

SHIV SHAKTI

**International Journal in Multidisciplinary and
Academic Research (SSIJMAR)**

Vol. 5, No. 6, December 2016 (ISSN 2278 – 5973)

**A Statistical Study on Rural Taluk Size Distribution of SC and ST Population in Tamil
Nadu State**

Dr.M.Kalaiarasi

Assistant Professor of Statistics,
Sri Ganesh College of arts and science,
Ammamet, Salem-14,
TamilNadu, India.

**Impact Factor = 3.133 (Scientific Journal Impact Factor Value for 2012 by Inno Space Scientific
Journal Impact Factor)**

Global Impact Factor (2013)= 0.326 (By GIF)

Indexing:



Ruralisation is the process of growth with respect to the area and population. Rural areas are the geographical regions which are located out of the town or cities. Rural areas are also known as countryside or a village in India and it has a very low density of population (population of 5000 people or less). The lowest unit of a rural area is a village which is defined as an area having minimum 100 hamlets. Mahatma Gandhiji believed and advocated that the prosperity of India lies on the empowerment of the villages. His main vision becomes the empowerment of Indian rural. Villages still remain as villages and India is quoted as a rich country with poor people. In rural areas agriculture is the largest and one of the most important sectors of the rural economy and contributes both to economic growth and employment. The rural Scheduled Castes (SC) population of Tamil Nadu in 2001 census is observed as 8.309 million and in 2011 as 9.5 million. It shows the increasing nature of rural Scheduled Castes population in Tamil Nadu State. The rural Scheduled Tribes (ST) population of Tamil Nadu in 2001 census is observed as 0.551 million and in 2011 as 0.66 million. It shows the increasing nature of rural Scheduled Tribes population in Tamil Nadu State. 2011 census data indicates that in India out of the 201.3 million Scheduled Castes populations 153.8 million live in rural areas but in 2001 census data indicates that out of the 166.6 million Scheduled Castes populations 133 million live in rural areas. Thus, the rural scheduled caste population are having increasing tendency during the period 2001-2011 at India level. 2011 census data indicates that in India out of the 104.2 million Scheduled Tribe populations 93.8 million live in rural areas but in 2001 census data indicates that out of the 84.3 million scheduled tribe population 77.3 million live in rural areas. Thus, the rural scheduled Tribe population are having increasing tendency during the period 2001-2011 at India level. Thus the growth of the rural SC and ST population in India as well as in Tamil Nadu motivates to carry out the statistical study on ruralisation in Tami Nadu state. Statistical models such as Exponential, Truncated exponential and lognormal models are proposed to study the nature of the rural taluk size distribution of SC and ST population using 2011 population census data of Tamil Nadu state.

The Scheduled castes (SC's) and Scheduled tribes (ST's) are two groups of historically disadvantaged people recognized in the constitution of India. During the period of British rule in the Indian subcontinent, they were known as the Depressed Classes. In modern literature, the Scheduled Castes are sometimes referred to as 'Dalits'. Since independence, the Scheduled castes and Scheduled tribes were given reservation status, guaranteeing political representation.

Since 1850's these communities were loosely referred to as depressed classes, with the scheduled tribes also being known as 'adivasis' (original inhabitants). Nearly 70 percent of dalits live in rural areas and over 90% work in the agricultural sector as unskilled or day laborers. Most of the remainder is employed in manual, unskilled labour jobs in urban areas. As per the census 2001, the total population of the scheduled castes in the country is 166.636 million which constitute 16.2 percent of the total population, the total population of ST is 84.326 million which accounts for 8.2 percent of the total population of the country. Majority of the ST population live in rural areas and their population is 10.4 percent of the total rural population of the country.

Census 1971 recorded 79,092,841 Scheduled Castes population which is increased to 201,378,372 in 2011 in India. The share of rural Scheduled Castes population was 88 percent in 1971 which reduced to 76.4 percent in 2011, consequently the Scheduled Caste urban population in India rose from 12 percent to 23.6 percent in the same period. Total Scheduled Castes population of India in terms of percentage indicates increasing trend in 1971 to 2011 except in 2001. The total, rural and urban Scheduled Castes population observe declining trend between 1991 and 2001. The percentage of Scheduled Castes population in rural area is higher than the percentage of total Scheduled Caste population. But overall higher growth rate is observed in urban Scheduled Castes population. As per 2011 census data, the population of scheduled castes is now 16.6 percent and scheduled tribes 8.6 percent together forming a quarter of the total Indian population. During the period 2001-2011, the scheduled caste population growth is 20.8 percent and scheduled tribe population is 23.7 percent in comparison to the total population.

