Jour. Ind. Soc. Ag. Statistics
Vol. XXXVII, No. 3 (1985), pp. 219-225

ON GENERALIZED LOGARITHMIC SERIES DISTRIBUTION

AMARENDRA MISHRA and DAMODAR TIWARY
Patna University, Patna-800 005

(Received : March, 1980)

SUMMARY

The generalized logarithmic series distribution (GLSD) has been defined also for the negative values of the parameters. The Jani and Shah's estimation procedure has been critically examined and an alternative estimation procedure has been suggested which is more efficient than the Jani and Shah's estimators in a certain range. The asymptotic efficiencies of these estimators relative to Jani and Shah's estimators have been derived and computed for certain values of the parameters. A relationship between the moments of the GLSD and those of the generalized negative binomial distribution (GNBD) has been obtained.

Keywords : Logarithmic series distribution, asymptotic efficiency, Negative Binomial distribution.

Introduction

A generalization of the logarithmic series distribution (LSD) with two parameters has been considered by Jain and Gupta [2], Jani [3] and Mishra [6] which provides superior fits to the observed distributions where the LSD is well fitted and satisfactory fits to the data which are not well fitted by the LSD. For a random variable x, the generalized logarithmic series distribution (GLSD) is defined by its mass function

$$P(x, \alpha, \beta) = \theta \binom{x\beta - 1}{x - 1} \frac{\alpha^x}{x} (1 - \alpha)^{x\beta - x} \tag{1.1}$$

where $\theta = -[\log (1 - \alpha)]^{-1}$, $0 < \alpha < 1$, $\alpha\beta < 1$,

$$x = 1, 2, 3, \ldots \text{ and } P(x, \alpha, \beta) = 0 \text{ if } x\beta - x + 1 \leqslant 0 \tag{1.2}$$

Recently, Mishra [6] and Jani and Shah [4] have discussed the ML and moments method of estimation of the two parameters α and β of the GLSD. In this paper it is observed that the GLSD can be defined for also the negative values of α and β such that $\alpha\beta < 1$ and thus the assumption of α and β being positive may be dropped. Jani and Shah's procedure for estimating the parameters by the method of moments has been discussed. Their procedure of estimation fails in a certain situation and hence a new estimation—procedure which is more efficient in a certain range has been suggested. A relationship between the moments of the GLSD and those of the generalized negative binomial distribution (GNBD) of Jain and Consul [1] has also been obtained.

2. GLSD for Negative Values of Parameters

Putting $\alpha = -a$ and $\beta = -b$ in (1.1) where a, b are positive quantities, we get

$$P(x, -a, -b) = \binom{xb+x-1}{x-1} \frac{1}{\log(1+a)} \left(\frac{a}{1+a}\right)^x \left(\frac{1}{1+a}\right)^{xb} \left(\frac{1}{x}\right) \tag{2.1}$$

Now if we substitute

$$\alpha_1 = \frac{a}{1+a} \quad \text{and} \quad \beta_1 = 1 + b \tag{2.2}$$

we obtain for $0 < \alpha_1 < 1$, $\alpha_1\beta_1 < 1$ and $\theta_1 = -[\log(1-\alpha_1)]^{-1}$

$$P(x, -a, -b) = \theta_1 \binom{x\beta_1 - 1}{x-1} \frac{\alpha_1^x}{x} (1-\alpha_1)^{x\beta_1 - x} \tag{2.3}$$

which is of the same form as of (1.1). Hence the assumption of α, β being positive may be dropped and the GLSD may be defined for also negative values of α and β such that $\alpha\beta < 1$. It is noted that same set of probabilities is obtained from (1.1) and (2.3) if the two sets of parameters $(-a, -b)$ and $(\alpha_1\beta_1)$ are related as in (2.2). It can also be seen that whereas for each set of negative values of the parameters $(-a, -b)$ there exists a corresponding set of positive values of the parameters $(\alpha_1\beta_1)$, for each set of positive values of the parameters $(\alpha_1\beta_1)$ the corresponding set of negative values of the same exists only if $\beta_1 > 1$ otherwise $b = \beta_1 - 1$ is negative which contradicts the assumption of b being positive.

