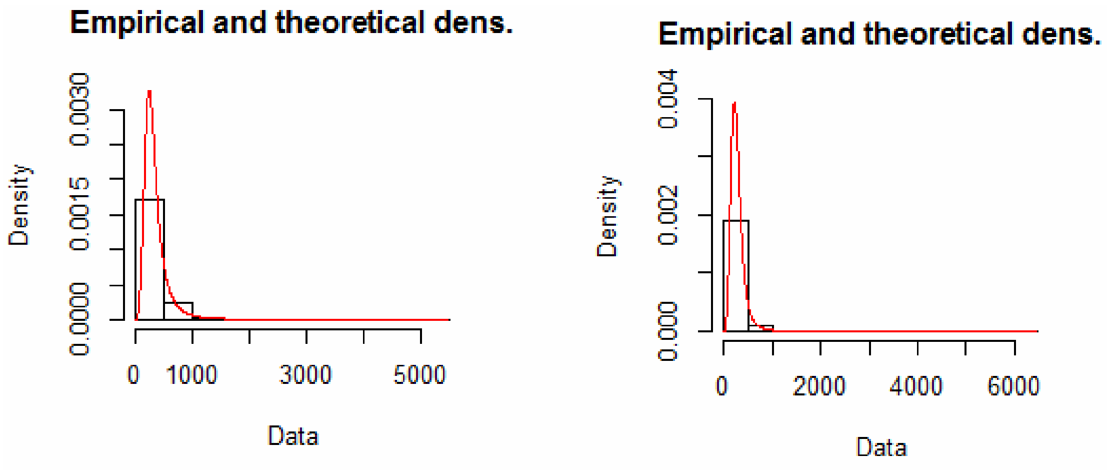


Figure (3) Fitting Burr Type XII distribution to the income data for urban and rural sectors in 2008/2009

This figure shows that the Burr distribution is better in urban sector than rural sector for all income levels.

(a) (b)



(c) (d)

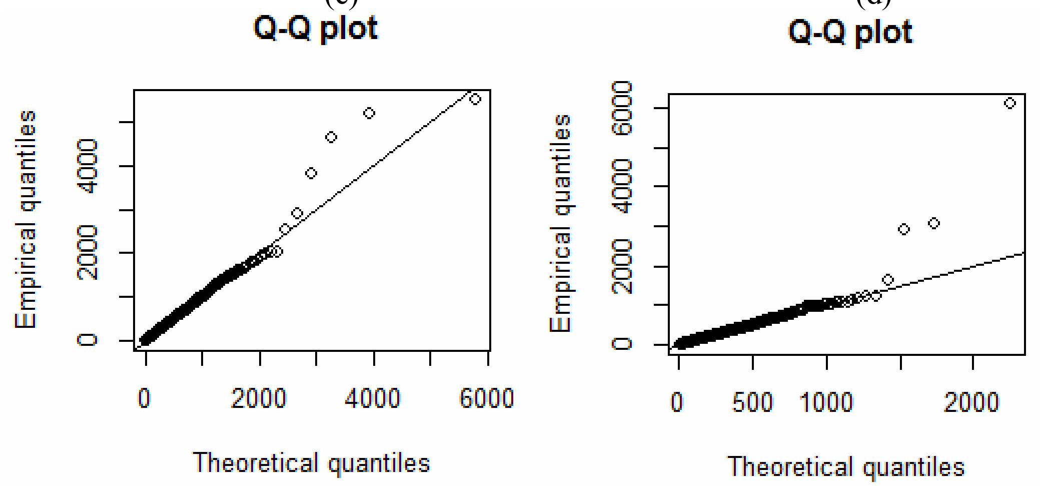
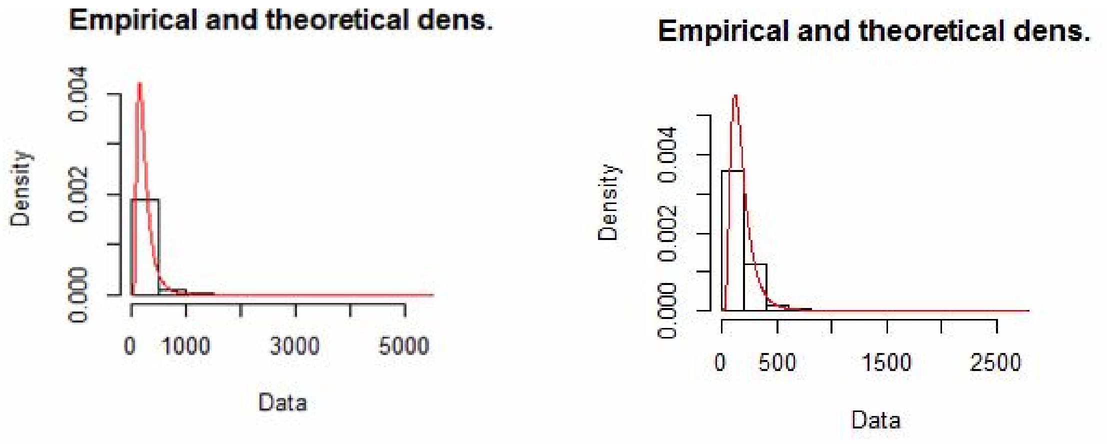