Primary census abstract (2011) indicates that the total scheduled caste population in 2011 is 16.6 percent and in 2001 is 16.2 percent with a growth rate of 20.8 percent. The rural scheduled caste population is 18.5 percent in 2011 and in 2001 is 17.9 percent with a growth rate of 15.7 percent. The total of scheduled tribe population is 8.6 percent in 2011 and in 2001 is 8.2 percent with a growth rate of 23.7 percent. The scheduled tribe population in rural area in 2011 is 11.3 percent and in 2001 is 10.4 percent with the growth rate of 21.3 percent. Thus, the SC and ST population in rural areas of India are having increasing tendency.

As per the census 2001, the total SC population is 11.857 million. Of this, 8.309 million are in rural areas and 3.548 million in urban areas. As per the census 2011, the total SC population is 14.4 million. Of this, 9.5 million are in rural areas and 4.9 million in urban areas. The SC

population shows an increase of 2.5 million and marked a decadal growth of 21.8 percent. As per the census 2001, the total ST population is 0.651 million. Of this, 0.551 million are in rural areas and 0.1 million in Urban areas. As per the census 2011, the total ST population is 0.79 million. Of this, 0.66 million are in rural areas and 0.13 million in Urban areas. The ST population is increased by 0.139 million constituting a decadal growth of 22 percent. Thus, the SC and ST population in rural areas of Tamil Nadu state are having increasing tendency.

Chandana R. C. (1980) analyzed distributional pattern of Scheduled Caste population in India. Gosal (1982) analyzed a comprehensive study of population growth in India. K. Sita (1986) explained about pattern of Scheduled Tribes in Maharashtra by using the district wise data, and took into consideration the growth and distribution of population. Tripathi (1999) has been undertaken study of Scheduled Caste population in view of socio-economic profile of India. Singh Deva N. (2006) examined demographic aspects of Scheduled Caste and Scheduled Tribe of Manipur by using census data from 1901-2001. Lakshmana (2008) explained the decadal variations of child population growth in Karnataka state by using census data from 1981-2001. Joachim Kaldasch (2014) studied an evolutionary model of the city size distribution. None of the above said papers deal with the study of rural taluk size distribution. The rural population census (2001 & 2011) data has the increasing tendency at India level as well as at Tamil Nadu state level. Thus the growth of the rural population in Tamil Nadu states motivates to carry out the statistical study on ruralisation in Tamil Nadu state. Statistical models such as exponential and lognormal models are proposed to study the nature of the rural taluk size distribution of population using 2001 and 2011 population census data of Tamil Nadu state. The suitability of these four models are to be tested empirically using statistical test.

Objectives

To study the growth nature of rural taluk size distribution of Scheduled castes (SC) and Scheduled tribes (ST) population in Tamil Nadu state using the statistical models such as exponential and lognormal model.

To test the suitability of the two statistical models using 2011 census SC and ST Population data of Tamil Nadu state.

Data Source

Census 2011 SC and ST population data is applied for analyzing the nature of taluk size distribution in Tamil Nadu state.

Methodology

Definition

Variable

Number of population living in rural taluk is defined as rural taluk size. It is referred as a random variable because changes in rural taluk size population are random in nature.

Rural area

An area outside the limits of a city or town or a designated on the basis of commercial, industrial, or residential center are the rural area. Rural areas are characterized by farms, vegetation and open spaces. Farmlands, homes and buildings are spread out with larger distance. Rural is the natural environment and the opposite of the city life and pertaining to the less-populated areas.

Taluk size distribution

Taluk size is the no. of population living in the taluk. Taluk size distribution is the distribution of Taluk with respect to their population size.

Models

A model establishes the main variables involved and connects them by means of mathematical statements. Thus, a model simplifies the reality and points out the basic features of the problem and denotes a set of hypothesis.

Methods

Method of Moments

X_1, X_2, \dots, X_n be a random sample of size n drawn from a given population having the probability density function $f(x; \theta_1, \theta_2, \dots, \theta_k)$ with k parameters $\theta_1, \theta_2, \dots, \theta_k$.

