International Journal of Business Economics and Management Research
Vol. 8, Issue 1, January 2017 Impact Factor: 4.963 ISSN: (2229-4848)
www.skirec.org Email Id: skirec.org@gmail.com

A STUDY ON THE THEORETICAL RELATIONSHIP BETWEEN RANK SIZE RULE AND LOGNORMAL RURAL TALUK SIZE DISTRIBUTION

Dr.M.Kalaiarasi

Assistant Professor of Statistics,

Sri Ganesh College of arts and science,

Ammapet, Salem-14,

TamilNadu, India.

Abstract

It highlights the theoretical relationship between rank size rule and Rural Taluk Size Distribution.

Introduction

Rural place located outside the city or town with population less than or equal to five thousand is called as a rural area. Classification of rural taluk size with respect to the size is called as Rural Taluk Size Distribution. It has the substantive interest in many socio and demographic fields. The upper tail of the lognormal distribution confirms to the rank size rule (Aitchison and Brown, 1957). The lognormal model is plausible model of rank size rule because lognormal model generates the rank size rule.

An attempt has been made to study the relationship between the rank size rule and the Rural Taluk Size Distribution when rural taluks are ranked in population size from largest to smallest.

Order statistics

The function $X_{(k)}$ of $(X_1, X_2, \ldots, X_n)$ that takes on the value $x_{(k)}$ in each possible sequence $(x_1, x_2, \ldots, x_n)$ of values assumed by $(X_1, X_2, \ldots, X_n)$ is known as the k^{th} order statistic (or) statistic of order k $\{X_{(1)}, X_{(2)}, \ldots, X_{(n)}\}$ is called the set of order statistics for $(X_1, X_2, \ldots, X_n)$.

Lognormal model

The basic underlying assumption of the lognormal model is "the law of proportionate effect", or Gibrat's law. Aitchison and Brown (1957) define the law as the process where "the change in the variate at any step in the process is a random proportion of the previous value of the variate". That is, if the growth rate of a unit between time t_0 and t_1 is defined as,

$$g_1 = \frac{S_{t_1} - S_{t_0}}{S_{t_0}} \qquad (1)$$

Then this number will be constant across all size classes. The growth rates are independent of size.

Aitchison and Brown (1957) show how Gibrat's law generates a lognormal distribution. If we have the random variableS_t, then by (1)

$$\sum_{t=1}^{n} \frac{S_t - S_{t-1}}{S_{t-1}} = \sum_{t=1}^{n} g_t \qquad (2)$$

and if the jump for each period is small.

$$\sum_{t=1}^{n} \frac{S_t - S_{t-1}}{S_{t-1}} \sim \int_{S_0}^{S_n} \frac{dS}{S} = \log S_n - \log S_0 \qquad (3)$$

(1) and (3) ⇨

$$\text{Log } S_n = \log S_0 + g_1 + g_2 + \cdots + g_n \qquad (4)$$

By the central limit theorem, the sum of independent random variables is asymptotically normal. Hence, the g's are normally distributed and $S_n$is lognormally distributed. Thus, Gibrat's law implies that growth rate is independent of size; the resulting distribution of size of rural taluks will be lognormal.

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

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

The random variable Y=log x is normal random variable having mean μ and variance$\sigma^2$where, log x denotes the natural logarithm of the rural population of taluks.

The distribution function of the lognormal model is,

$$F_X(x) = \text{P}\ (\text{X} \leq x)$$

$$= \frac{1}{2} + \frac{1}{2}\text{erf}\left(\frac{Inx-\mu}{\sigma\sqrt{2}}\right)$$

whereerf denotes the error function associated with the normal distribution.

The estimates of $\mu and \sigma are$

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{n} log_e x_i$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} f(log_e x_i)^2}{N} - \left(\frac{\sum_{i=1}^{n} f log_e x_i}{N}\right)^2}$$

where 'N' denotes the total number of rural taluks and 'n' denotes the number of rural taluk classes.