3. Jani and Shah's Method of Estimation

The mean and variance of the GLSD are

$$\mu_1' = \frac{\alpha\theta}{1-\alpha\beta} \tag{3.1}$$

$$\mu_2' = \alpha\theta(1 - \alpha - \alpha\theta - \theta\beta\alpha^2)/(1 - \alpha\beta)^3 \qquad (3.2)$$

and hence

$$\alpha^2\theta^2 - Q(1 - \alpha) = 0 \qquad (3.3)$$

where

$$Q - \mu_1'^3/(\mu_1'^2 + \mu_2) = \alpha^2\theta^2/(1 - \alpha) \qquad (3.4)$$

Jani and Shah solved the equation (3.3) for α by the method of iteration after replacing μ_1' and μ_2 by their respective estimates sample mean $\bar{x}$ and sample variance S^2 of the observed data in the expression for Q. For obtaining the initial value of α, they expanded θ into power series of α and neglected the higher powers of α which gave

$$\alpha^2 - 12(Q' - 1)(1 - \alpha) = 0 \qquad (3.5)$$

where $Q' = \bar{x}^3/(\bar{x}^2 + S^2)$

(i) Jani and Shah have mentioned that the term α^3 and the terms of higher power than this have been neglected. But in the expansion of (3.4) the coefficient of α^3 becomes zero and hence it does not appear in the power-series of α. This point has perhaps remained unobserved by them. The term α^4 appears with coefficient $-1/240$ and hence the terms having the absolute value $\alpha^4/240$ and smaller than this have been neglected. This means that the value of α is sufficiently near to the true value given by (3.3) and hence one may leave the iteration if not very much accurate result or a quick result is required.

(ii) For comparing the asymptotic efficiency relative to the ML estimators, Jani and Shah obtained different elements m_{rs}; $r, s = 1, 2$, of the asymptotic variance — covariance matrix M to the order N^{-1} of the moment estimators obtained from (3.3). These elements are in very messy forms. However, to obtain the efficiency, the value of $|M|$, the determinant of the matrix M is required. The value of $|M|$ is

$$|M| = [\theta\alpha(1 - \alpha)/\mu_1'^2\mu_2(2 - \alpha - 2\alpha\theta)]^2 (\mu_2\mu_4 - \mu_2^3 - \mu_3^2) \qquad (3.6)$$

(iii) It may also be noted from (3.4) that one of the properties of the GLSD is that for $0 < \alpha < 1$ the ratio $\mu_1'^3/\mu_2'$ (2.4) is always greater than one, for

$$\mu_1'^3/\mu_2' = \alpha^2\theta^2/(1-\alpha) = (1+\alpha+\alpha^2+\alpha^3+\ldots)/(1+\alpha/2+\alpha^2/3+\ldots)^2 \qquad (3.7)$$

and the coefficient of $\alpha^r (r = 2, 3, \ldots)$ in the denominator is less than unity, the coefficient of α^r in the numerator and equal to unity for $r = 0$ and 1. It is, however, not the peculiarity of the GLSD alone. It is true of the LSD too. But curiously enough this property of the LSD has gone unobserved and unidentified so far. It may be of interest to identify the characteristic which this ratio measures or represents.

Although the ratio $\mu_1'^3/\mu_2'$ is always greater than one, the case of this ratio for the sample moments (Q') being less than one cannot be ruled out. The ratio may be less than one due to sampling fluctuation and in that case Jani and Shah's estimation-procedure fails. An alternative estimation procedure is suggested which does not suffer from this limitation.

4. Alternative Estimators

An estimation-procedure using the third moment of the GLSD is suggested which does not require the application of iteration method for the solution of any equation. The third moment of the GLSD is given by

$$\mu_3' = \theta\alpha(1-\alpha)\,[1 - 2\alpha + \alpha\beta(2-\alpha)]/(1-\alpha\beta)^5 \tag{4.1}$$

From (3.1), (3.2) and (4.1) we have

$$(2-\alpha)^2/(1-\alpha) = (\mu_1'\mu_3' - 3\mu_2'^2)^2/\mu_1'\mu_2'^3 \tag{4.2}$$

The two roots of α are given by

$$\alpha = 1 - A/2 \pm \sqrt{A^2/4 - 1} \tag{4.3}$$

where $A = -2 + [\mu_1'\mu_3' - 3\mu_2'^2]^2/\mu_1'\mu_2'^3$

The roots will be real only if $A \geqslant 2$ i. e. if

$$(\mu_1'\mu_3' - 3\mu_2'^2)^2/\mu_1'\mu_2'^3 \geqslant 4$$

The estimator of α, α^* may be obtained by replacing the population moments by the respective sample moments and from (3.1) the estimator of β, β^* as

$$\beta^* = \frac{1}{\alpha^*} - \frac{\theta^*}{\bar{x}} \tag{4.4}$$

It is verified that two sets of values of parameters, one positive and another negative obtained from (4.3) and (4.4) are related as in (2.2). The same case is with the two sets of estimates obtained by Jani and Shah's procedure.