Figure (4) Fitting Burr Type XII distribution to the income data for urban and rural sectors in 2012/2013

Figure (4) illustrates that the Burr distribution for rural sector is the same as urban sector and the fit is good for low and middle levels.

(a) (b)



(c) (d)

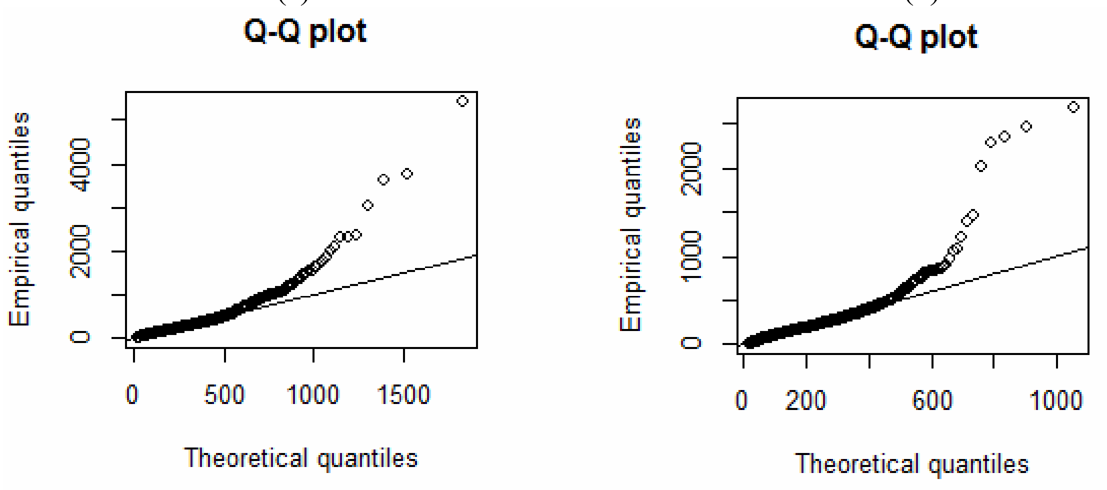


Figure (5) Fitting the lognormal (3P) distribution to the income data for urban and rural sectors in year 2008/2009

It is clear from this figure that the lognormal distribution fits only low levels of income.

(a) (b)

