International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2015): 6.391

A Study on the Concentration of Urban Population in Tamilnadu State

Dr. I. Venkatesan

Assistant Professor of Statistics, Sri Balamurugan Arts & Science College, Sathappadi-636451

1. Introduction

Much attention has been paid recently to the distribution of sizes of cities. City is the one of the components of urban system. The growth of urban population is due to the effect modernization, economic development; social, cultural, political changes and government policy. When these forces are in operation, the path of urbanization may be stochastic in nature.

The growth of urbanization particularly the growth of city population may be studied using the certain growth models. Random growth model was applied to explain the size distribution of city (Simon 1955). Size distribution of city in an urban system can be closely approximated by Pareto distribution [Auerbach (1913) and Zipf (1949)]. The empirical distribution of city sizes are having in skew nature rank size rule, pareto model are used to study the path of the city process. Changes in the size distribution of city are related to the differential growth. The urban system, analyzed by the various threshold sizes, examines urban growth rate.

2. Urbanization

The study of urbanization has certain importance all over the world and more so in the developing countries like India urbanization is the process by which villages turn into towns and towns develop into cities. A town may not be a resources but urban life is often associated with economic advances.

A more precise definition of 'Urban' was introduced for the first time by the census commissioner of India in 1981. According to the definition, the 'Urban places' is defined as "all other places which satisfy the following criteria. All places with a municipal area, corporation cantonment notified area committee". This definition of an urban area is arbitrary and is not the same as used in other countries. The urbanization process in the India has the following features in terms of the percentage of the total population the level of urbanization is low. In terms of absolute numbers, the level of urbanization is rather high in comparison with most of the developing countries and many developed countries. The rank of population growth was higher in urban areas than the rural areas in each decade.

Urban social structure reflects the attitudes, values and norms of the peasant society of rural India. When the definition of urban places is given in terms of the criteria of urbanization, it involves an inherent circularity, (E.g) an urban trait is one found in cities and a city is a place characterized by urban traits. The term urbanism denotes a distinct quality of human community, a special mode of existence or way of life which is the characteristic of the city. Urbanization is a process of population, concentration, diffusion of the influence of urban centers of rural hinterland and increases in the share of working force engaged in non-agricultural occupations.

3. Models

Model is a simplification of reality. A model is an abstract representation of reality which clears what is relevant to a particular question at as particular time and neglects all other aspects. A model establishes the main variables involved and connects them by means of mathematical statements. A model involving a random variable or chance factor is called as a stochastic model or a probabilistic model.

Urban Population Models

To study the path of urbanization in any urban areas, certain city size distributions are used namely Pareto city size distribution; rank size model and exponential distribution are used to describe the skew distribution. These models may be viewed as urban population models.

Exponential Model

For most urban areas the distribution of population appears to confirm approximately to the negative exponential distribution. This can be expressed as follows,

$$D(u) = Ae^{-b(u)}$$

Where $D(u)$ represents the density of population at distance 'i.e.' from the centre. 'A' represents the density of the city centre. 'b' represents the rate at which population declines per unit of distance from the city centre and 'r' represents the radius of the city.

The value of 'b' the slope parameter of the exponential distribution, represents a relative measure of compactness of the city or its degree of decentralization. By contrast the value of 'A' is an absolute measure indicating maximum level of density (ie) density at the level of the city. It will be seen later that within a national system of cities for a given points in time. A can be regarded as on index of the overall mass of the urban population.

Pareto Model

The Pareto distribution is named after an Italian – born Swish Professor of Economics, Vilified Pareto (1843 – 1923). Pareto law deals with distribution of income over population and can be states as follows,

$$N = Ax^{-a}$$

Where N is the number of persons having income x greater than A, Pareto distribution is one of the skew distribution is used to describe the city sizes when the frequency distribution of city size is strongly skewed.

The study of urban composition of a country is important for a variety of reasons. The developed countries of today had undergone changes as was observed in the demographic. Transition we looked at the birth and death rate. Similarly change could also be observed in respect to the urban population in the developed countries increased substantially with economic development.

Objectives

To examine the spatial concentration by means of computing spatial Gini coefficient using statistical model for the Tamil Nadu city population data.

Gini Coefficient

The Gini coefficient is a measure of statistical dispersion most prominently used as a measure of inequality of income distribution or inequality of wealth distribution. It is defined as a ratio with values between 0 and 1: A low Gini coefficient indicates more equal income or wealth distribution, while a high Gini coefficient indicates more unequal distribution. 0 corresponds to perfect inequality (everyone having exactly the same income) and 1 corresponds to perfect inequality (where one person has the income, while everyone else has zero income).

