Available online at www.ijpab.com

DOI: http://dx.doi.org/10.18782/2582-2845.8689

ISSN: 2582 – 2845

Ind. J. Pure App. Biosci. (2021) *9*(3), 151-155

Indian Journal of Pure & Applied Biosciences

Peer-Reviewed, Refereed, Open Access Journal

Research Article

On Methods of Estimation for Generalized Logarithmic Series Distribution and Its Application to Counts of Red Mites on Apple Leaves

Fehim J Wani[1*], T. A. Raja[1], S. Maqbool[2], M. Iqbal Jeelani[3] and Farheen Naqash[4]

[1]Division of Agricultural Economics & Statistics, FoA, Wadura, SKUAST-Kashmir

[2]Division of Animal Genetics and Breeding, FVSc & AH, SKUAST-Kashmir

[3]Division of Statistics and Computer Science, Chatha, SKUAST-Jammu

[4]School of Agricultural Economics & Horti-Business Management, FoH, Shalimar, SKUAST-Kashmir

*Corresponding Author E-mail: faheemwani@skuastkashmir.ac.in

Received: 19.04.2021 | Revised: 23.05.2021 | Accepted: 1.06.2021

ABSTRACT

The Generalized Logarithmic Series Distribution (GLSD) adds an extra parameter to the usual logarithmic series distribution and was introduced by Jain and Gupta (1973). This distribution has found applications in various fields. The estimation of parameters of generalized logarithmic series distribution was studied by the methods of maximum likelihood, moments, minimum chi square and weighted discrepancies. The GLSD was fitted to counts of red mites on apple leaves and it was observed that all the estimation techniques perform well in estimating the parameters of generalized logarithmic series distribution but with varying degree of non-significance.

Keyword: *GLSD, Estimation, Parameters.*

INTRODUCTION

The generalized logarithmic series distribution (GLSD) characterized by two parameters α and β was first obtained by Jain and Gupta (1973) and its probability function is given by:

$$P(X=x) = -\frac{1}{\log(1-\alpha)}\frac{\Gamma(\beta x)}{x!\Gamma(\beta x - x + 1)}\alpha^{x}(1-\alpha)^{\beta x - x} \qquad ; \; x = 1,2,3,\ldots$$

$$0 < \alpha < 1, \; |\alpha\beta| < 1$$

Since GLSD is generalization of logarithmic series distribution (LSD). GLSD will reduce to logarithmic series distribution (LSD) when taking $\beta = 1$.

Cite this article: Wani, F. J., Raja, T. A., Maqbool, S., Iqbal Jeelani, M., & Naqash, F. (2021). On Methods of Estimation for Generalized Logarithmic Series Distribution and Its Application to Counts of Red Mites on Apple Leaves, *Ind. J. Pure App. Biosci. 9*(3), 151-155. doi: http://dx.doi.org/10.18782/2582-2845.8689

GLSD is a member of Consul and Shenton (1972) family of Lagrangian probability distribution and also Gupta (1974) modified power series distribution (MPSD). The same distribution has been obtained by many more authors. The distribution has found applications in various fields. Jain and Gupta (1973) applied it to the William's data on number of papers by entomologists, Rao (1981) applied it to the study of correlation between two types of children in a family, and Hansen and Willenkens (1990) used it in the risk theory in a problem related to the total claim size upto time t. The estimation of GLSD has been studied by many researchers, where as Gupta (1974) and Jani (1977) examined its minimum variance unbiased (MVU) estimation. Mishra (1979) discussed its maximum likelihood (ML) estimation, Jani and Shah (1979) discussed the method of moments of its estimation. Wani et al (2016) compared lagrangian probability distributions for counts of red mites on apple leaves in Kashmir valley.

ESTIMATION OF PARAMETERS

The various methods of estimation for estimating the parameters of Generalized Logarithmic Series distribution are as follows:

Maximum likelihood Estimation of GLSD

Consider a random sample of size N taken from the GLSD and let the observed frequencies be $f_x; x = 1,2...k$ so that $\sum_{x=1}^{k} f_x = N$ where k is the largest of the observed values having non-zero frequencies. The likelihood equation of the GLSD can be written as

$$L = \frac{\theta^N \alpha^{N\bar{x}} \prod_{x=1}^{k}\prod_{j=1}^{x-1}(\beta x - j)^{f_x} (1-\alpha)^{(\beta-1)N\bar{x}}}{\prod_{i=1}^{k}(X_i!)^{f_x}}$$

The log likelihood function is given as

$$LogL = N\log\theta + N\bar{x}\log\alpha + \sum_{x=1}^{k}\sum_{j=1}^{x-1} f_x \log(\beta x - j) + (\beta-1)N\bar{x}\log(1-\alpha) - \sum f_i \log X_i$$

The two likelihood functions can be obtained as

$$\frac{\partial LogL}{\partial\alpha} = -\frac{N\theta}{(1-\alpha)} + \frac{N\bar{x}}{\alpha} - \frac{(\beta-1)N\bar{x}}{(1-\alpha)} = 0 \qquad (3.1.2.1)$$

$$\frac{\partial LogL}{\partial\beta} = \frac{-N\bar{x}}{\theta} + \sum_{x=1}^{k}\sum_{j=1}^{x-1}\frac{xf_x}{\beta x - j} = 0 \qquad (3.1.2.2)$$

Where $\bar{x}$ is the sample mean. From equation $(3.1.2.2)$, we get

$$\beta = \frac{1}{\alpha} - \frac{\theta}{\bar{x}} \qquad (3.1.2.3)$$

Putting this equation in $(3.1.2.2)$, we get

$$\varphi(\alpha) = \frac{N\bar{x}}{\theta} + \sum_{x=1}^{k}\sum_{j=1}^{x-1}\frac{xf_x}{\left(\frac{1}{\alpha} - \frac{\theta}{\bar{x}}\right)x - j} = 0 \qquad (3.1.2.4)$$

Which can be solved for $\hat{\alpha}$ and $\hat{\beta}$, the M.L. estimators of α and β respectively.