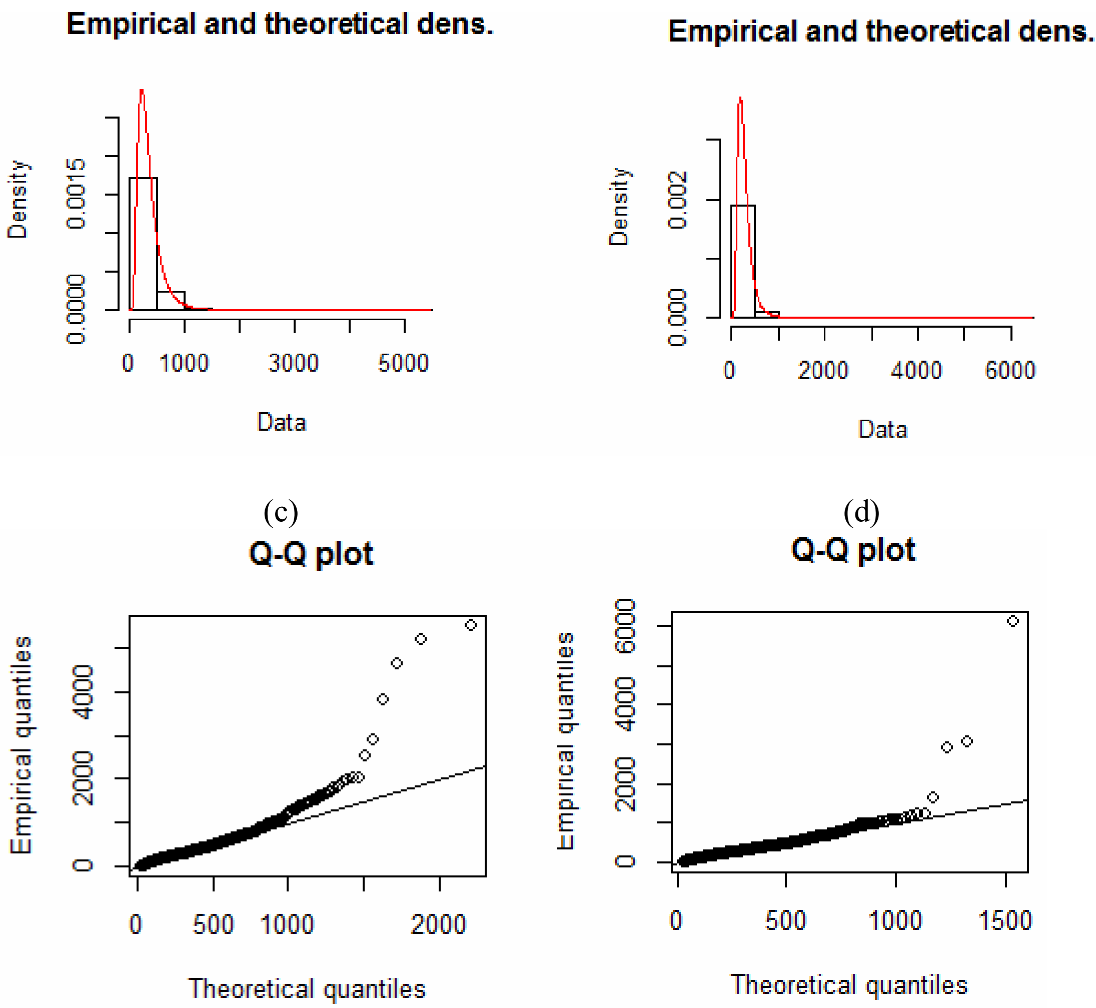


Figure (6) Fitting the lognormal (3P) distribution to the income data for urban and rural sectors in year 2012/2013.

Figure (6) shows that the lognormal distribution is a good fit to the empirical distribution. The fit is good for low and middle levels of income.

7. Income Inequality Measure

The Gini coefficient is a measure of income inequality which is an indication of social welfare. It is defined as a ratio with values between 0 and 1. 0 corresponds to perfect income equality (i.e. everyone has the same income) and 1 corresponds to perfect income inequality (i.e. one person has all the income, while everyone else has zero income). Gini coefficient can be expressed mathematically in terms of the distribution's parameters estimates for the three models as follows.

For Dagum distribution:

$$G = \frac{\Gamma(\hat{p})\Gamma\left(2\hat{p}+\frac{1}{\hat{a}}\right)}{\Gamma(2\hat{p})\Gamma\left(\hat{p}+\frac{1}{\hat{a}}\right)} - 1 \qquad \text{(Dagum, 1977)} \qquad (22)$$

For Burr Type XII distribution:

$$G = 1 - \left(\frac{\Gamma(\hat{k})\Gamma\left(2\hat{k}-\frac{1}{\hat{a}}\right)}{\Gamma(2\hat{k})\Gamma\left(\hat{k}-\frac{1}{\hat{a}}\right)}\right) \qquad \text{(McDonald, 1984)} \qquad (23)$$

For lognormal (3P) distribution:

$$G = \lambda\left(2\Phi\left(\frac{\hat{\sigma}}{\sqrt{2}}\right) - 1\right) \qquad \text{(Groll and Lambert, 2013)} \qquad (24)$$

where Φ *is the* $N(0,1)$ *distribution function* and $\lambda = \frac{1}{1+\hat{\gamma}e^{-(\hat{\mu}+\frac{1}{2}\hat{\sigma}^2)}}$

The corresponding Gini coefficients were estimated from the parameters estimates of the models applying the Mathcad 14 software to solve Equations (22), (23) and (24).