The Gini coefficient is defined as a ratio of the areas on the Lorenz curve diagram. If the area between the line of perfect equality and Lorenz curve is A, and the area under the Lorenz curve is B, then the Gini coefficient is $A/(A+B)$. Since $A+B=0.5$, the Gini coefficient, $G=\frac{A}{0.5}=2A=1-2B$. If the Lorenz curve is represented by the function $Y=L(X)$, the value of B can be found with integrations and :

$$G=1-2\int_0^1 L(X)\,dX$$

Lorenz Curve

The Lorenz curve is a graphical representation of the cumulative distribution function of a probability distribution, it is a graph showing the preparation of the distribution assumed by the bottom y% of the values. It is often used to represent income distribution, where it shows for the bottom x% of households, what percentage y% of the total income they have. The percentage of income on the y – axis. It can also be used to show distribution of assets. In such case many economists consider it to be a measure of social inequality. It was developed by Max O.Lorenz in 1905 for representing income distribution.

Methods

Maximum Likelihood Estimation

Let $x_1,x_2,\dots,x_n$ be a random sample of size n from a population with density function $f(x,\theta)$. Then the likelihood function of the sample values $x_1,x_2,\dots,x_n$, usually denoted by $L=L(\theta:x)$ is their joint density function such as

$$L=f(x_1,\theta),f(x_2,\theta),\dots,f(x_n,\theta)=\prod_{i=1}^{n} f(x_i,\theta)$$

L gives the relative likelihood that the random variables assume a particular set of values $x_1,x_2,\dots,x_n$. For the given sample $x_1,x_2,\dots,x_n$ L becomes a function of the variable θ of the parameter. The likelihood function L is maximized by using the maxima, minima principle.

(ie) $\frac{\partial \log L}{\partial\theta}=0,\frac{\partial^2 \log L}{\partial\theta^2}<0$

For θ are the conditions for the function L attaining the maximum value. The equation $\frac{\partial \log L}{\partial\theta}=0$, is a likelihood equation. The solution of this equation is called as maximum likelihood estimate of θ.

Pareto Model

The urban size x is assumed to follow Pareto distribution $f(x,a)$ is described as

$$f(x,a)=\frac{ax^{*a}}{x^{a+1}},a>0,x>x^*$$

Where x^* is the threshold urban size. Parameters $a\ and\ x^*$ are estimated by using the method of maximum likelihood (Rao.C.R.1973).

The likelihood function of the Pareto distribution is

$$L(a:x)=\prod_{i=1}^{n}\frac{ax^{*a}}{x^{a+1}}$$

$$L(a:x)=\frac{a^n(x^{*a})^n}{\prod_{i=1}^{n}x^{a+1}}$$

$$\log L(a:x)=n\log a+na\log x^*-\sum_{i=1}^{n}(a+1)\log x$$

Differentiating this with respect to a,

$$\frac{\partial\log L(a:x)}{\partial a}=\frac{n}{a}+n\log x^*-\sum_{i=1}^{n}\log x$$

$$\frac{\partial\log L(a:x)}{\partial a}=0\implies n\log x^*-\sum_{i=1}^{n}\log x=-\frac{n}{a}$$

$$\sum_{i=1}^{n}\log x-n\log x^*=\frac{n}{a}$$

$$\frac{\sum_i\log x}{n}-\log x^*=\frac{1}{a}$$

$$x^*=Mini(x_i)_{i\le 1}$$

The expressions for p and q are obtained as follows,

$$p_x=F_x(x)$$

$$=\int_{x^*}^{x}\frac{ax^{*a}}{t^{a+1}}dt$$

$$p_x=1-(x^*/x)^a\ \dots(1)$$

$$q_x=\int_{x^*}^{x}tf(t,a)dt\bigg/\int_{x^*}^{\infty}tf(t,a)dt$$

$$q_x=\frac{l_1}{l_2}$$

$$Where,l_1=\int_{x^*}^{x}tf(t,a)dt$$

$$=\int_{x^*}^{x}\frac{ax^{*a}}{t^{a+1}}dt$$

$$= (a/a-1)\, x^*[(x^*/x)^{a-1} - 1]$$

$$Where, l_2 = \int_{x^*}^{\infty} t f(t,a) dt$$

$$= \int_{x^*}^{\infty} \frac{a x^{*a}}{t^{a+1}} dt$$

$$= (a/a-1)\,(-x^*)$$

$$q_x = l_1/l_2$$

$$= (a/a-1)\, x^*[(x^*/x)^{a-1} - 1]/(a/a-1)\,(-x^*)$$

$$q_x = [1-(x^*/x)^{a-1}] \ldots\ldots(2)$$

The function $q = f(p)$ has been obtained by using (1) and (2)

