Estimation of an Appropriate Function for the Household Income Distribution of Bangladesh and Income Inequality

Md.Sharif, Hossain

https://doi.org/10.15017/3000239

出版情報：経済論究. 106, pp.113-132, 2000-03-31. 九州大学大学院経済学会

Estimation of an Appropriate Function for the Household Income Distribution of Bangladesh and Income Inequality

Hossain Md. Sharif

Abstract

Many distributions have been developed as descriptive models for the size distribution of income. In this paper, at the first stage, we considered the two parameters models, like Lognormal and Gamma families and at the second stage we considered the three and four parameters models like Log-t, Generalized Gamma, Singh-Maddala and Generalized Beta distributions as descriptive models for the household income distribution of Bangladesh. All of these considered functions do not fit the real data of the household income distribution of Bangladesh very well. That is why, the principle purpose of this paper is to select an appropriate distribution that fits the size distribution of household income of Bangladesh very well. And another attempt has been made to calculate some important measures of income inequality on the basis of the appropriate distribution. For the purpose of comparing the distributions, we also used a non-linear maximization technique and adopted the parameters value, probabilities, sum of squared errors criteria, sum of absolute errors criteria and the chi-square value of the different distributions. On the basis of these criteria the relative performance of the different distributions is compared.

An empirical application, including a comparison the considered distributions to demonstrate the better fit is made to household income data of Bangladesh for the years, 1995-96, 1991-92, 1988-89 and 1985-86. In our collected data, measurement errors arise from approximation and rounding off. That is why, on the basis of the selected distribution we have calculated some inequality measures, like, Relative Mean Deviation, Variance, Standdard Deviation of Logarithm, Coefficient of Variation, Lorenz Curve and the Gini Coefficient for the years, 1995-96, 1991-92, 1988-89 and 1985-86.

Finally for the empirical analysis the set of data has been collected from BBS (Bangladesh Buruae of Statistics) publication like Household Expenditure Survey[1].

Table of Content

List of Tables :

List of Figures :

1) Acknowledgement

The first and foremost words of thanks from me are for my reverend teacher Professor Chikayoshi Saeki who provided me for the higher study, at the Faculty of Economics, Department of Economic Engineering of Kyushu University in Japan.

I express my deep indebtness to my honorable teacher Dr. A.B.M.A. Sobhan Miah, Professsor, Department of Statistics, Jahangirnagar University, Savar, Dhaka, Bangladesh, for his vigilant encouragement during the entire phase of the work.

I also acknowledge my gratefulness to Mr. Shougo Fukui who is a Master's student at the Faculty of Economics, Department of Economic Engineering of Kyushu University, Japan, for his hearty and sincere co-operations.

I like to express my thanks to my friend Most. Noor-E-Shahreen Tuly for her valuable advice.

At last I gratefully acknowledge the financial support of (MONBUSHO) Japanese Government for a research at the Faculty of Economics, Department of Economic Engineering, Kyushu University, Japan.

1 Introduction

The most attractive way to characterize an empirical distribution is by means of a small number of parameters. Preferably, these parameters should have a clear interpretation and define a theoretical distribution with a good fit to the empirical data. These statements hold for statistical distributions in general.

In past decades, many distributions have been developed as descriptive models for the size distribution of income. These includes, among others, the Lognormal, Log-t, Gamma, Beta, Singh-Maddala, Pareto and Weibull distributions.

The best-known families of theoretical distributions, used so far in the field of income distributions, are the Pareto, Lognormal and the Gamma families (see, Cramer (1969), Aitchison and Brown (1957), and Salem and Mount (1974). These families, however different in other respects, have in common that their members are characterized by two parameters. The interpretation of these parameters does not present great problems, but as regards fit, their performance is rather poor, which is probably due to oversimplification. Moreprecisely, members of the Pareto family very well fit the higher incomes. Members of the Lognormal family perform well in the central part of the distribution but their tails are not heavy enough to fit empirical data well; compare Aitchison and Brown (1957) and Cramer (1969). The weak points of the Gamma family are similar to that of the Lognormal family. In many applications, the Singh-Maddla distribution provides a better fit than the Gamma which performs much better than the Lognormal (McDonald and Ransom, Salem and Mount and Singh and Maddala). Thurow adopted the Beta distribution as a model for the distribution of income and this model includes the Gamma as limiting case; hence, the Beta distribution will provide at least as good a fit as the Gamma. The Singh-Maddala distribution includes the Weibull and Fisk distributions as special cases.

Recently, the Generalized Gamma has been used by Atoda, Suruga and Tachibanaki, Esteban, Kloek and Van Dijk and Taille. Esteban demonstrates that the Generalized Gamma has similar tail behavior or includes the Lognormal, Weibull, Gamma, Exponential, Normal and Pareto distributions as special or limiting cases. However the Beta, Singh-Maddala and Fisk distributions are not included as members of this class of distributions.

Such models should provide a reasonably close approximation to the true distribution and their parameters should be simple to estimate and also to interpret in an economically meaningful way. In recent years, the distributions which have been developed as descriptive models for the size distribution of income, are described below shortly ;

2 Some Important Functions for the Size Distribution of Income

2.1 Pareto Distribution

In general, the Pareto distribution of type three is given by,

$$F(x)=1-e^{-kx}(\alpha+\beta x)^{-a} \tag{1}$$

where, α, β, k and a are the parameters.

If we put, $k=0$, then the it is called the Pareto distribution of type two, which is given by,

$$F(x)=1-(\alpha-\beta x)^{-a} \tag{2}$$

where, α, β, and a are the parameters.