If μ_r' is the r^{th} moment about origin will be the function of the k parameters $\theta_1, \theta_2, \dots, \theta_k$ and μ_r' is defined as follows.

$$\mu_r' = \int_{-\infty}^{\infty} x^r f(x; \theta_1, \theta_2, \dots, \theta_k) dx, \quad r = 1, 2, \dots, k.$$

Estimate the parameters $\theta_1, \theta_2, \dots, \theta_k$ in terms of $\mu_1', \mu_2', \dots, \mu_k'$ by the method of moments (i.e.) by solving the k equations $\mu_r', r = 1, 2, \dots, k$ and replacing the population moments $\mu_r', r = 1, 2, \dots, k$ by the sample moments $m_r', r = 1, 2, \dots, k$, the r^{th} moment about origin are obtained as,

$$\mu_r' = m_r', \quad r = 1, 2, \dots, k.$$

$$m_r' = \sum_{i=1}^n X_i^r / n \quad \text{where, } m_r' \text{ is the } r^{\text{th}} \text{ sample moments about origin and } \mu_r' \text{ is the } r^{\text{th}}$$

Population moments about origin.

Graphical Method

Diagrams and graphs are used to present pattern of rural taluk size distribution and they are described as follows.

Diagrams

Diagrams are based on scale but not confined to points or lines. There are various geometrical shapes such as bars, circles, square, etc. Diagrams are visual presentation of categorical and geographical data. It furnishes only approximate information. Diagrams are more appealing to the eyes and even laymen can understand the concept easily under study.

Graphs

Graphs are more appropriate to represent the time series data and the frequency distribution. Graphs are more precise and accurate than the diagrams. Graphs can be effectively used to study the slopes, rate of change and forecasting. Frequency curve, O give curve, trend line etc., are graphs representing the data related to the study.

Techniques

Statistical models

Exponential model

Let X be a random variable having the p.d.f,

$$f(x) = \begin{cases} \lambda e^{-\lambda x}; & x > 0, \\ 0 & ; \text{ otherwise} \end{cases}$$

Here, $\lambda > 0$ is the parameter of the distribution and is estimated by the method of moments

The distribution function of the exponential model is,

$$F_X(x) = P(X \leq x) = 1 - e^{-\lambda x}$$

The mean and variance are obtained as,

$$E(x) = \frac{1}{\lambda}, \quad V(x) = \frac{1}{\lambda^2}$$

The parameter of the exponential model is,

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

Truncated exponential model

Let X be a random variable representing rural taluk size and it has the exponential probability density function as,

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & ; x > 0, \lambda > 0 \\ 0 & ; \text{ otherwise} \end{cases}$$

The truncated probability density function $g(x)$ of the random variable X when X is truncated at $X=a$ is stated as,

$$g(x) = \begin{cases} \lambda e^{\lambda a} e^{-\lambda x} ; x > a \\ 0 & ; \text{otherwise} \end{cases}$$

where $\lambda > 0$ is the parameter of the distribution and is estimated by the method of moments.

The distribution function of the truncated exponential model is,

$$F_X(x) = P(X \leq x) \\ = 1 - e^{-\lambda(x-a)} \text{ for all } x > a; \lambda > 0$$

The estimates of parameter of the truncated exponential model is obtained as,

$$\hat{\lambda} = \frac{1}{\bar{X} - a}$$

Lognormal model

Let X be a random variable having the lognormal p.d.f,

$$f(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}} , x > 0 \\ 0; \text{otherwise} \end{cases}$$

The random variable $Y = \log x$ is normal random variable having mean μ and variance σ^2 estimated by the method of moments, which $\log x$ denotes the natural logarithm of the rural population of taluks.

The distribution of the lognormal model is,

$$F_X(x) = P(X \leq x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{\ln x - \mu}{\sigma\sqrt{2}}\right)$$

Where erf denotes the error function associated with the normal distribution.

The estimates of μ and σ^2 are

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^n \log_e x_i, \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n f (\log_e x_i)^2}{N} - \left(\frac{\sum_{i=1}^n f \log_e x_i}{N} \right)^2}, \text{ Where, N is the sample size.}$$

The mean and variance of the lognormal model are obtained as,

$$\bar{X} = e^{\hat{\mu} + \frac{\hat{\sigma}^2}{2}}, \quad S^2 = e^{2\hat{\mu}} (e^{2\hat{\sigma}^2} - e^{\hat{\sigma}^2})$$

Statistical measures

Mean

The mean of the variable $X_i, i=1, 2, 3 \dots n$ is defined as,

$$\bar{X} = \frac{1}{N} \sum_{i=1}^n X_i$$

In case of frequency distribution,

$$\bar{X} = \frac{1}{N} \sum_{i=1}^n f_i X_i, \quad f_i, i=1, 2, 3, \dots, n \text{ is the frequency of the variable } X_i$$

Variance

The variance of a variable $X_i, i=1, 2, 3 \dots n$ is defined as,

$$\mu_2 = V(X) = \frac{1}{N} \sum_{i=1}^n f_i ((X_i - \bar{X})^2)$$

Where, $\bar{X} = \frac{1}{N} \sum_{i=1}^n f_i X_i, f_i$ is the frequency of the variable X_i .