$$1 - p_x = (x^*/x)^a$$

$$\frac{x^*}{x} = (1-p_x)^{1/a} \ldots\ldots\ldots(3)$$

$$1 - q_x = (x^*/x)^{a-1}$$

$$\frac{x^*}{x} = (1-q_x)^{\frac{1}{a}-1} \ldots\ldots(4)$$

From (3) and (4),

Equ (3) = Equ (4) $\Rightarrow (1-q_x)^{\frac{1}{a}-1} = (1-p_x)^{1/a}$

$$(1-q_x) = (1-p_x)^{a-1/a}$$

Since the urban size distribution is known, $q = f(p)$ is obtained by eliminating x.

$$(ie)\; q = 1-(1-p)^{a-1/a}; 0 \le p \le 1$$

The graph of the function $q = f(p)$ is Lorenz curve representing the variation in the size of the urban areas. Gini's concentration ratio is stated as

$$\rho = \frac{Area\ between\ the\ Lorenz\ curve\ and\ the\ line\ of\ equal\ distribution}{Total\ area\ under\ diagonal}$$

= 1- 2A, Where A is the area under the Lorenz curve.

$$A = \int_0^1 F(p)dp$$

$$= \int_0^1 [1-(1-p)^{(a-1)/a}]dp$$

$$= \frac{a-1}{2a-1}$$

$$\rho = 1 - \frac{2(a-1)}{(2a-1)}$$

$$\rho = \frac{1}{2a-1}$$

Gini's concentration for urban population using Pareto model is estimated.

$$\hat{\rho} = \frac{1}{(2\hat{a}-1)}$$

4. Exponential Model

Let x be the urban size and it follows an exponential distribution $f(x:\lambda)$, where $f(x:\lambda)$ is described as

$$f(x:\lambda) = \lambda e^{-\lambda x}, x > 0$$

The parameter λ is estimated by using method of maximum likelihood.

The likelihood function of the exponential model is

$$L(\lambda:x) = \prod_{i=1}^{n} \lambda e^{-\lambda x}$$

$$L(\lambda:x) = \lambda^n e^{-\lambda \sum_i x_i}$$

$$\log L(\lambda:x) = n\log\lambda - \lambda\sum_i x_i$$

Differentiating this with respect to λ,

$$\frac{\partial \log L(\lambda:x)}{\partial\lambda} = \frac{n}{\lambda} - \sum_i x_i$$

$$\frac{\partial \log L}{\partial\lambda} = 0 \Rightarrow \frac{n}{\lambda} - \sum_i x_i = 0$$

$$n - \lambda\sum_i x_i = 0$$

$$\lambda\sum_i x_i = n$$

$$\lambda = \frac{n}{\sum_i x_i}$$

$$\hat{\lambda} = \frac{1}{\bar{x}}$$

Expressions for p and q are obtained as follows

$$p_x = F_x(x)$$

$$= \int_0^x \lambda e^{-\lambda t}\, dt$$

$$p_x = 1 - e^{-\lambda x} \ldots\ldots(5)$$

$$q_x = \int_0^x t\lambda e^{-\lambda t}\, dt \Big/ \int_0^\infty t\lambda e^{-\lambda t}\, dt$$

$$q_x = \frac{l_1}{l_2}$$

$$Where, l_1 = \int_0^x t\lambda e^{-\lambda t}\, dt$$

$$= \lambda\left[te^{-\lambda t}/(-\lambda)\right]_0^x + \frac{1}{\lambda}\int_0^x e^{-\lambda t}\, dt$$

$$= \frac{1}{\lambda} - \frac{1}{\lambda}e^{-\lambda t} - xe^{-\lambda t}$$

$$Where, l_2 = \int_0^\infty t\lambda e^{-\lambda t}\, dt$$

$$= \lambda\left[te^{-\lambda t}/(-\lambda)\right]_0^\infty + \frac{1}{\lambda}\int_0^\infty e^{-\lambda t}\, dt$$

$$= \frac{1}{\lambda}$$

$$q_x = l_1/l_2 = \left[\frac{1}{\lambda} - \frac{1}{\lambda}e^{-\lambda t} - xe^{-\lambda t}\right]/[1/\lambda]$$

$$q_x = 1 - e^{-\lambda x} - \lambda xe^{-\lambda x}$$

$$q_x = p_x - \lambda xe^{-\lambda x}$$

Since the distribution of x is known, $q = f(p)$ is obtained by eliminating x.