If we put, $\alpha=0$, and $\beta=\frac{1}{x_0}$, then it is called the Pareto distribution of type one, which is given by,

$$F(x)=1-\left(\frac{x}{x_0}\right)^{-a} ; x>x_0 \tag{3}$$

where, "a" is the parameters and x_0 is the minimum income.

2.2 Lognormal Distribution

If the logarithm of income is normally distributed with mean μ and variance σ^2, then we can say that income variable x is lognormally distributed with mean μ and variance σ^2, and we can write that,

$$x \sim LN(\mu, \sigma^2) \Longleftrightarrow , \log x \sim N(\mu, \sigma^2) \tag{4}$$

The frequency density function of the Lognormal distribution is given by,

$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}.x} e^{-\frac{(\log x-\mu)^2}{2\sigma^2}} ; x>0 \tag{5}$$

where, μ and σ^2 are the parameters.

2.3 Gamma Distribution

In general the p.d.f. of the Gamma distribution is given by,

$$f(y)=\frac{e^{-y}y^{a-1}}{\Gamma a} ; y>0 \tag{6}$$

If we put, $y=bx$, then the p.d.f. of the Gamma distribution is given by,

$$f(x)=\frac{b^a}{\Gamma a} e^{-xb}x^{a-1} ; x>0 \tag{7}$$

where, a and b are the positive parameters and Γa is the gamma function, which is given by,

$$\Gamma a=\int_0^\infty e^{-x}x^{a-1}dx \tag{8}$$

This distribution is proposed by Salem and Mount in 1974.

2.4 Log-t Distribution

If we put, $t=\log x$, (where x is income) to a general Student density, then the resulting density is called the log-Student or log-t distribution, which is given by,

$$f(x)=\frac{v^{v/2}}{\beta(1/2,\, v/2)cx}\left[v+\frac{(\log(x)-\log(m))^2}{c^2}\right]^{-(v+1)/2};\, x>0 \tag{9}$$

where, m, c and v are the parameters ($m>0, v>0, c>0$) and β is a beta funnction. This distribution is modified by Kloek and Van Dijk in 1978.

2.5 Generalized Gamma Distribution

The p.d.f. of the generalized gamma distribution is given by,

$$f(x)=\frac{ax^{ac-1}e^{-(x/b)^a}}{b^{ac}\Gamma c};\, x>0 \tag{10}$$

where, a, b and c are the parameters and Γ is a gamma function. If we put $a=1$, then the generalized gamma distribution is called the Salem and Mount Gamma distribution. If we put, $a=-1$ or $a=-2$, then it is called the inverese -1 Gamma or inverse -2 Gamma distribution.

2.6 Singh-Maddala Distribution

The two functions most often used are the Pareto function and the Lognormal. The Pareto function fits the data fairly well toward the higher income levels but the fit is poor toward the lower income levels. If one consider the entire range of income perhaps the fit may be the better for the Lognormal but the fit toward the upper end is far from satisfactory. Therefore, Singh and Maddala (1976) has derived a function based on the concept of hazard rate or failure rate which has been widely used for deriving distributions in reliability theory and for the analysis of the distribution of life times. The function is given by,

$$f(x)=\frac{abcx^{(b-1)}}{(1+ax^b)^{c+1}} \tag{11}$$

where, a, b and c are the parameters.

2.7 Generalized Beta Distribution

The p.d.f. of the generalized Beta distribution of first and second kind are defined as,

GB1

$$f(x)=\frac{ax^{(ap-1)}(1-(x/b)^a)^{(q-1)}}{b^{ap}\beta(p,\, q)};\, 0\le x\le b \tag{12}$$

GB2

$$f(x)=\frac{ax^{(ap-1)}}{b^{ap}\beta(p,q)((1+(x/b)^{a})^{(q-1)}};x>0 \tag{13}$$

where, a, b, p and q are the parameters.

These function are derived by, McDONALD in 1984. These distributions can be shown to include the Beta of the first kind (GB1) considered by Thurow, the Beta of the second kind (GB2), the Singh-Maddala, the Lognormal, Gamma, Weibull, Fisk and Exponential distributions as special or limiting cases.

Pareto function fits the data very well for the higher income classes. But, Bangladesh is a less developed country in the world. That is why, for the empirical analysis Pareto function is not applicable in case of Bangladesh. Therefore, for the empirical analysis we did not consider the Pareto function.

Here, we considered the Lognormal, Gamma, Log-t, Generalized Gamma, Singh-Maddala, and Generalized Beta distributions. All of these distributions do not fit the data of the household income distribution of Bangladesh very well. That is why, the main objective of this paper, is to select an appropriate distribution among these that fits the data of the household income distribution of Bangladesh very well. And another objective of this paper, is to calculate some important measures of income inequality on the basis of the appropriate distribution

To select the appropriate function that fits the data very well, next we will do the empirical analysis.

3 Empirical Analysis to Select an Appropriate Function as a Descriptive Model for the Household Income Distribution of Bangladesh

The principle objective of this work is to select an appropriate distribution among Lognormal, Gamma, Log-t, Generalized Gamma, Singh-Maddala, and Generalized Beta distributions, that fits the distribution of household income of Bangladesh very well. And on the basis of the appropriate distribution, we calculate some important measures of income inequality. That is why, we have to need the empirical analysis on the basis of the real data set. So, we have to collect a real data set for the empirical analysis. Next we describe about the data collection

3.1 Data Collection

3.1.1 Introduction

For the empirical analysis of the size distribution of household income of Bangladesh we have to collect a set of data. Data have been obtained on the variables, Monthly Household Income Group in Taka, Number of Households corresponding to the income group, Average Monthly Income Per Household in Taka corresponding to the income group. The variable x_i representing the average monthly income per household corresponding to the ith group, n_i representing the number of household corresponding to the ith group. On the basis of the x_i and n_i we will estimate the different functions for the household income distribution of Bangladesh for different years.