Beta distribution

Let X be a random variable representing rural taluk size and it has the Beta one distribution with parameters (r, n-r+1), then its probability density function is stated as,

$$\text{f}(x) = \frac{1}{B(\text{r},\text{n}-\text{r}+1)}(x)^{r-1}(1-x)^{n-r}\ ;\ B(\text{r},\text{n}-\text{r}+1) > 0,\ 0 < x < 1$$

where $\text{B}(\text{r},\text{n}-\text{r}+1)$is the Betafunction.

The mean and variance of the rural taluk size distribution are obtained as,

$$\text{E (X)} = \frac{\text{r}}{\text{n}+1},$$

$$\text{V (X)} = \frac{\text{r}(\text{n}-\text{r}+1)}{(\text{n}+1)^2(\text{n}+2)}$$

Distribution of rth order statistics

Let $X_{(r)}$be the r^{th} order statistics, then its probability density function is stated as,

$$g_r(\text{x}_\text{r}) = \frac{n!}{(r-1)!(n-r)!}\ [F(\text{x}_\text{r})]^{\text{r-1}}\,[1-F(\text{x}_\text{r})]^{\text{n-r}}\ \text{f}(\text{x}_\text{r})$$

where F is the common distribution function of X.

Rank Size Rule

The relation,

$X_{(r)}R_{(r)}^{q} = C$, for all r = 1, 2, 3 . . . n,

where n is the no. of taluks,

$X_{(r)}$ is the size of the r^{th} ranked taluks,

$R_{(r)}$ is the rank of the r^{th} taluks, C and q are constants.

is called as rank size rule.

Rank size rule has been described probabilistically through an application of order statistics to study the relationship between rank size rule and Rural Taluk Size Distribution.

At a certain given point of time, 'n' taluks in a state are ordered or ranked

according to their sizes $X = [X_{(1)}, X_{(2)}, \ldots X_{(n)}]$. A. Okabe (1979) assumed that the set of observed values of X consists of 'n' taluk size values. These values are sampled according to the same RTSDF (X).As the observed Taluk Size values are sampled, ranked 'n' Taluk Size values $X = [X_{(1)} \; X_{(2)} \ldots X_{(r)} \ldots X_{(n)} ; X_{(1)} < X_{(2)} < \ldots < X_{(n)}]$ are probabilistic. Then expected Taluk Size is obtained as,

$$E(X) = \{E[X_{(1)}]\, E[X_{(2)}] \ldots E[X_{(r)}] \ldots [E[X_{(n)}]\}$$

By using the expected Taluk Size, the Rank Size Rule is recalled as expected Rank Size Rule.

i.e., $E[X_{(r)}]\, R_{(r)}^{q} = C$, for all r = 1, 2. . . n,

Strong Expected Rank Size rule

A. Okabe (1979) established Strong Expected Rank Size Rule as,

$E[X_{r/n}]\, R(r) = C(n)$ when $r = r^{*}, r^{*}+1, \ldots, n$

$n = r^{*}+1, r^{*}+2, \ldots$

where r^* is the minimum positive integer such that R(r) > 0, R(r) rank function, is dependent of n, and C (n) is a constant function of 'n'.

The relationship between Rural Taluk Size Distribution and the Strong expected rank size rule is explained in the following section.

Lognormal Rural Taluk Size Distribution and Strong expected rank size rule

Let $X_1, X_2, X_3, \ldots, X_n$ be the Rural Taluk Size random variable sampled from the distribution function F and rural taluk size are independent identically distributed with lognormal probability density function,

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

when $X_{(r)}$ be the r[th] order Statistics, its probability density function corresponding to $X_{(r)}$ is obtained as,

$$f[x_{(r)}] = \begin{cases} \frac{1}{X_{(r)}\sigma\sqrt{2\pi}} e^{-\frac{(\log X_{(r)}-\mu)^2}{2\sigma^2}} & , x_{(r)}>0, -\infty < \mu < \infty \text{ and } \sigma > 0 \\ 0 & ; \text{otherwise} \end{cases}$$