(ie) $q = p - \lambda e^{-\lambda}$

Gini's concentration ratio, based on the function $= p - \lambda e^{-\lambda}$, is obtained as follows.

(ie) ρ = 1- 2A, Where A is the area under the Lorenz curve.

$$A = \int_0^1 f(p)dp$$

$$= \int_0^1 [p - \lambda e^{-\lambda}]dp$$

$$A = (1/2) - \lambda e^{-\lambda}$$

$$\rho = 1 - 2[(1/2) - \lambda e^{-\lambda}]$$

$$\rho = 2\,\lambda e^{-\lambda}$$

Using Exponential model, Gini's concentration ratio for urban population was estimated as

$$\hat{\rho} = 2\,\hat{\lambda} e^{-\hat{\lambda}}$$

Year	Estimates of parameters (Pareto Model)		Estimation of Gini's Concentration Ratio
	$\hat{a}$	x^*	$\hat{\rho}$
1981	4.8852	125000	0.1140
1991	4.8804	125000	0.1141
2001	3.4722	125000	0.1682
2011	2.3725	125000	0.2670

Year	Estimates of parameters $\hat{\lambda}$ (Exponential Model)	Estimation of Gini's Concentration Ratio $\hat{\rho}$
1981	0.0000040	0.0000080
1991	0.0000042	0.0000084
2001	0.0000033	0.0000066
2011	0.0000025	0.0000050

Lorenz Curve using Pareto Model (1981)

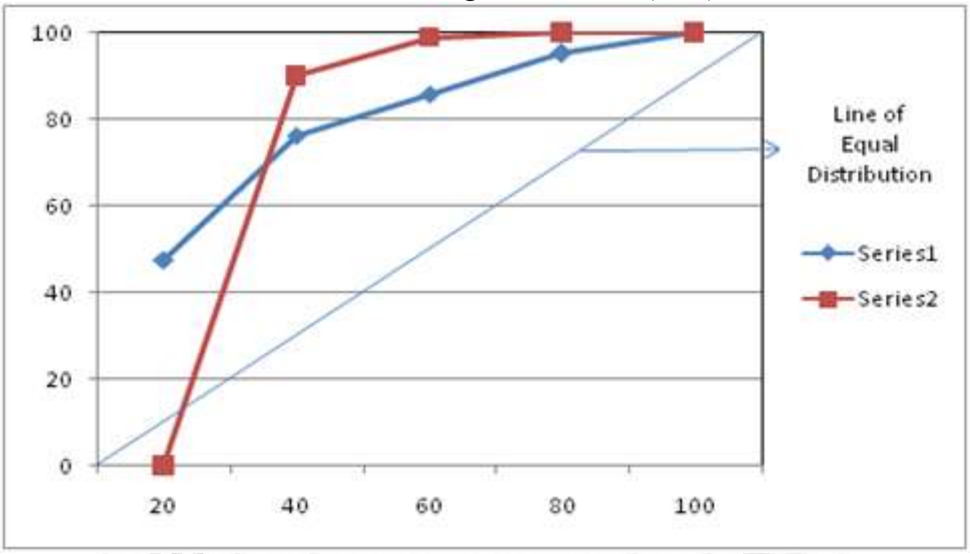

Figure 1

X-Axis = City Size **Y-Axis = No. of Cities**

Series1 = Emperical Curve **Series2 = Estimated Curve**

Lorenz Curve using Pareto Model (1991)