Next we will discuss about the Nature and Source of Data for the empirical analysis.

3.1.2 The Nature and Source of Data for Empirical Analysis

The success of any statistical and econometrical analysis ultimately depends on the availability of the appropriate data. It is therefore essential that we spend some time discussing the nature, sources, and limitations of the data that may arise in empirical analysis. The difficulty lies in the availability and nature of the data.

A particular problem facing the researcher is to obtain the appropriate data. In Bangladesh a reliable data is a golden deer. It is often difficult to obtain good reliable data with the necessary information required for a particular analysis. Though data are not available for statistical or econometrical research in Bangladesh we start with as best as we have. Missing value is a great problem in some data. Even in some experimentally collected data, errors of measurement arise from approximation and rounding off.

Because of all these and many other problems, the researchers should always keep in mind that the results of research may be affected by the quality of the data. Therefore, if in given situations the researchers find that the results of the research are "unsatisfactory" the case may be not that they used the wrong model but the quality of the data was poor.

Some difficulties in finding suitable data are frequently encountered. Income groups are perheps best indicated by equal income groups, but we also used unequal income groups. The measurement of income presents particular difficulties, because there are a lot of wrong informations.

For this study data were collected from BBS (Bangladesh Bureau of Statistics) publication "Household Expenditure Survey". For empirical analysis data are collected for the periods 1995-96, 1991-92, 1988-89 and 1985-86.

It is to be noted that, however, for printing mistake in the publications highly accurate results may not be obtained.

On the basis of the given sample observations for the years 1995-96, 1991-92, 1988-89 and 1985-86 we estimate the different distributions. Next, we will discuss about the estimation technique.

3.2 Estimation Technique

For the purpose of comparing the distributions, that fits the size distribution of household income of Bangladesh very well, we considered the sample data for the years 1995-96, 1991-92, 1988-89 and 1985-86. Let, x_i be the average monthly income per household corresponding to the ith household income group, and $f(x_i, \theta)$ is the probability of the x_i observation, where ($i=1$, 2.........n) and θ is the parameter space.

The likelihood function is given by ;

$$L(\theta)=f(x_1, \theta)f(x_2, \theta), \ldots\ldots\ldots f(x_n, \theta) \tag{14}$$

The log of the likelihood function is given by,

$$\text{Log } L(\theta)=\sum_{i=1}^{n}\log f(x_i, \theta) \tag{15}$$

Then the estimators are obtained by maximizing the likelihood or log likelihood function. The first stage of our numerical work consisted of estimating the Lognormal and Gamma families by means of the maximum likelihood method. We estimated the parameters value, sum of squared errors criteria, sum of absolute errors criteria and the chi-square value. These estimated results are reported with the table no. 1 and 2 for the years 1995-96, 1991-92, 1988-89 and 1985-86.

In the second stage of our nemerical work, we considered three and four parameters families like Log-t, Generalized Gamma, Singh-Maddala and Generalized Beta distributions. The estiamtors are obtained by using the non-linear maximization procedure (by using MATHEMATICA). These results of this estimation for the years 1995-96, 1991-92, 1988-89 and 1985-86 are reported with the table no. 3, 4, 5, and 6. For comparison the distributions we calculated the sum of squared errors criteria, absolute errors criteria and the chi-square value[2]. Also the estimated parameters value of the different distributions are also presented with these tables ;

(1) Estimated Parameters Value, Sum of Squared Errors Criteria, Absolute Errors Criteria and the χ^2 Value of the Log-Normal Distribution

Table No. 1

Year	Parameters Value		SSE Criteria, SAE Criteria and the χ^2 Value		
	μ	σ^2	SSE	SAE	χ^2 Value
1995-96	1.50179	1.10431	0.042034	0.647158	10092.6
1991-92	1.49288	1.06453	0.0344461	0.647654	6967.87
1988-89	1.25056	0.905588	0.0520946	0.766331	7879.99
1985-86	0.870203	0.968523	0.182391	1.15695	14848.3

(2) Estimated Parameters Value, Sum of Squared Errors Criteria, Absolute Errors Criteria and the χ^2 Value of the Gamma Distribution

2) The SSE, SAE and χ^2 values are obtained by evaluating,

$$\sum_{i=1}^{n}\left(\frac{n_i}{N}-f_i\theta\right)^2\sum_{i=1}^{n}\left|\left(\frac{n_i}{N}-f_i(\theta)\right)\right| \text{and} \sum_{i=1}^{n}\frac{(n_i-Nf_i(\theta))^2}{Nf_i(\theta)}, \text{ where } N=\Sigma_{i=1}^{n}n_i.$$

Table No. 2

Year	Parameters Value		SSE Criteria, SAE Criteria and the χ^2 Value		
	a	b	SSE	SAE	χ^2 Value
1995-96	1.13754	0.15366	0.0271964	0.558406	10023.4
1991-92	1.20225	0.168795	0.0258679	0.607981	6644
1988-89	1.37095	0.261282	0.0348798	0.692285	7232.93
1985-86	1.25988	0.337602	0.10574	0.9445	12193.1

(3) Estimated Parameters Value, Sum of Squared Errors Criteria, Absolute Errors Criteria and the χ^2 Value of the Log-t Distribution