Coefficient of variation

The coefficient of variation is 100 times the coefficient of dispersion based upon the standard deviation as follows:

$$\text{Coefficient of variation} = 100 \frac{\sqrt{\mu_2}}{\bar{X}}$$

Where, $\mu_2 = \frac{1}{N} \sum_{i=1}^n f_i ((X_i - \bar{X})^2), \bar{X} = \frac{1}{N} \sum_{i=1}^n f_i X_i$

Professor Karl Pearson suggested that the measure coefficient of variation is the percentage variation in the mean, standard deviation being considered as the total variation in the mean.

Skewness

The coefficient of Skewness based upon the moments about mean as follows.

$$\beta_1 = \frac{\mu_3}{\mu_2^{3/2}}, \text{ Where, } \mu_2 = \frac{1}{N} \sum_{i=1}^n f_i ((X_i - \bar{X})^2),$$

$$\mu_3 = \frac{1}{N} \sum_{i=1}^n f_i ((X_i - \bar{X})^3)$$

Kurtosis

The coefficient of Kurtosis based upon the moments about mean as follows.

$$\beta_2 = \frac{\mu_4}{\mu_2^2}, \text{ Where, } \mu_2 = \frac{1}{N} \sum_{i=1}^n f_i ((X_i - \bar{X})^2),$$

$$\mu_4 = \frac{1}{N} \sum_{i=1}^n f_i ((X_i - \bar{X})^4)$$

Test statistic

Chi square test

If $X_i, i=1,2,3,\dots,n$ are n independent normal variates with mean μ_i and variances $\sigma_i^2, i=1,2,3,\dots,n$, then

$$\chi^2 = \sum \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2, \text{ is a chi-square } (\chi^2) \text{ variate with } n \text{ degrees of freedom.}$$

The χ^2 statistic is stated as,

$$\chi^2 = \sum_{i=1}^n \left[\frac{(O_i - E_i)^2}{E_i} \right] \sim \chi^2_{(n-1)} \text{ when } \sum_{i=1}^n O_i = \sum_{i=1}^n E_i$$

Where, O_i is the observed no. rural taluk and E_i is the expected no. of rural taluk

Normal test

A random sample X_1, X_2, \dots, X_n of size n has been drawn from a normal population with mean μ and variance σ^2 . The sample mean \bar{X} is distributed normally with mean μ and variance σ^2/n when X represents rural taluk size.

ie, $\bar{X} \sim N(\mu, \sigma^2/n)$. The standard normal variate based on the sample mean \bar{X} is described as,

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

The standard normal variate Z is distributed as normal distribution with mean 0 and variance 1 when H_0 is true. In such a situation the standard normal statistic is called normal test statistic for testing the null hypothesis H_0 . The normal test statistic, for testing the equality of the two sample means, the normal test statistic is described as follows:

Let X_1 be the observed taluk rural size and X_2 be the expected taluk rural size, N_1 be the observed no. of rural taluks and N_2 be the expected no. of rural taluks from the population with means μ_1 and μ_2 and variances S_1^2 and S_2^2 respectively.

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}} \sim N(0, 1)$$

Where, \bar{X}_1 = Mean of the observed taluk size, \bar{X}_2 = Mean of the expected taluk size.

Find the Z table value at the desired level of significance and compared with the Z calculated value.

H_0 is accepted if $|Z| < Z_\alpha$ at $\alpha\%$ level of significance

H_0 is rejected if $|Z| > Z_\alpha$ at $\alpha\%$ level of significance

EMPIRICAL RESULTS

Rural population data in Tamil Nadu State provided by 2011 census of India are used to test the goodness of fit of the statistical models proposed in the present investigation of Rural Taluk Size Distribution. Rural Taluk Size Distribution of **SC and ST** population based on 2011 census resembles the skew distribution as seen in Table no.1.

The Rural Taluk Size Distribution of SC and ST population has been studied using 2011 Tamil Nadu census data. Empirical Statistical models such as Exponential, Truncated exponential and lognormal models were proposed to study the nature of distribution of rural taluk size. The empirical results are summarized below.