Using the differential method (5) the asymptotic variance-covariance matrix M^* of the estimators α^* and β^* has been obtained to the order N^{-1}

as $M^* = N^{-1}[m_{rs}]$. The elements $m_{r,s} : r, s = 1, 2$ are given by

$$m_{11} = \sum_{i=1}^{3} \sum_{j=1}^{3} A_i A_j \sigma_{ij}$$

$$m_{12} = m_{21} = D_1 m_{11} + D_2 \sum_{j=1}^{3} A_j \sigma_{ij} \tag{4.5}$$

$$m_{22} = 2D_1 m_{12} - D_1^2 m_{11} + D_2^2 \sigma_{11}$$

where

$$A_1 = (1-\alpha)(2-\alpha)(\mu_1'\mu_3' + 3\mu_2'^2)/\alpha\mu_1' (\mu_1'\mu_3' - 3\mu_2'^2)$$
$$A_2 = -3(1-\alpha)(2-\alpha)(\mu_1'\mu_3' + \mu_2'^2)/\alpha\mu_2'(\mu_1'\mu_3' - 3\mu_2'^2)$$
$$A_3 = 2(1-\alpha)(2-\alpha)\mu_1'/\alpha(\mu_1'\mu_3' - 3\mu_2'^2) \tag{4.6}$$
$$D_1 = \theta^2/\mu_1'(1-\alpha) - 1/\alpha^2$$
$$D_2 = \theta/\mu_1'^2$$
$$\sigma_{ij} = \mu_{i+j}' - \mu_i \mu_j'$$

5. Comparision of Asymptotic Efficiency

The joint asymptotic efficiency of the suggested estimators (α^*, β^*) relative to the Jani and Shah's estimators, say $(\bar{\alpha}, \bar{\beta})$ is given by

$$E = |M|/|M^*| \tag{5.1}$$

The asymptotic efficiency E has been computed in the restricted sample space $1 - (1-\alpha)/\alpha\theta \leqslant \alpha\beta \leqslant 1$ for $\beta = .81 (.02) .89, .93, 1.00, 1.03, 2.03$ and $\alpha = .1(.2).7$ and tabulated in Table 1

TABLE 1—ASYMPTOTIC EFFICIENCIES (%) OF THE ESTIMATORS (α^*, β^*) RELATIVE TO $(\bar{\alpha}, \bar{\beta})$

β/α	0.1	0.3	0.5	0.7
0.81	108.92	124.70	140.70	158.68
0.83	105.30	117.08	131.67	141.27
0.85	102.10	110.14	120.83	127.97
0.87	99.42	105.37	111.62	116.13
0.89	97.06	99.85	102.17	103.94
0.93	93.19	83.26	84.24	87.65
1.00	87.95	66.37	47.67	32.24
1.03	86.07	61.39	42.73	28.61
2.03	51.30	26.21	—	—

The table shows that (α^*, β^*) are more efficient than $(\bar{\alpha}, \bar{\beta})$ for $\beta \leqslant .89$ except the cases $\alpha = .1$, $\beta = .87$, $.89$ and $\alpha = .3$, $\beta = .89$. It may be noted from (1.2) that for $.81 \leqslant \beta \leqslant .89$ the number of classes with non-zero probability would be from 5 to 9. Hence if the GLSD is to be fitted to the cases where the number of observed classes is between 5 and 9 the alternative estimation-procedure is expected to give more efficient result than Jani and Shah's procedure.

6. Relationship between Moments of GLSD and those of GNBD

Jain and Consul (1) defined a generalized negative binomial distribution (GNBD) for $0 < \alpha < 1$, $\alpha\beta < 1$, $n > 0$ as

$$P(x, n, \alpha, \beta) = \frac{n\Gamma(n + x\beta)}{x!\,\Gamma(n + x\beta - x + 1)}\,\alpha^x(1 - \alpha)^{n+x\beta-x} \qquad (6.1)$$

$$x = 0, 1, 2, \ldots$$

such that $P(x, n, \alpha, \beta) = 0$ for $x \leqslant m$ if $n + m\beta < 0$.