Table (5) The observed and estimated Gini coefficients for rural and urban sectors in the two periods 2008/2009 and 2012/2013.

sector	years	models	estimated Gini coefficient (G_E)	observed Gini coefficient (G_O)	$\lvert G_O - G_E \rvert$ accuracy of the Gini coefficient
urban	2008/2009	Dagum	0.343	0.353	0.010
		Lognormal(3p)	0.332		0.021
		Burr XII	0.353		0.000
	2012/2013	Dagum	0.318	0.323	0.005
		Lognormal(3p)	0.312		0.011
		Burr XII	0.323		0.000
rural	2008/2009	Dagum	0.301	0.318	0.017
		Lognormal(3p)	0.302		0.016
		Burr XII	0.305		0.013
	2012/2013	Dagum	0.270	0.274	0.004
		Lognormal(3p)	0.283		0.009
		Burr XII	0.271		0.003

From Table (5) in the two periods 2008/2009 and 2012/2013 it is noticed that, the rural sector has less income inequality than the urban sector. In both sectors the income inequality decreased through the two years. The last column in this table indicates the accuracy of the Gini coefficient which is used as an additional criterion of goodness of fit of the models. We can see that the fit is good for Burr distribution followed by Dagum and lognormal comes at last for urban sector in the two periods. For rural sector the fit is good equally for the three distributions.

8. Conclusions

This paper compares the ability of 3 probability distributions to fit income data of rural and urban sectors of Egypt for two periods of time .The goodness of fit of the three models is evaluated using combined criteria (analytical criteria, visual criteria and the accuracy of the Gini coefficient).

Using the combined evaluation criteria, the Dagum distribution is considered the best fitting model over a large part of the lower and middle income levels but not at the upper levels of income.
The Burr distribution gives similar results to Dagum except for the accuracy of the Gini coefficient for urban sector in year 2008/2009.
The lognormal distribution fits the data well using visual representation and the accuracy of the Gini coefficient but gives poor results using the analytical measures.

References

Candino, F. and Domma, F. (2013). 'The Beta-Dagum Distribution.' *Communications in Statistics*, Vol. 42, Issue 22, pp. 4070-4090.

CAPMAS August (2010). Income, Expenditure & Consumption Survey Vol. I.

CAPMAS April (2014). Income, Expenditure & Consumption Survey Vol. I.

Dagum, C. (1977). 'A New Model of Personal Income Distribution: Specification and Estimation.' *Economie Applique'e, Tomo XXX*, No. 3, pp. 413–436.

Gertel, H. R., Giuliodori, R, Auerbach, P.F and Rodríguez, A.F. (2003). 'A Parametric Estimation of Personal Income Distribution in Argentina using the Dagum Model.' *Special Issue of the Journal of the Inter-American Statistical Institute (IASI), "Estadística"* Vol. 55, pp. 164-165.

Groll. T. and Lambert. J.P. (2013). 'The Pro-Poorness, Growth and Inequality Nexus: some Findings from a Simulation Study.' *Review of Income and Wealth*, Vol. 59, Issue 4, pp. 776–784.

Huang, S. and Oluyede, B. O. (2014). 'Exponentiated Kumaraswamy-Dagum Distribution with Applications to Income and Lifetime Data.' *Journal of Statistical Distributions and Applications*, pp. 1-8.

Lukasiewicz, P., Karpio, K. and Orłowski, A. (2010). 'The Models of Personal Incomes in USA.' *Proceedings of the 5th Symposium on Physics in Economics and Social Sciences*, Vol. 121, pp. 25–27.

McDonald, J.B. (1984). 'Some Generalized Functions for the Size Distribution of Income.' *Econometrica*, Vol. 52, No.3, pp. 1-18.

Oluyede, B. O., Huang, S., and Pararai, M. (2014). 'A New Class of Generalized Dagum Distribution with Applications to Income and Life time Data.' *Journal of Statistical and Econometric Methods*, Vol. 3, No. 2, pp.125-151.

Tartaľová, A. (2013). 'Modeling Income Distribution in Slovakia.' *Statistics and Economy Journal*, Vol. 93, No.2, pp. 1-10.

Younan, W. (1984). The Derivation of the Distribution Function for the Egyptian Income Distribution and Measuring Its Inequality. M.A. Thesis, Dept. of Statistics, Faculty of Economics and Political Science, Cairo University. Unpublished.

Younan, W. Y. (1988). On Modeling Statistical Data and Measuring the Distribution Inequality with Special Reference to Income Data. PHD Thesis, Dept. of Statistics, Faculty of Economics and Political Science, Cairo University. Unpublished.