The standardized lognormal variable,

$$\zeta = \frac{X_{(r)}-\mu_1'}{\sqrt{\mu_2}}$$

$$= \frac{e^{Z_{(r)}\sigma+\mu} - e^{\mu+\frac{\sigma^2}{2}}}{[e^{\sigma^2}(e^{\sigma^2}-1)\, e^{2\mu}]^{\frac{1}{2}}}$$

where μ_1' is the mean of the lognormal distribution and μ_2 is the variance of the lognormal distribution.

$$\mu_1' = e^{\mu+\frac{\sigma^2}{2}},$$

$$\mu_2 = e^{\sigma^2}(e^{\sigma^2}-1)\, e^{2\mu}$$

$$Z_{(r)} = \frac{\log X_{(r)}-\mu}{\sigma}$$

$\Rightarrow \text{Log } x_{(r)} = Z_{(r)}\sigma + \mu$

$$x_{(r)} = e^{Z_{(r)}\sigma + \mu}$$

$$\therefore \quad \zeta = \frac{e^{Z_{(r)}\sigma + \mu} - e^{\mu + \frac{\sigma^2}{2}}}{[e^{\sigma^2}(e^{\sigma^2}-1)\, e^{2\mu}]^{\frac{1}{2}}}$$

$$= \frac{e^{\mu}\,(e^{Z_{(r)}\sigma} - e^{\frac{\sigma^2}{2}})}{e^{\mu}\,[\,e^{\sigma^2}\left(e^{\sigma^2}-1\right)]^{\frac{1}{2}}}$$

$$= \frac{e^{Z_{(r)}\sigma} - e^{\frac{\sigma^2}{2}}}{[\,e^{\sigma^2}\left(e^{\sigma^2}-1\right)]^{\frac{1}{2}}} \rightarrow (1)$$

As $\sigma \to 0$,

$$(1) \to Z_{(r)} = \frac{\log X_{(r)} - \mu}{\sigma} \to N(0, 1)$$

i.e. $\zeta \to$ Unit normal (or) standardized normal distribution as $\sigma \to 0$.

i.e. Lognormal distribution gives a good approximation to normal distribution.

$$Z_{(r)} = \frac{\log X_{(r)} - \mu}{\sigma} \sim N(0, 1) \text{ as } n \to \infty$$

The probability density function of $Z_{(r)}$ is

$$f[z_{(r)}] = \frac{1}{\sqrt{2\pi}} e^{\frac{-z_{(r)}^2}{2}} \qquad \text{when } -\infty < z_{(r)} < \infty$$

The distribution function of $Z_{(r)}$ is stated as,

$$F[z_{(r)}] = \int_{-\infty}^{Z_{(r)}} f[z_{(r)}]\, dz_{(r)}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_{(r)}} e^{\frac{-z_{(r)}^2}{2}}\, dz_{(r)}$$

The r^{th} order Rural Taluk Size Distribution is stated as,

$$f[z_{(r)}] = \frac{1}{B(r, n-r+1)} f[z_{(r)}]\, [F(z_{(r)})]^{r-1}\, [1-F(z_{(r)})]^{n-r}$$

Let $F[z_{(r)}] = U_{(r)}$

The r[th] order Rural Taluk Size Distribution is,

$$g\,[z_{(r)}] = \frac{1}{B(r,n-r+1)}\, f\,[z_{(r)}]\,[U_{(r)}]^{r-1}\,[1-U_{(r)}]^{n-r}\frac{1}{f\,[z_{(r)}]}$$

$$= \frac{1}{B(r,n-r+1)}[U_{(r)}]^{r-1}\,[1-U_{(r)}]^{n-r}, \quad 0 < U_{(r)} < 1$$

which is called as a beta one distribution with parameters (r, n-r+1)