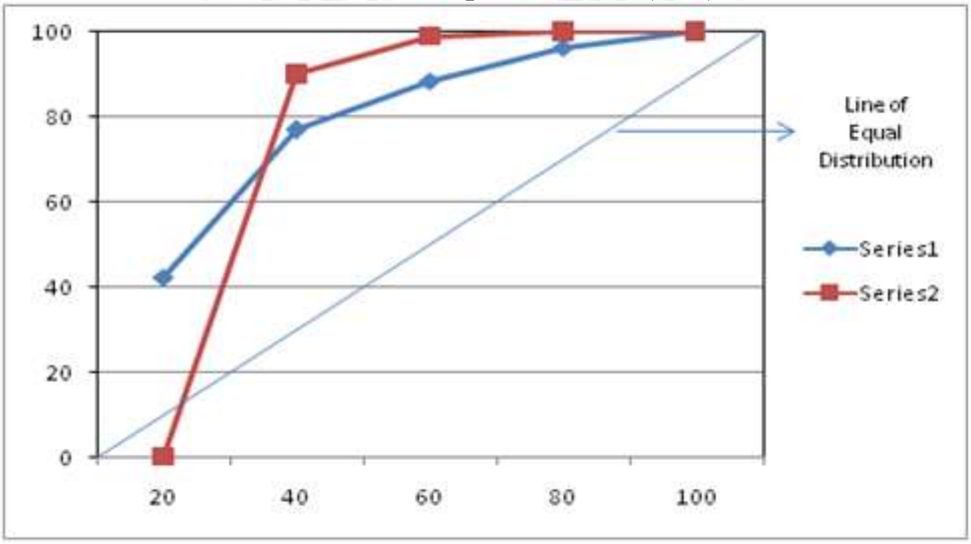

Figure 2

X-Axis = City Size **Y-Axis = No. of Cities**

Series1 = Emperical Curve **Series2 = Estimated Curve**

International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2015): 6.391

Lorenz Curve using Pareto Model (2001)

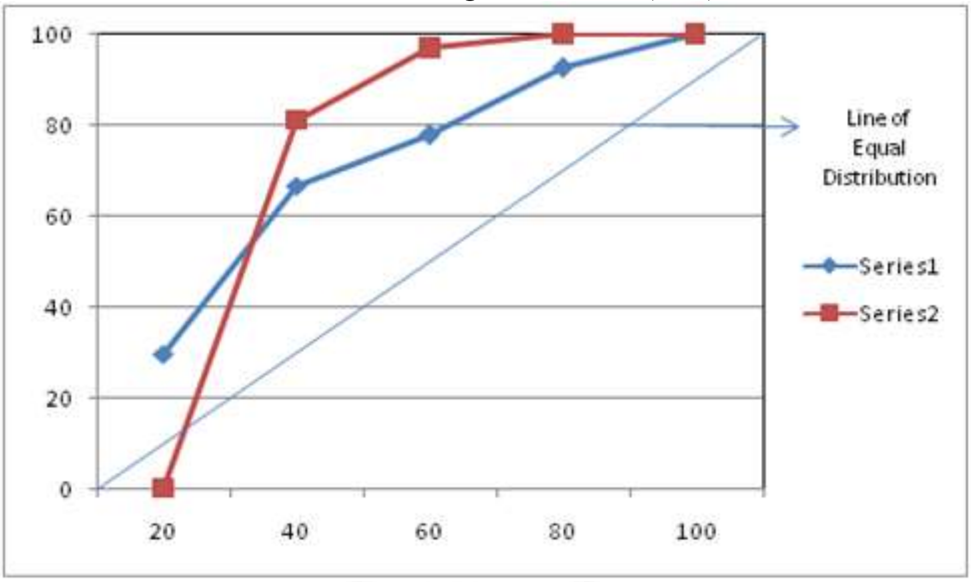

Figure 3

X-Axis = City Size **Y-Axis = No. of Cities**
Series1 = Emperical Curve **Series2 = Estimated Curve**

Lorenz Curve using Pareto Model (2011)

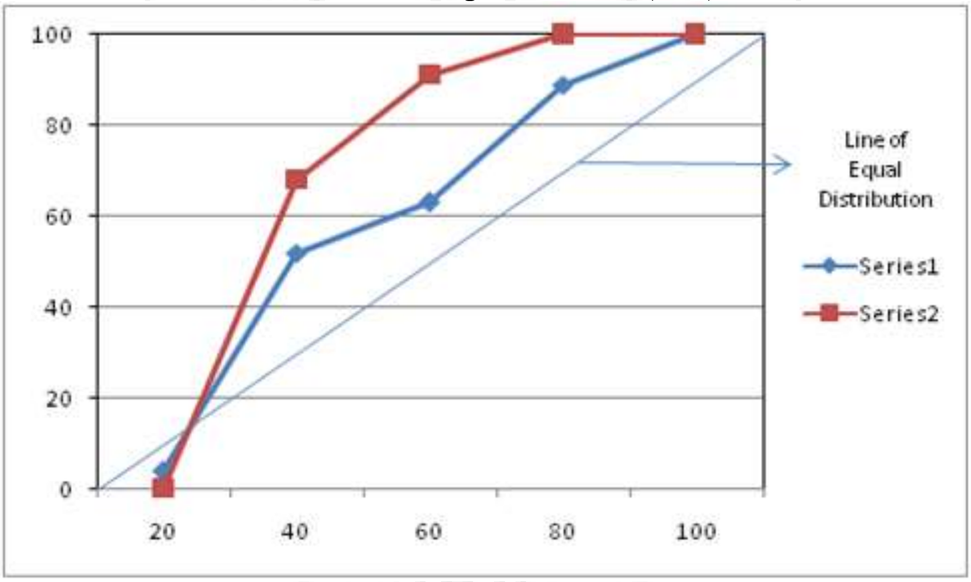

Figure 4

X-Axis = City Size **Y-Axis = No. of Cities**
Series1 = Emperical Curve **Series2 = Estimated Curve**