Table No.1 EMPIRICAL DISTRIBUTION OF RURAL TALUK SIZE IN 2011

Rural Taluk Size (in'000)	No. of rural taluks
	2011
0-20	34
20-40	70
40-60	47
60-80	22
80-100	16
100-120	11
120-140	4
>140	2
Total	206

Exponential Model

Exponential Model is fitted using **2011** population data given in the Table no. 1 as follows.

The mean of the empirical rural taluk size distribution is obtained as, $\bar{X} = 47.5728$,

The estimates of the parameter of the exponential model is obtained as,

$$\hat{\lambda} = 0.0210 > 0$$

The fitted model is stated as,

$$f(x) = \begin{cases} (0.021)e^{(-0.021)x} & ; x > 0 \\ 0 & ; \text{otherwise} \end{cases}$$

The expected no. of rural taluks is obtained using the fitted function $f(x)$ and it is presented in Table no. 2 below.

Table no. 2 Observed and the expected distribution of rural taluks for

SC & ST Population – 2011

Rural Taluk Size (in'000)	Observed no. of rural taluks	Probability values f(x)	Expected no. of rural taluks
0-20	34	0.3429	71
20-40	70	0.2253	46
40-60	47	0.1481	30
60-80	22	0.0973	20
80-100	16	0.0639	13
100-120	11	0.0420	9
120-140	4	0.0276	6
>140	2	0.0529	11
Total	206	1.0000	206

Using the informations given in table 3, statistical measures of the rural taluk size distribution are computed and presented as follows.

The mean of the rural taluk size, $\bar{X} = 46.5048$

The variance of the rural taluk size, $S^2 = 1646.0364$

The coefficient of variation, $C.V = 87\%$

The measures of skewness, $\beta_1 = 1.2958$

The measures of Kurtosis, $\beta_2 = 3.372$

Normal test Statistic value, $|Z| = 0.2994$ is less than the table value of $Z = 1.96$ at 5% level of significance. It implies that the null hypothesis is accepted at 5% level of significance.

χ^2 test statistic value, $\chi^2 = 50.8037$ is greater than the table χ^2 value at 5% level of significance with 6 d.f is 12.592. It shows that the null hypothesis is rejected at 5% level of significance; the goodness of exponential model to the Rural Taluk Size Distribution in 2011 is not good.

Truncated Exponential Model

The rural taluk size random variable X is truncated at the point X =20 in 2011. Hence the empirical Rural Taluk Size Distribution of **SC and ST** population becomes the truncated Rural Taluk Size Distribution and presented in Table -3.

Table -3Truncated Rural taluk size distribution of SC and ST

Population –2011

Rural Taluk Size (in'000)	Number of rural taluks
	2011
20-40	70
40-60	47
60-80	22
80-100	16
100-120	11
120-140	4
>140	2
Total	172

Truncated Exponential Model is fitted using **2011** population data given in the Table-3 as follows.

The mean of the empirical rural taluk size distribution is obtained as,

$$\bar{X} = 55$$

The estimates of the parameter of the truncated exponential model is obtained as,

$$\hat{\lambda} = 0.0286$$

The fitted model is stated as,

$$f(x) = \begin{cases} (0.0286) e^{(0.0286)(20)} e^{(-0.0286)x} & ; x > 20 \\ 0 & ; \text{otherwise} \end{cases}$$

The expected number of rural taluks is obtained using the fitted function $f(x)$ and it is presented in Table -4 below:

Table -4 Observed and the expected distribution of rural taluks for

SC and ST Population – 2011

Rural Taluk Size (in'000)	Observed number of rural taluks	Probability values $f(x)$	Expected number of rural taluks
20-40	70	0.4356	75
40-60	47	0.2458	42
60-80	22	0.1387	24
80-100	16	0.0783	13
100-120	11	0.0442	8
120-140	4	0.0249	4
>140	2	0.0325	6
Total	172	1.0000	172

Using the information given in Table -4, statistical measures of the truncated rural taluk size distribution are computed and presented as follows:

The mean of the rural taluk size, $\bar{X} = 55.2325$

The variance of the rural taluk size, $S^2 = 988.8994$

The coefficient of variation, $C.V = 57\%$

The measures of skewness, $\beta_1 = 1.9689$

The measures of Kurtosis, $\beta_2 = 4.3488$

Normal test Statistic value, $|Z| = 0.0219$ is less than the table value of $Z = 1.96$ at 5% level of significance. It implies that the null hypothesis is accepted at 5% level of significance. i.e., “there is no difference in means between the empirical and fitted value”.