We have the rth moment of the GLSD

$$\mu_r' = \theta \sum_{x=1}^{\infty} x^r \frac{\Gamma(x\beta)}{x!\Gamma(x\beta - x + 1)}\,\alpha^x(1 - \alpha)^{x\beta-x} \qquad (6.2)$$

$$= \theta\alpha \sum_{x=0}^{\infty} (1 + x)^{r-1} \frac{\Gamma(x\beta + \beta)}{x!\,\Gamma(x\beta + \beta - x)}\,\alpha^x(1 - \alpha)^{x\beta-x+\beta-1}$$

$$= \theta\alpha \sum_{j=0}^{r-1} \binom{r-1}{j} \left[\sum_{x=0}^{\infty} x^j \frac{\beta + x\beta - 1}{\beta - 1} \cdot \frac{\beta - 1}{\beta + x\beta - 1} \cdot \frac{\Gamma(x\beta + \beta)}{x!\,\Gamma(x\beta + \beta - x)}\,\alpha^x(1 - \alpha)^{x\beta+\beta-x-1} \right]$$

$$\text{for } \beta \neq 1$$

$$= \theta\alpha \sum_{j=0}^{r-1} \binom{r-1}{j} \sum_{x=0}^{\infty} x^j \frac{\beta - 1}{\beta + x\beta - 1} \cdot \frac{\Gamma(x\beta + \beta)}{x!\,\Gamma(x\beta + \beta - x)}\,\alpha^x (1 - \alpha)^{x\beta-\beta-x-1} + \sum_{x=0}^{\infty} x^{j+1} \frac{\beta}{\beta - 1} \cdot \frac{\beta - 1}{\beta + x\beta - 1} \cdot \frac{\Gamma(x\beta + \beta)}{x!\,\Gamma(x\beta + \beta - x}\,\alpha^x(1 - \alpha)^{x\beta-\beta-x-1} \Big]$$

$$= \theta\alpha \sum_{j=0}^{r-1} \binom{r-1}{j} \left[\sum_{x=0}^{\infty} x^j\, P(x, \beta - 1, \alpha, \beta) + \sum_{x=0}^{\infty} x^{j+1} \frac{\beta}{\beta - 1}\, P(x, \beta - 1, \alpha, \beta) \right]$$

The expression within bracket is the sum of the jth and $(j+1)$th moments of the GNBD with $\beta - 1$ in place of n. Denoting the rth moment of the GNBD (6.1) bp $M'_r(n)$ we have the rth moment of the GLSD

$$\mu'_r = \alpha\theta \sum_{j=0}^{r-1} \binom{r-1}{j} \left[M'_j(\beta - 1) + \beta/(\beta - 1) M'_{j+1}(\beta - 1) \right] \quad (6.3)$$

$$r = 1, 2, 3, \ldots$$

$$\beta \neq 1$$

In particular, we have

$$\begin{aligned} \mu'_1 &= \alpha\theta\,(1 + \lambda M'_1) \\ \mu'_2 &= \alpha\theta\,(1 + \overline{\lambda + 1}\, M'_1 + \lambda M'_2) \\ \mu'_3 &= \alpha\theta\,(1 + \overline{\lambda + 2}\, M'_1 + \overline{2\lambda + 1}\, M'_2 + \lambda M'_3) \\ \mu'_4 &= \alpha\theta\,(1 + \overline{\lambda + 3}\, M'_1 + \overline{3\lambda + 1}\, M'_2 + \overline{3\lambda + 1}\, M'_3 + M'_4) \end{aligned} \quad (6.4)$$

where $\lambda = \beta/(\beta - 1)$ and $M'_r = M'_r(\beta - 1)$

Jain and Consul obtained the first four moments of the GNBD in which substituting $n = \beta - 1$, M'_r may be obtained. These relationships may be used for obtaining the moments of the GLSD knowing the moments of the GNBD.

REFERENCES

[1] Jain, G. C. and Consul, P. C. (1971) : A generalised negative binomial distribution, *Siam. Journ. Appl. Math.*, **21** : 501-513.

[2] Jain, G. C. and Gupta, R. P. (1973) : A logarithmic type distribution, *Trabjos Estadist*, **24** : 99-105.

[3] Jani, P. N. (1977) : Minimum variance unbiased estimation for some left-truncated modified power-series distributions, *Sankhya B*, **39** : 258-278.

[4] Jani, P. N. and Shah, S.M. (1979) : On fitting of the generalized logarithmic series distribution, *Journ. Ind. Soc. Agri. Stat.*, **31**(3) : 1-10.

[5] Kendall, M. G. and Stuart, A. (1969) : *The Advanced Theory of Statistics*, vol. 1, Charle's Griffin and Co., London.

[6] Mishra, A. (1979) : Generalizations of some discrete distributions, Ph.D. *thesis*, Patna University, Patna.