Lorenz Curve using Exponential Model (1981)

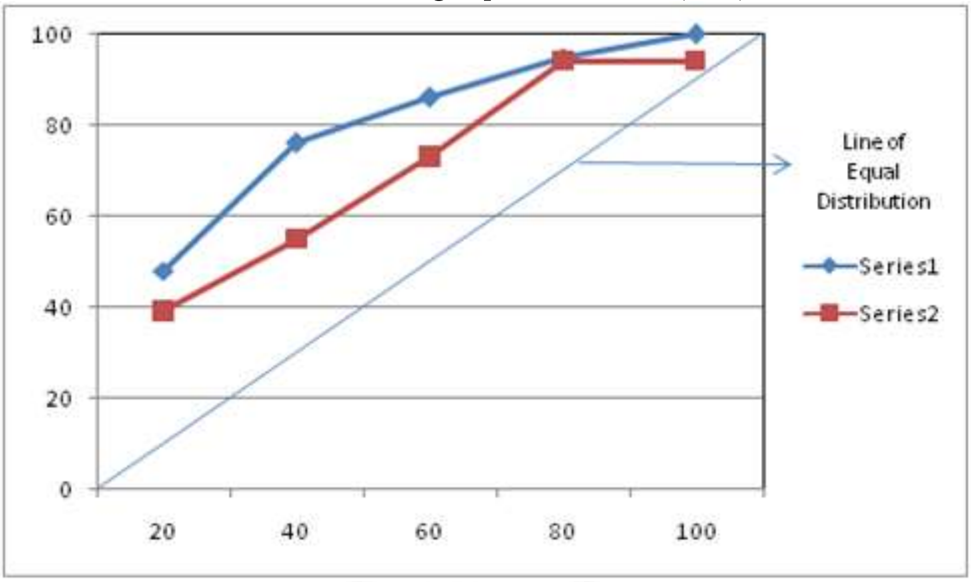

Figure 5

X-Axis = City Size **Y-Axis = No. of Cities**
Series1 = Emperical Curve **Series2 = Estimated Curve**

Lorenz Curve using Exponential Model (1991)

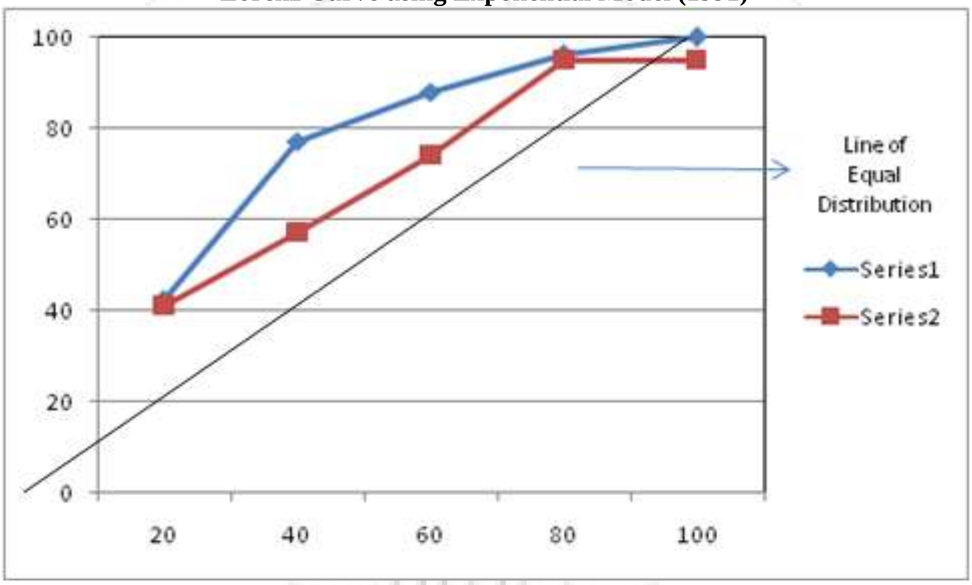

Figure 6

X-Axis = City Size **Y-Axis = No. of Cities**
Series1 = Emperical Curve **Series2 = Estimated Curve**

Lorenz Curve using Exponential Model (2001)