χ^2 test statistic value, $\chi^2 = 4.5125$ is less than the table χ^2 value at 5% level of significance with 4 d.f is 9.49. It shows that the null hypothesis is accepted at 5% level of significance; the goodness of truncated exponential model to the Rural Taluk Size Distribution in 2011 is good.

Lognormal Model

Lognormal model is fitted using **2011** population data given in the Table no.1 as follows.

The estimates of the parameters μ and σ^2 in the normal distribution are obtained as,

$$\hat{\mu} = 3.6257, \hat{\sigma} = 0.7320$$

Based on $\hat{\mu}$ and $\hat{\sigma}$, mean and variance of the lognormal distribution are obtained as,

The mean of the empirical rural taluk size distribution,

$$\bar{X} = 49.0873$$

The variance of the empirical rural taluk size distribution,

$$S^2 = 1709.1371$$

The fitted model is described as,

$$f(x) = \begin{cases} \frac{1}{x(0.7320)\sqrt{2\pi}} e^{\frac{-1}{2(0.7320)^2}(\log x - 3.6257)^2} & ; x > 0 \\ 0 & ; \text{otherwise} \end{cases}$$

The expected no. of rural taluks is obtained using the fitted function $f(x)$ and it is presented in table no.5 below.

Table no.5 Observed and the expected distribution of rural taluks for

SC & ST Population – 2011

Rural Taluk Size (in'000)	Observed no. of rural taluks	Probability values f(x)	Expected no. of rural taluks
0-20	34	0.1947	40
20-40	70	0.3397	70
40-60	47	0.2046	42
60-80	22	0.1100	23
80-100	16	0.0606	13
100-120	11	0.0341	7
120-140	4	0.0202	4
>140	2	0.0361	7
Total	206	1.0000	206

Using the informations given in table no.5, statistical measures of the rural taluk size distribution are computed and presented as follows.

The mean of the rural taluk size, $\bar{X} = 47.1845$

The variance of the rural taluk size, $S^2 = 1170.7135$

The coefficient of variation, $C.V = 72\%$

The measures of skewness, $\beta_1 = 1.6184$

The measures of Kurtosis, $\beta_2 = 4.2922$

Normal test Statistic value, $|Z| = 0.0373$ is less than the table value of $Z = 1.96$ at 5% level of significance. It implies that the null hypothesis is accepted at 5% level of significance.

χ^2 test statistic value, $\chi^2 = 6.7894$ is less than the table χ^2 value at 5% level of significance with 4 d.f is 9.49. It shows that the null hypothesis is accepted at 5% level of significance; the goodness of lognormal model to the Rural Taluk Size Distribution in 2011 is good.

SUMMARY AND COCLUSION

Summary

Inferences based on Statistical test

Inferences about the suitability of the models are obtained using 2011 census population data and presented as follows.

Chi-square test of goodness of fit

Table-6 Inferences based on χ^2 Statistics- Rural Taluk SC&ST population

Model	Test Statistic Chi-square values	Inference
Exponential	$\chi^2=50.8037$ $\chi^2_{0.05}$ with 6 d. f is 12.592	rejected at 5% level of significance
Truncated Exponential model	$\chi^2=4.5125$ $\chi^2_{0.05}$ with 4 d. f is 9.49	accepted at 5% level of significance
Lognormal	$\chi^2=6.7894$ $\chi^2_{0.05}$ with 4 d. fis 9.49	accepted at 5% level of significance

$$\chi^2 = \sum_{i=1}^n [(O_i - E_i)^2 / E_i], \quad \text{when } \sum_{i=1}^n O_i = \sum_{i=1}^n E_i$$

H_0 : The fit of the models such as exponential, Truncated exponential and lognormal to the rural taluk size distribution of the SC & ST population are good.

Based on the χ^2 test, the fit of the exponential model to the rural taluk size distribution of SC & ST population is not good because the null hypothesis (H_0) is rejected at 5% level of significance. Hence the exponential model is not preferred to represent the rural taluk size distribution of SC & ST population in the year 2011. The fit of the truncated exponential and lognormal models to the rural taluk size distribution of SC & ST population is good because

the null hypothesis (H_0) is accepted at 5% level of significance. Hence the truncated exponential and lognormal models are preferred to represent the rural taluk size distribution of SC & ST population in the year 2011.