Table No. 3

Year	Parameters Value			SSE Criteria, SAE Criteria and the χ^2 Value		
	v	c	m	SSE	SAE	χ^2 Value
1995-96	42.9259	1.03587	4.49593	0.040486	0.640189	9963.36
1991-92	74.5118	1.02387	4.46493	0.0335288	0.64338	6876.54
1988-89	48.502	0.940739	3.49508	0.0507215	0.759212	7776.94
1985-86	48.1099	0.972867	2.38776	0.178223	1.14833	14699.8

(4) Estimated Parameters Value, Sum of Squared Errors Criteria, Absolute Errors Criteria and the χ^2 Value of the Generalized Gamma Distribution

Table No. 4

Year	Parameters Value			SSE Criteria, SAE Criteria and the χ^2 Value		
	a	b	p	SSE	SAE	χ^2 Value
1995-96	0.365629	0.0260321	7.06416	0.0347716	0.584319	9740.07
1991-92	0.4454	0.145237	5.08121	0.0290334	0.621705	6674.66
1988-89	0.427564	0.054753	6.40255	0.0423271	0.722679	7462.32
1985-86	0.409464	0.0314422	6.37555b	0.14504	1.05605	13566.5

(5) Estimated Parameters Value, Sum of Squared Errors Criteria, Absolute Errors Criteria and the χ^2 Value of the Singh and Maddala Distribution

Table No. 5

Year	Parameters Value			SSE Criteria, SAE Criteria and the χ^2 Value		
	a	b	c	SSE	SAE	χ^2 Value
1995-96	0.0382132	1.27226	2.70879	0.0299582	0.55912	9478.68
1991-92	0.0286295	1.25516	3.54956	0.0266904	0.608633	6540.99
1988-89	0.0272355	1.32575	4.49741	0.0359595	0.691986	7169.02
1985-86	0.0870002	1.36997	2.48415	0.122767	0.991541	12730.5

(6) Estimated Parameters Value, Sum of Squared Errors Criteria, Absolute Errors Criteria and the χ^2 Value of the Second Kind of Beta Distribution

Table No. 6

Year	Parameters Value				SSE Criteria, SAE Criteria and the χ^2 Value		
	a	b	p	q	SSE	SAE	χ^2 Value
1995-96	0.3227	17.7459	14.7685	22.7536	0.0400459	0.635259	10016.4
1991-92	0.3877	23.7754	9.9688	18.6288	0.0320256	0.636655	6803.34
1988-89	0.3975	29.6034	10.4711	23.8295	0.0474501	0.740008	7668.64
1985-86	0.3717	8.9439	12.4612	20.0609	0.172087	1.13477	14528.6

where,
SSE : Sum of Squared Errors Criterion
SAE : Sum of Absolute Errors Criterion
Next we will discuss about the empirical results.

3.3 Empirical Results

Now, from the estimated results of the sum of squared errors criteria, absolute errors criteria, chi-square value and the parameters value we cam make the following comments.

(1) From the estimated results of the sum of squared errors criteria, absolute errors criteria we can say that Lognormal and Log-t distributions do not fit the real data very well than other distributions.

(2) From the estimated results of the sum of squared errors criteria, and absolute errors criteria of the Generalized Beta distribution we can say that this distribution fits the data better than Lognormal and Log-t distributions but it less fits the data than Gamma, Generalized Gamma and Singh-Maddala distribution.

(3) From the estimated results of the of sum of squared errors criteria, absolute errors criteria of the Gamma, Generalized Gamma and Singh-Maddala distributions we can say that the value of the sum of suared errors criteria and absolute errors criteria of the Gamma distribution is smaller than others. Therefore we can conclude that the Gamma distribution fits the data better than Singh-Maddala and Generalized Gamma distributions.

(4) Finally on the basis of these criteria it can be concluded that as a descriptive model for the household income distribution of Bangladesh, the Gamma function is more appropriate than other.

(5) From the chi-square value it is little bit difficult for us to find which distribution fits very well for the size distribution of household income of Bangladesh. But, from the chi-square value roughly we can say that the results of the Gamma and Sing-Maddala distributions are better than other distributions.

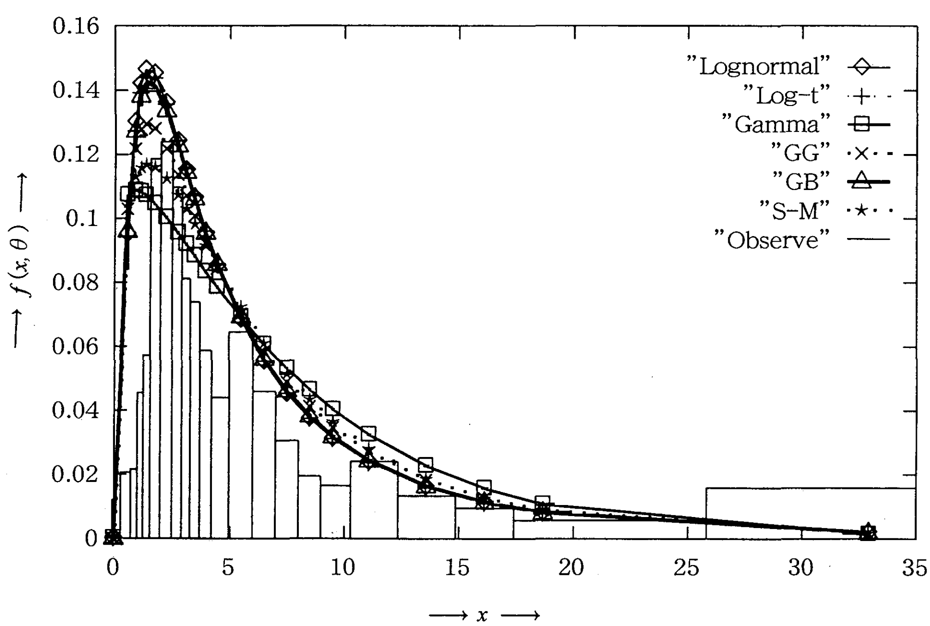
(6) From the estimated parameters value of the different distributions we can conclude that there exists a changing pattern of the individual household income in Bangladesh from year to year.