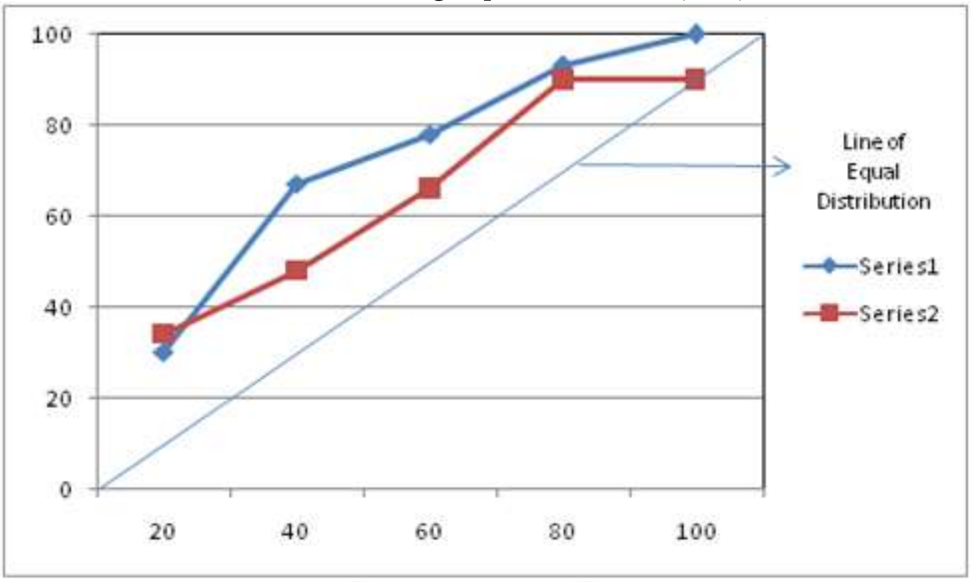

Figure 7:
X-Axis = City Size **Y-Axis = No. of Cities**
Series1 = Emperical Curve **Series2 = Estimated Curve**

Lorenz Curve using Exponential Model (2011)

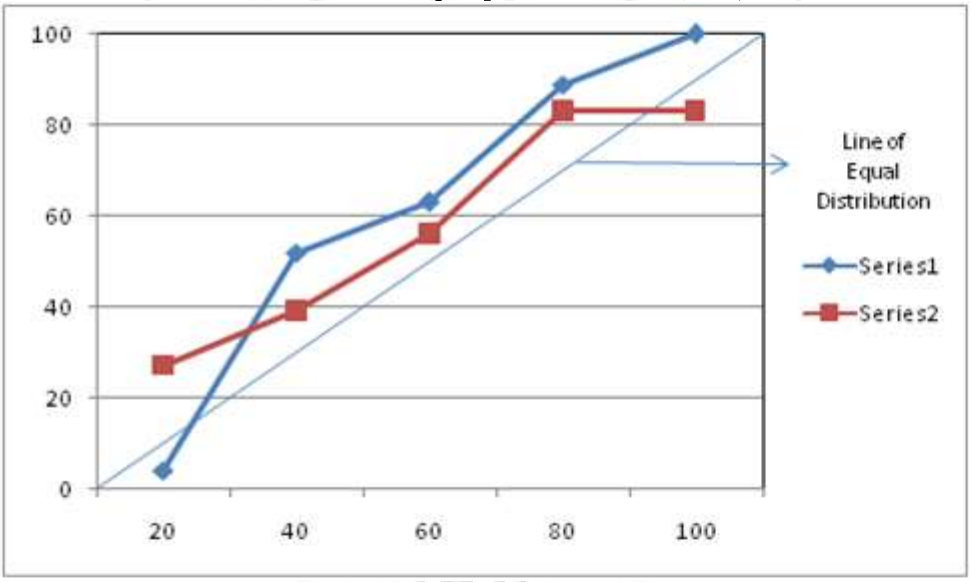

Figure 8:
X-Axis = City Size **Y-Axis = No. of Cities**
Series1 = Emperical Curve **Series2 = Estimated Curve**

5. Summary and Conclusion

Summary

Concepts such as urbanization, growth of urbanization, determines of urbanization were explained briefly. Models, urban population, models, statistical model such as Pareto distribution and Exponential distribution were explained. The object of these studies has been mentioned as to study the concentration of urban population in Tamil Nadu using statistical model. Related literatures have been reviewed and summarized the results as follows.

Much of urbanization is accommodated in smaller and medium size cities. Growth of the individual city sizes is explained by changes in local market conditions, technological change and changes in national policies. The spatial concentration has been examined by means of computing spatial Gini coefficient.