Statistical measures- 2011 SC & ST population

Model wise averages of rural taluk size were obtained and presented in the Table no.9 below.

Table-7. Averages of Rural Taluk Size - SC and ST population 2011

MODELS	AVERAGE RURAL TALUK SIZE
Empirical model	47.5728
Exponential	46.5048
Truncated exponential	55.2325
Lognormal	47.1845

The average lognormal rural taluk size distribution is deviated in least from the average rural taluk size of empirical distribution and exponential and truncated exponential models are having the larger deviations in the year 2011. It indicate that lognormal model may be preferred than the exponential and exponential models.

The variability of rural taluk size for the fitted distributions such as exponential and lognormal models are measured in terms of coefficient of variation and presented in the Table-8.

Table-8 Coefficient of variation of rural taluk size -SC&ST population 2011

MODELS	COEFFICIENT OF VARIATION
Exponential	87%
Truncated exponential	57%
lognormal	72%

The coefficient of variation in both the years due to Truncated exponential model is minimum than the exponential and lognormal models. It shows that the truncated exponential model may be preferred than the exponential and lognormal models because less amount of variation in truncated exponential rural taluk size distribution has been observed.

The skew nature of the models were studied in terms of measures of skewness β_1 . It was obtained and presented in the Table-9.

Table-9 Skewness values of rural taluk size -SC&ST population 2011

MODELS	Skewness measure(β_1)
Exponential	1.2958
Truncated exponential	1.9689
lognormal	1.6184

β_1 Values are greater than zero in the three models. Hence the exponential, truncated exponential and lognormal models are referred as a skew model. Hence these models may be used to describe the taluk size distribution.

The kurtosis of the four models are obtained and presented in the Table-10.

Table-10 Kurtosis values of rural taluk size -SC&ST population 2011

MODELS	Kurtosis measure(β_2)
Exponential	3.372
Truncated exponential	4.3488
lognormal	4.2922

Based on the SC & ST population exponential model having β_2 values where very closed to three than the other models in both the year 2011. It indicate that these models are moving towards the symmetric nature in the long run.

Graph

Graphs are drawn using rural taluk SC and ST Population data of 2011 census and presented in the Figure.1

Curves of proposed models- SC and ST population population 2011

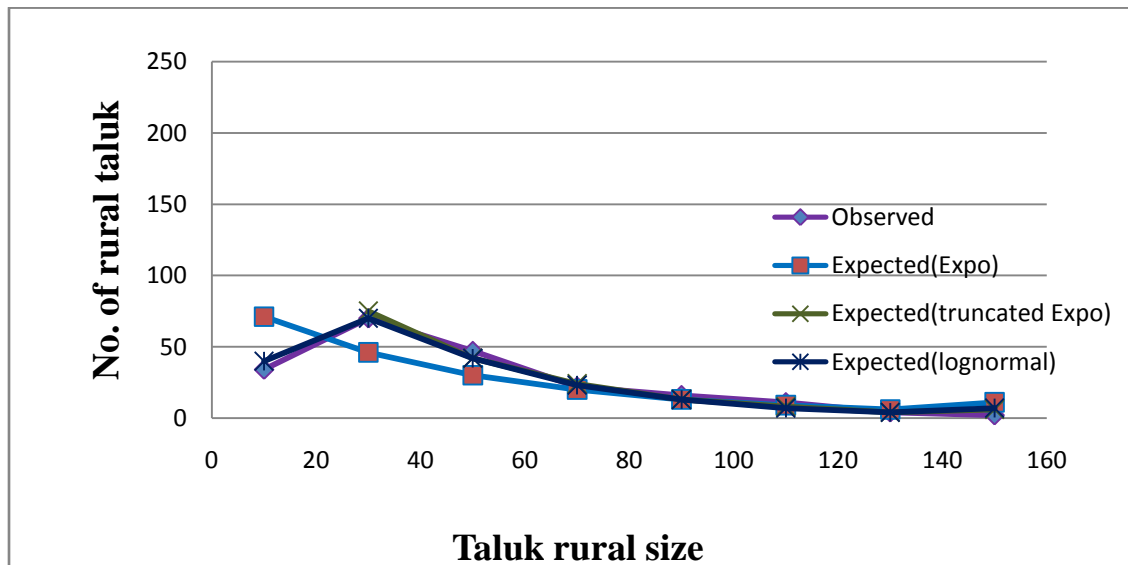


Fig.1

Fig.1 indicates that the expected curve of truncated exponential and lognormal model are very closed to the observed curve but the expected curve of exponential model is deviated from the observed curve. Thus, truncated exponential and lognormal model are more suitable than the other model and also these two models are the best description of the growth nature of the rural taluk SC and ST population in the year 2011.