$$E\,[U_{(r)}] = \frac{r}{r+n-r+1} = \frac{r}{n+1}$$

$$V\,[U_{(r)}] = \frac{r(n-r+1)}{(n+1)^2(r+n-r+1+1)}$$

$$= \frac{r(n-r+1)}{(n+1)^2)(n+2)}$$

$$E\,[U_{(r)}] = \frac{r}{n+1} = E\,[F\,(z_{(r)})]$$

By using the probability integral transformation,

$$F^{-1}[U_{(r)}] = z_{(r)} \qquad [\because U_{(r)} = F\,(z_{(r)})]$$

$$z_{(r)} = F^{-1}\,[U_{(r)}]$$

$$E\,[Z_{(r)}] = E\,[F^{-1}(U_{(r)})]$$

$$\leq F^{-1}[E\,(U_{(r)})] \qquad [\because \text{Jensen's inequality } f\,[E(X)] \leq E\,[f(x)]$$

where, f is a convex and monotonic increasing function]

$$= F^{-1}\left(\frac{r}{n+1}\right)$$

$$= \frac{r}{n+1}$$

$$= \frac{g(r)}{n+1}, \qquad \text{where, } g(r) = r$$

$$\therefore E[Z_{(r)}] \leq \frac{g(r)}{n+1}.$$

when $F^{-1}[U_{(r)}]$ is linear function in $U_{(r)}$, $F^{-1}(\frac{1}{n+1})$ is linear, $g(r) = r$

$$E\,[Z_{(r)}] = \frac{r}{n+1}$$

$E[Z_{(r)}].\frac{1}{r} = \frac{r}{n+1}$

$E[Z_{(r)}]\, R(r) = c(n)$

where $R(r) = \frac{1}{r}$,

$c(n) = \frac{1}{n+1}$

The strong expected rank size rule,

$E[Z_{(r)}]\, R(r) = c(n)$

where $R(r) = \frac{1}{r}$,

$c(n) = \frac{1}{n+1}$

is satisfied by the Lognormal Rural Taluk Size Distribution because

$F^{-1}[U_{(r)}]$: inf $\{ Z_{(r)}: F[z_{(r)}] \geq \frac{r}{n+1} \}$

is satisfied by Lognormal Rural Taluk Size Variable.

Conclusion

A Rural Taluk Size Distribution is related to Rank Size Rule in terms of Strong expected Rank Size rule. It is shown that Strong expected Rank Size rule is satisfied by Lognormal Rural Taluk Size Distribution. Lognormal model confirm the real distribution of rural taluk size. The present investigation suggests to the future researchers for analyzing the nature of rural taluk size in all states of India.

REFERENCE

1. Aitchison. J. and J.A.C. Brown (1957) "The Lognormal Distribution", Cambridge University Press, Cambridge.

2. Atsuyuki OKABE (1979) "A theoretical relationship between the rank size rule and city size distributions".Regional science and urban economics. Elsevier, ISSN 0166-0462, Volume-9, 1979, pages 21-40.

3. R.K .Gupta (2004).”RURAL DEVELOPMENT IN INDIA”.Atlantic publishers and distributors, New Delhi.

4. S.C.Gupta and V.K. Kapoor (2007).”FUNDAMENTALS OF MATHEMATICAL STATISTICS”.Sultan and sons, New Delhi.

5.Joachim Kaldasch (2014). “Evolutionary Model of the City Size Distribution”.Hindawi Publishing Corporation, ISRN Economics, Volume 2014, Article ID 498125, 6 pages.

6. ParimalMukhopadhyay (2009). “MATHEMATICAL STATISTICS”. Books and allied (P) Limited,Cacutta.

7. V.K. Rohatgi (1985). “AN INTRODUCTION TO PROBABILITY THEORY AND MATHEMATICAL STATISTICS”.John Wiley and sons, Inc.

8. H.C. Saxena, P.U. Surendran (1990).”STATISTICAL INFERENCE”. Sultan chand and company, Ltd., New Delhi.

9. S. SuddenduBiswas, (2011). “MATHEMATICAL STATISTICS”.Narosa publishing house, New Delhi.