The concentration of rural population using probabilistic model have been examined using Tamil Nadu rural population data as per 1981 census. Dr.B.Renganathan studied the concentration of rural population in the villages of Tamil Nadu by developing a Gini's concentration ratio using probabilistic mode.

Methods and techniques frequently used in this study were explained and presented in methodology. Spatial concentrations of urban population have been studied by means of estimating the Gini's concentration ratio using Pareto and Exponential models. Gini's concentration are obtained and presented as follows.

6. Analytical Results

Statistical Models	Estimate of Pareto	Estimate of Gini's concentration ratio
Pareto Model	$\hat{a} = \left[\frac{\sum_i f_i \log x_i}{\sum_i f_i} - \log x^*\right]^{-1}$ $x^* = Mini(x_i)_{i \leq 1}$	$\hat{\rho} = \frac{1}{(2\hat{a} - 1)}$
Exponential Model	$\hat{\lambda} = \frac{1}{\bar{x}}$	$\hat{\rho} = 2\,\hat{\lambda}e^{-\hat{\lambda}}$

Using Tamil Nadu city population data based on 1981, 1991, 2001 and 2011 census report estimates the parameter and Gini's concentration ratio were obtained and presented as follows.

7. Empirical Result

Year	Models	
	Pareto Model	Exponential Model
1981	$\hat{a}$ =4.8852 x^* =125000 $\hat{\rho}$ =0.1140	$\hat{\lambda}$ =0.0000040 $\hat{\rho}$ =0.0000080
1991	$\hat{a}$ =4.8804 x^* =125000 $\hat{\rho}$ =0.1141	$\hat{\lambda}$ =0.0000042 $\hat{\rho}$ =0.0000084
2001	$\hat{a}$ =3.4722 x^* =125000 $\hat{\rho}$ =0.1682	$\hat{\lambda}$ =0.0000033 $\hat{\rho}$ =0.0000066
2011	$\hat{a}$ =2.3725 x^* =125000 $\hat{\rho}$ =0.2670	$\hat{\lambda}$ =0.0000025 $\hat{\rho}$ =0.0000050

8. Conclusions

Gini's concentration ratio is gradually increases over the span of 40 years, though the threshold city sizes are 125000. The increasing tendencies of concentration of urban population in Tamil Nadu state have been observed empirically. It has been interpreted due to the annexation of rural areas with urban areas.

Figures 1, 2, 3 and 4 indicates that the difference between the Lorenz curve and the line of equal distribution has been observed using Pareto model. Similarly differences were also observed in Figure 5, 6, 7 and 8 using Exponential model. But the Lorenz curve based on exponential model is very close to line of equal distribution than the Lorenz curve based on Pareto model. These differences have been attributed as the inequality existing in the city size distribution in Tamil Nadu state.

Measures such as providing urban facilities in rural areas are to be implemented to reduce the problems due to urban concentration in Tamil Nadu state.

References

[1] CENSUS OF INDIA 1981, Series – 20, Tamil Nadu Part II – A

[2] CENSUS OF INDIA 1991, Series – 23, Tamil Nadu Provisional Population Data

[3] CENSUS OF INDIA 2001, Series – 34, Tamil Nadu Provisional Population Total Director of Census Opreation, Tamil Nadu, Chennai.

[4] GUPTA.S.C. and KAPOOR V.K. (1998) – "FUNDAMENTALS OF MATHEMATICAL STATISTICS" – Sultan Chand and Sons.

[5] LYLE.W. Cook and DEAN .C. MUMME (1981) – " Estimation of Pareto parameters by Numerical Methods" – STATISTICAL DISTRIBUTIONS IN SCIENTIFIC WORK – Vol - 5, 127-132.

[6] MISRA B.D.(1981) – "AN INTRODUCTION TO THE STUDY OF POPULATION" – South Asian Publishers Pvt Ltd, New Delhi.

[7] RENGANATHAN.B (2005) – "CONCENTRATION OF RURAL POPULATION USING PROBABILISTIC MODELS" – Journal of Extention and Research Vol VII, No 1 and 2.

[8] ROHATGI V.K.(1976) – "AN INTRODUCTION TO PROBABILITY THEORY AND MATHEMATICAL STATISTICS" – Wiley Eastern Ltd, First Edition.

[9] VERNON HENDERSON.J.HYOUNG GUN WANG (2006) – "URBANIZATION AND CITY GROWTH THE ROLE OF THE INSTITUTIONS"- Paper presented at LSE, the Wider Institution and the Regional Association International meetings.