In general to get the clear idea of the distribution that fits the data very well, it is better to represent the predicted probabilities of the different distributions graphically. That is why, we also calculated the value of $f(x, \theta)$ of the different distributions for different years.

Next, to get the clear idea about different distributions, the observed and these predicted probabilites of the different distributions for the years 1995-96, 1991-92, 1988-89 and 1985-86 are also presented with the figure 1, 2, 3 and 4. We ploted the income variable x (in thousand), along the X-axis and corresponding probability $f(x, \theta)$, along the Y-axis.

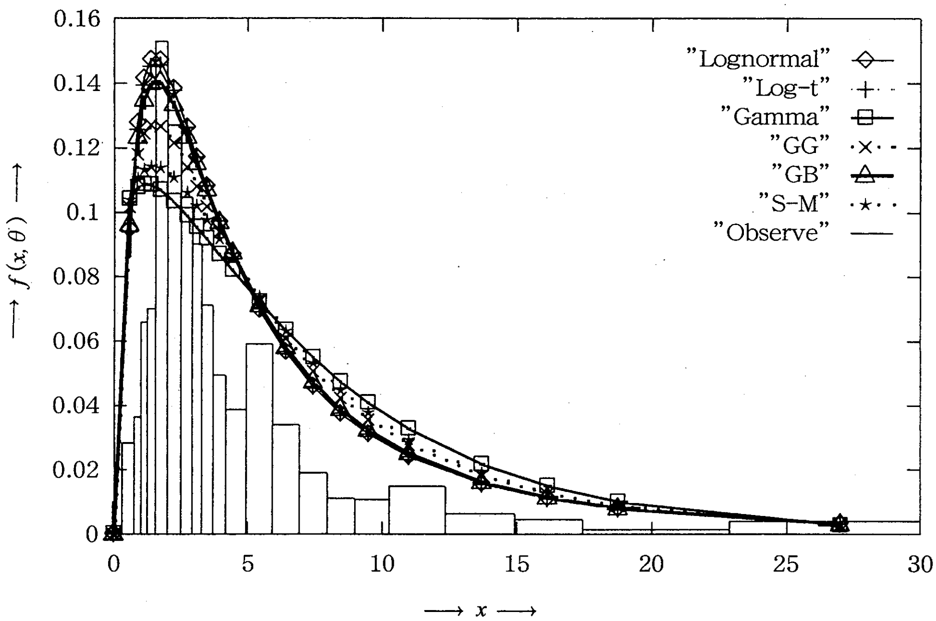
(1) Graphical Representation of the Observed and Predicted Probabilities of the Different Distributions for the Year 1995-96

Figure No. 1



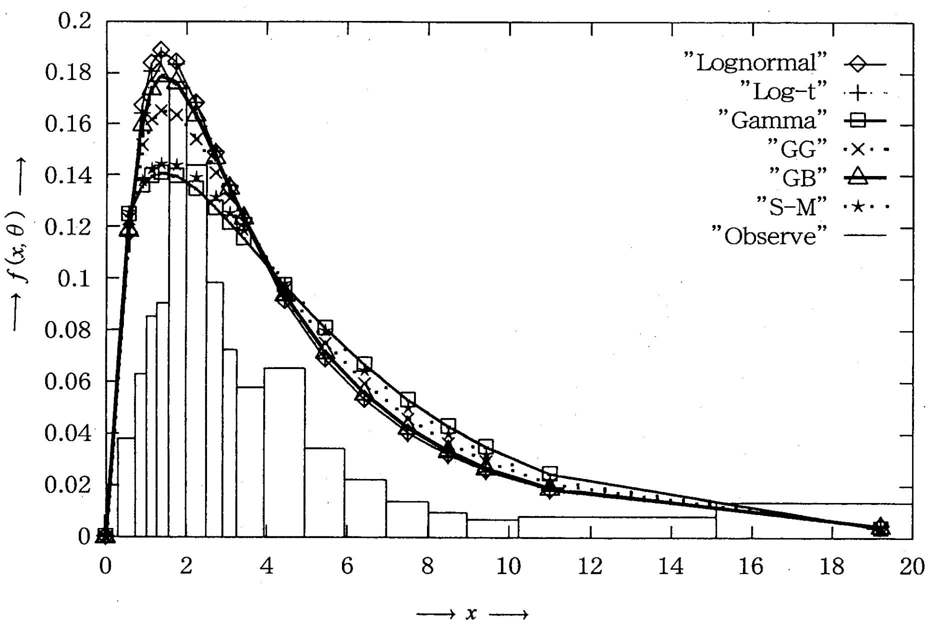
(2) Graphical Representation of the Observed and Predicted Probabilities of the Different Distributions for the Year 1991-92