Conclusion

The exponential, truncated exponential and lognormal models were fitted to the empirical rural taluk size distribution based on the 2011 census data for SC & ST population. Inferences were made that lognormal for rural taluk size distribution for SC and ST population in 2011 in Tamil Nadu State has been identified as a best model because

- * its goodness of fit is good at 5% level of significance,
- * it is having a average taluk size very closed to the average taluk size of empirical distribution,
- * it is having a skew nature because β_1 values are greater than zero,
- * it has a least amount of variation in taluk size than the other models,
- * its curve is moving towards the curve of empirical taluk size distribution.

Thus, the lognormal model is used to describe the nature of the rural taluk size distribution of SC and ST population. The present investigation is useful to analyze the nature of rural taluk size in rural system in Tamil Nadu State. Rural taluk size can be forecasted using lognormal model which will be useful to the policy makers and planners for rendering the services to the rural population.

REFERENCE

1. Chandana R. C. (1980): "Distributional Pattern of Scheduled Caste Population in India", Population Geography, Vol. 2, Pp. 78-89.
2. Directorate of census operations, Tamilnadu, Chennai: primary census Abstract, Census-2001&2011.
3. Gosal G. S. (1982): "Recent Population Growth in India", Population Geography, Vol. IV, Pp. 30-51.
4. R.K .Gupta (2004). "RURAL DEVELOPMENT IN INDIA". Atlantic publishers and distributors, New Delhi.

- 5.S.C.Gupta and V.K. Kapoor (2007).”FUNDAMENTALS OF MATHEMATICAL STATISTICS”.Sultan and sons, New Delhi.
- 6.Joachim Kaldasch (2014). “Evolutionary Model of the City Size Distribution”.Hindawi Publishing Corporation, ISRN Economics, Volume 2014, Article ID 498125, 6 pages.
7. Lakshmana (2008): “The Decadal Variation of Child Population Growth in Karnataka State”, the Deccan Geographer, Vol. 44, No. 1, Pp. 11-23.
8. ParimalMukhopadhyay (2009). “MATHEMATICAL STATISTICS”. Books and allied (P) Limited,Cacutta.
9. V.K. Rohatgi (1985). “AN INTRODUCTION TO PROBABILITY THEORY AND MATHEMATICAL STATISTICS”.John Wiley and sons, Inc.
10. H.C. Saxena, P.U. Surendran (1990).”STATISTICAL INFERENCE”. Sultan chand and company, Ltd., New Delhi.
11. K. Sita (1986): “Spatial Pattern of Scheduled Caste Tribes in Maharashtra”, the Geographer, vol. 37, Pp 53-59.
12. Singh Deva N. (2006): “A Study of Demographic Aspects of Scheduled Caste and Scheduled Tribes in Manipur”, the Deccan Geographer, Vol. 44, No. 1, Pp. 17-28.
13. S. SuddenduBiswas, (2011). “MATHEMATICAL STATISTICS”.Narosa publishing house, New Delhi.
14. Tripathi Rajmani (1999): “Socio-Economic Profile of Scheduled Caste Population in India”, Geographical Review of India, Vol. 61, No. 2, Pp. 47-59.

Website reference

- 1.<http://thpindia.org/2011/02/15/an-overview-of-the-socio-cultural-economic-and-political-scenario/>
2. <http://www.census2011.co.in/census/state/tamil+nadu.html>
3. <http://www.yourarticlelibrary.com/population/growth-of-population-in-india-1901-to-2001-with-statistics/39653/>
4. Famous tamilnadu.blogspot.com/2011/05/demographics-of-tamilnadu.html.
5. en.wikipedia.org/wiki/scheduled-castes-and_scheduled-Tribes

6. vikaspedia.in/social-welfare/scheduled-caste-welfare-1/status-report.
7. [Census of india.gov.in/census.../scheduled-castes-and scheduled-tribes.aspx](http://Census%20of%20india.gov.in/census.../scheduled-castes-and%20scheduled-tribes.aspx)
8. [http://en.wikipedia.org/wiki/List of Developmental administrative units of Tamil Nadu](http://en.wikipedia.org/wiki/List_of_Developmental_administrative_units_of_Tamil_Nadu).
9. www.businessdictionary.com/definition/rural.html.