Figure No. 2



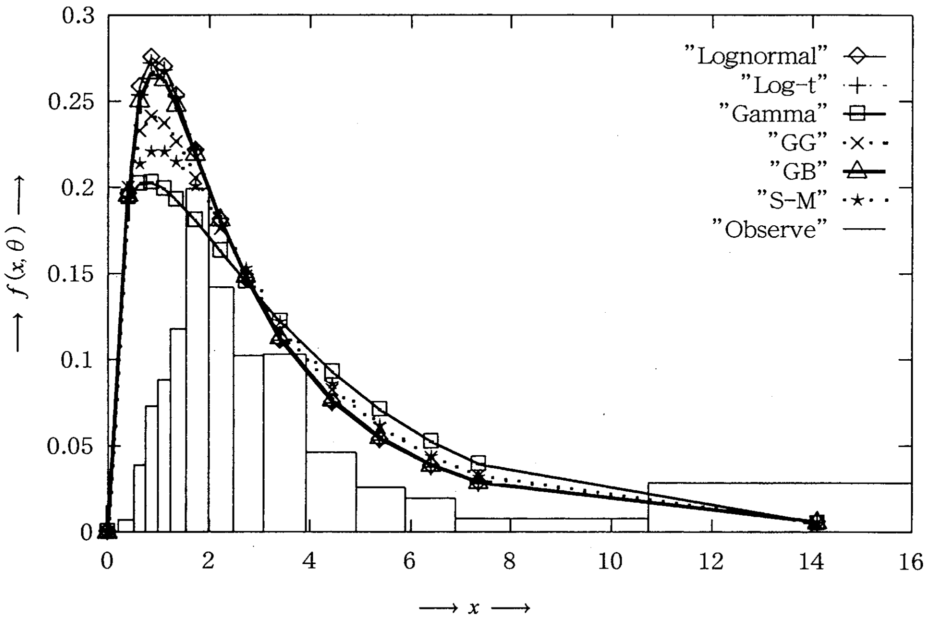
(3) Graphical Representation of the Observed and Predicted Probabilities of the Different Distributions for the Year 1988-89

Figure No. 3



(4) Graphical Representation of the Observed and Predicted Probabilities of the Different Distributions for the Year 1985-86

Figure No. 4



where,

GG : indicates generalized gamma function
GB : indicates generalized beta function
S-M : indicates singh-maddala function

From the Graphical representation, we can make the following comments.

3.4 Results From the Graphical Representation

(1) A preliminary plot showed that there is no significant difference between Lognormal and Log-t distributions as descriptive models for the size distribution of household income of Bangladesh

(2) From the graphical representation we see that all of the functions fit the lower income groups very well. But Lognormal, Log-t and Generalized Beta distributions fit the lower income classes better than other.

(3) From the graphical representation we see that Singh-Maddala and Gamma distributions fit the middle income classes well than other distributions. But comparatively, the Gamma function fits the middle income classes better than Singh-Maddala function.

(4) Finally, from the above points, it can be concluded that as a descriptive model for the size distribution of household income of Bangladesh, Gamma function is better than other functions.

(5) Also from the graphical representation, we can easily understand that there is a significant change of the individual household income from years 1985-86 to 1988-89 and 1988-89 to 1991-92. But there is a small change of the individual household income from year 1991-92 to 1995-96.

(6) From the graphical representation it can be concluded that the middle income classes of the household income of Bangladesh are not increased significantly from year to year.

Next, on the basis of the Gamma function, we calculate some important measures of income inequality.

4 Some Important Measures of Income Inequality

Inequality is not a uni-dimentional homogenous concept. There are many aspects of inequality and one measure of it may for some purposes be more meaningful than the other. The type of index required must depend on the purpose for which the enquiry is made. Hence, a single measure of the distribution of income, which is statisticaly devisable is quite inadequate for discribing all dimensions that must included in any adequate analysis of income inequality.

Basically, measures of income inequality may be divided into two categories

(1) measures that describe inequality in an object sense employing some statistical measure

(2) in terms of some normative notion of social welfare such that is higher degree of inequality corresponds to a lower level of social welfare for a given total of income.

Some measures of inequality are discussed below ;

4.1 Relative Mean Deviation

This is the well known statistical measure of looking at the whole distribution by comparing the income level of each with the mean income. This measure is calculated by summing up the absolute values of all the differences of the income levels from the mean and then dividing it by the total income.

Although this measure takes into consideration the whole distribution of income, yet it is insensitive to transfer of income from a poorer to a richer person

4.2 Variance

This measure of inequality is arrived at by adding up the squares of the differences between different income levels and the mean income. This means that any transfer of income from a poorer to a richer person, other things remaining the same, always increase the varinace. However, this measure is also not independent of mean income level. Thus the mean income levels of different income distrbutions of income will obviously have its influence on variance.

4.3 Coefficient of Variation

It is a square root of the variance divided by the mean income level. Thus this measure of inequality is independent of mean income level. This measure is also sensitive to income transfers for all income levels but it attaches the same weight irrespective of the level of income from which these transfers take place.

4.4 Standard Deviation of Logarithm

Logarithmic transformation has several advantages over actual values in calculating standdard deviation, since it eliminates the arbitrariness of the units and therefore of the absolute levels, a change of unit which takes the form of a multiplication of the absolute values, comes out in the logarithmic form as an addition of a constant, and therefore when pairwise differences are taken they vanish.

This measure gives greater weight to income transfers at the lower levels of income. But this measure also depends on mean incomes as the differences only from the mean are taken.

4.5 Lorenz Curve

Lorenz curve shows the cumulative proportions of income received by cumulative proportions of receipents arranged in order from the poorest to the richest proportions of people. A Lorenz curve joins the two corners diametrically opposite to each other of a unit square.

Let $f(X, \theta)$ is the pdf of the sample observation X, where θ is the parameter space. The proportion of units having an income less than or equal to x is defined by $F(x)$ and is given by ;

$$F(x)=\int_0^x f(X, \theta)dx, 0\leq F(x)\leq 1$$

The proportional share of total income of the units having an income less than or equal to x is defined as $Q(x)$ and is given by ;

$$Q(x)=\frac{1}{\mu}\int_0^x Xf(X, \theta)dx, 0\leq Q(x)\leq 1$$

where, μ is the mean value of the distribution exists.

Then the Lorenz curve is the relationship between the variable $F(x)$ and $Q(x)$.

If the income distribution is perfectly equal then the relationship between $F(x)$ and $Q(x)$ will be diagonal. If there is any inequality then the bottom income groups will enjoy a proportional lower share of income. Thus it is obvious that any Lorenz curve must lie below the diagonal unless the distribution is perfectly equal and its slope increasingly rise at any rate not fall, as we come to richer and richer sections of the population. If the Lorenz curves of the two distributions of income intersect each other than we can not say unamgiously as to which distribution is more unequal. If the Lorenz curves do not intersect, then we can say unambigiously that the distribution closer to the diagonal is more equal than the other.

4.6 Gini-coefficient

Gini-Coefficient is derived from the Lorenz curve is probably the most widely used of inequality of income. The Gini-Coefficient is the ratio of difference between the line of absolutely equality (the diagonal) and the Lorenz curve. It is clear that the value of Gini-Coefficient will lie between 0 and 1. For absolute equality of distribution, the diagonal and the Lorenz curve will coincide and as such the difference between them will be zero. Again for absolute inequality of income distribution the difference between the diagonal and the Lorenz curve will be the whole triangle will be unity. And intermediate distribution of income will lead to a value of Gini-Coefficient anything between zero and unity.

In taking differences over all pairs of incomes, the Gini-Coefficient avoids the total concentration on difference around the mean, which the variance, coeffcient of variation and the standdard deviation of logarithms do, and also it seems to be a more direct approach as it avoids the arbitrary squaring procedure of the rest of the others mentioned above. This measure is sensitive to transfers from the richer to the poorer sections of the population at every level and gives greater weight to such transfers at the middle income groups.

Next we will discuss about the empirical analysis of some important inequality measures.

5 Empirical Analysis of Some Important of income inequality Measures

From the empirical analysis of the household income distribution of Bangladesh, it has been found that as a descriptive model for the size distribution of the household income of Bangladesh, the Gamma distribution is more appropriate. Now on the basis of the Gamma function we calculated some important measures of income inequality of Bangladesh. The estimated values of the income inequality measures are also reported with the following table.

(1) Estimated Values of Some Important Measures of Income Inequality

Table No. 7

Year	RMD	Rank	SDL	Rank	CV	Rank	GC	Rank
1995-96	0.754506	4	1.05086	4	0.89127	4	0.4781	4
1991-92	0.736348	3	1.03176	3	0.853216	3	0.4415	3
1988-89	0.691869	1	0.951624	1	0.807434	1	0.4164	1
1985-86	0.729798	2	0.984136	2	0.840317	2	0.4371	2

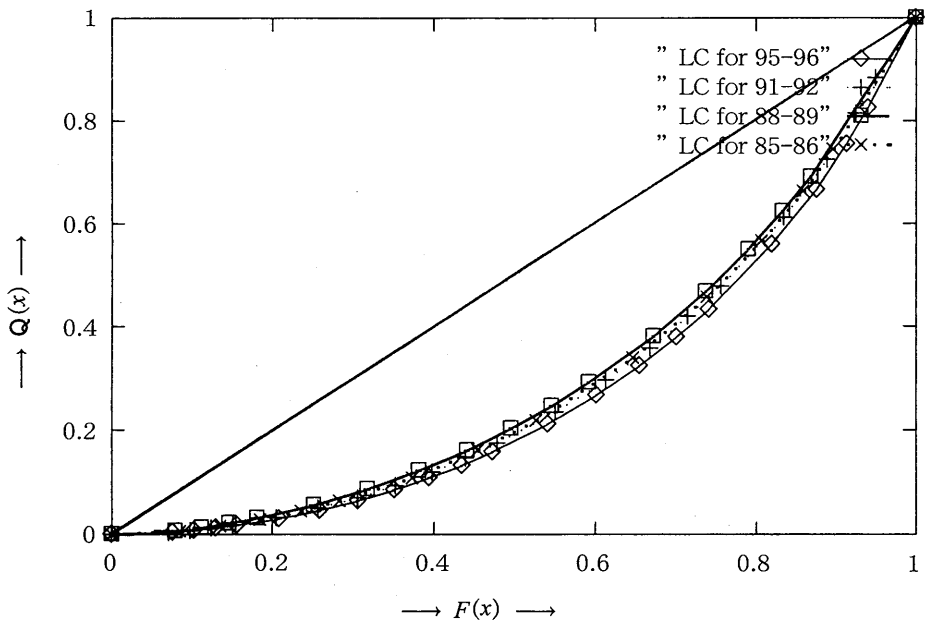
where,

RMD : indicates relative mean deviation ;
SDL : indicates standard deviation of logarithm
CV : indicates coefficient of variation ;
GC : indicates Gini coefficient

Also to get the clear idea of the inequality of the household income distribution of Bangladesh for the different years, we draw the Lorenz curves on the basis of the estimated values of $F(x)$ and $Q(x)$ for different years. We plot $F(x)$, along the X-axis and $Q(x)$, along the Y-axis.

(2) Lorenz Curves for the Years 1995-96, 1991-92, 1988-89 and 1985-86

Figure No. 5



LC : indicates Lorenz curve

Next we will discuss about the empirical results of some important inequality measures.

5.1 Empirical Results of Some Important Measures of Income Inequality

From the estimated values of the different measures, which are reported with the table no 7, it has been found that in 1995-96 the household income distribution of Bangladesh is more unequal with compared to the household income distribution of the years 1991-92, 1988-89 and 1985-86, but the distribution is less unequal in 1988-89. Also from the figure 5, it can be easily understood that the Lorenz curve for the year 1988-89 is closer to the diagonal line and for the year 1995-96 is further from the diagonal line. Thus from this graphical representation it can be concluded that the degree of inequality of the household income distribution of Bangladesh is higher for the year 1995-96 and was lower for the year 1988-89. Also from the estimated results it can be concluded that the income inequality of Bangladesh is increased from year to year. This means, that in Bangladesh the poorer are going to be more poorer and the richer are going to be more richer. Next we will pass for overall discussion and conslusion.

6 Discussion and Conclusion

In past decades different distributions have been developed as descriptive models for the size distribution of income, like Pareto, Lognormal, Gamma, Log-t, Generalized Gamma, Singh-Maddala, Beta, Gneralized Beta, Weibull and Exponential distributions. But in this paper for the numerical analysis, at the first stage we considered the two parameters models, like Lognormal and Gamma families. At the second stage we considered the three and four parameters models, like Log-t, Generalized Gamma, Singh-Maddala and Generalized Beta distributions. Any model of income distribution involves some compromise between how well the model fits real data and how easily the parameters can be estimated and interpreted. The two parameters of Gamma density can be directly related to measure of inequality and of proportionate growth and both of these concepts have obvious relevance for economists. All of these consideded distributions do not fit the data of the household income distribution of Bangladesh very well.

That is why, the principle purpose of this paper is to select an appropriate distribution that fits the data of the household income of Bangladesh very well than other distributions. And another attempt has been made to calculate some important measures of income inequality of Bangladesh on the basis of the appropriate function.That is why we have done empirical analysis on the basis of the real data set. A set of data has been collected for the years 1995-96, 1991-92, 1988-89 and 1985-86 from the BBS (Bangladesh Bureau of Statistics) publication like Household Expenditure Survey. For comparing the relative performance of the different distributions we estimated the sum of squared errors criteria, sum of absolute errors criteria and the chi-square value of the different distributions for the years 1995-96, 1991-92, 1988-89 and 1985-86. For the two parameters models we used the maximum likelihood method and for the three and four parameters models we used the non-linear maximum likeihhod method. The results are obtained by using MATHEMATICA program. The estimated results of the different distributions are also reported with the table no. 1, 2, 3, 4, 5 and 6. From the estimated results of sum of squared errors criteria, absolute errors criteria we found that the Gamma function fits the data of the household income distribution of Bangladesh very well than other distributions. Also we estimated the parameters value of the different functions for different years. From the estimated parameters value of the different distributions, it can be concluded that there is a significant change of the individual household income from year 1985-86 to 1988-89 and 1988-89 to 1991-92, but from year 1991-92 to 1995-96 there is no significant change.

To get the clear idea of the different distributions we also presented the observed and predicted probabilies graphically for the years 1995-96, 1991-92, 1988-89 and 1985-86. From the graphical representation we see that there is no significant difference of the Lognormal and Log-t distributions as descriptive models for the size distribution of household income of Bangladesh. Also from the graphical representation we see that all of the distributions fit the lower income classes very well. But Singh-Maddala and Gamma distributions fit the middle income classes better than other functions. But Gamma function fits the middle income classes better than Singh-Maddala function. That is why, finally it can be concluded that as a descriptive model for the size distribution of the household income of Bangladesh Gamma function is more appropriate than other functions. Also from the graphical representation we see that there is a significant change of the individual household income in Bangladesh from year 1985-86 to 1988-89 and 1988-89 to 1991-92. But there is a small change of the individual household income from year 1991-92 to 1995-96.

From the graphical represenation finally it can be concluded that the middle household income classes are not increased significantly from year to year in Bangladesh. Next, on the basis of the Gamma function we estimated some important measures of income inequality. These estimated results are reported with the table no 7. From the estimated results of the different measures it has been found that, the degree of inequality of the household income distribution of Bangladesh is higher for the year 1995-96, and was lower for the year 1988-89. Also to get the clear idea, we have drawn Lorenz curves on the basis of the estimated results for different years. From the Lorenz curves, it has been found that the degree of household income inequality is higher for the year 1995-96 and was lower for the year 1988-89. From the estimated results, it can be concluded that the inequality of the household income distribution of Bangladesh is increased from year to year. So, it can be concluded that in Bangladesh the poorer are going to be more poorer and the richer are going to be more richer.

Bibliography

1. Aigner (1970), "Estimation of Pareto's Law from Grouped Observation," *Journal of American Statistical Association*, Vol. 65.
2. Aitchison J. and J. A. C. Brown (1976), "The Lognormal Distribution," *Cambridge University Press*.
3. Cowell F. A. (1977), "Measuring Inequality," *Philip Allan*,
4. Dagum, C. (1983), "Income Distribution Models," *Encyclopedia of Statistical Science*, Vol. 4.
5. Gastworth, J. L. (1972), "The Estimation of the Lorenz Curve and Gini Index," *Review of Economics and Statistics*, Vol. 54.
6. Kloek, T. and H. K. Van Dijk (1978), "Efficient Estimation of Income Distribution Parameters," *Journal of Econometrics*, Vol. 8.
7. McDonald, J. B., (1984), "Some Generalized Function for the Size Distribution of Income," *Econometrica*, Vol. 52.

8. Metcalf, C. E. (1972), "An Econometric Model of the Income Distribution" *Markham Publishing Company*,

9. Salem, A. B. and T. D. Mount, (1974), "A Convenient Descriptive Model of Income Distribution ; The Gamma Density" *Econometrica*, Vol. 42.

10. Singh, S. K., and G. S. Maddala, (1976), "A Function For Size Distribution of Income," *Econometrica*, Vol. 44.