Moment Estimation of GLSD

The mean, the variance and the recurrence relation for higher moments is given by

$$\mu_1 = \frac{\alpha\theta}{(1-\beta\theta)} \qquad (3.1.4.1)$$

$$\mu_2 = \frac{\alpha\theta(1-\theta-\alpha\theta+\alpha\beta\theta^2)}{(1-\beta\theta)^3} \qquad (3.1.4.2)$$

$$\mu_{r+1} = \frac{\theta(1-\theta)}{1-\beta\theta}\frac{\partial\mu_r}{\partial\theta} + r.\mu_3\mu_{r-1} \qquad (3.1.4.3)$$

Since $\mu_2 < 0$ for $\beta\theta < 1-(1-\theta)/\alpha\theta$, we will obtain the moment estimators of the parameters of the GLSD for the sample space $1-(1-\theta)/\alpha\theta \le \beta\theta < 1$.

From equation $(3.1.4.1)$ and $(3.1.4.2)$, we have

$$\alpha^2\theta^2 - Q(1-\theta) = 0 \qquad (3.1.4.4)$$

where

$$Q = \mu^3/(\mu^2+\mu_2) \qquad (3.1.4.5)$$

The equation $(3.1.4.4)$ can be solved for θ by using the method of iterations. To obtain the initial value of θ expanding $\alpha^2\theta^2 - Q(1-\theta)$ into power series expansion and neglecting θ^3 and terms higher than that we get $(3.1.4.4)$ as

$$\theta^2 - 12(Q-1)(1-\theta) = 0$$

which gives, for $\theta > 0$

$$\theta = 2\left(\sqrt{Q'(Q'+1)}\right) - Q' \qquad (3.1.4.6)$$

where, $Q' = 3(Q-1)$ and Q is given by equation $(3.1.4.5)$

To get θ^*, the moment estimator of θ, μ and μ_2 are to be replaced by their respective estimates sample mean $\overline{X}$ and sample variance S^2 of the observed data.

Using equation $(3.1.4.1)$ and replacing μ by $\overline{X}$, we get

$$\beta^* = 1/\theta^* - \alpha^*/\overline{X}$$

Minimum chi-square for GLSD

Since

$$\chi^2 = \sum_{x=1}^{k}\frac{(n_x-p_x)^2}{p_x}$$

is approximately distributed as chi-square. Differentiating it with respect to α and β, we obtain

$$\sum_{x=1}^{k}(n_x-p_x)\left(1+\frac{n_x}{p_x}\right)\frac{\partial}{\partial\alpha}\log p_x = 0 \qquad (3.1.6.1)$$

$$\sum_{x=1}^{k}(n_x-p_x)\left(1+\frac{n_x}{p_x}\right)\frac{\partial}{\partial\beta}\log p_x = 0 \qquad (3.1.6.2)$$

Substituting the corresponding derivatives to $(3.1.6.1)$ and $(3.1.6.2)$ we get

$$\sum_{x=1}^{k}(n_x-p_x)\left(1+\frac{n_x}{p_x}\right)\left[\frac{N}{(1-\alpha)\log(1-\alpha)} + \frac{N\bar{x}}{\alpha} - \frac{(\beta-1)N\bar{x}}{(1-\alpha)}\right] \qquad (3.1.6.3)$$

$$\sum_{x=1}^{k}(n_x - p_x)\left(1+\frac{n_x}{p_x}\right)\left[N\bar{x}\log(1-\alpha)+\sum_{x=1}^{k}\sum_{j=1}^{x-1}\frac{xn_x}{(\beta x - j)}\right] \quad (3.1.6.4)$$

The resulting equations $(3.1.6.3)$ and $(3.1.6.4)$ are known as minimum chi-square equations.

Weighted discrepancies (WD) method

Let f_x denote the observed frequencies $x = 0,1,2,\ldots k$. Obviously, k is the largest of the observations. Let $N = \sum_{x=1}^{k} f_x$.

The corresponding relative frequencies are given by

$$n_x = \frac{f_x}{N}; \qquad x = 0,1,2,\ldots k \quad (3.1.7.1)$$

The log likelihood function can be written as

$$\log L = \sum_{x=1}^{k} Nn_x \log p_x(\alpha,\beta) \quad (3.1.7.2)$$

The likelihood equations are

$$\sum_{x=1}^{k} n_x \frac{\partial}{\partial\alpha}\log p_x = 0 \quad (3.1.7.3)$$

$$\sum_{x=1}^{k} n_x \frac{\partial}{\partial\beta}\log p_x = 0 \quad (3.1.7.4)$$

Again as $\sum_{x=1}^{k} p_x = 1$ $(3.1.7.5)$

$$\sum_{x=1}^{k} p_x \frac{\partial}{\partial\alpha}\log p_x = 0 \quad (3.1.7.6)$$

$$\sum_{x=1}^{k} p_x \frac{\partial}{\partial\beta}\log p_x = 0 \quad (3.1.7.7)$$

Subtracting $(3.1.7.6)$ from $(3.1.7.3)$ and $(3.1.7.7)$ from $(3.1.7.4)$, we get

$$\sum_{x=1}^{k}(n_x - p_x)\frac{\partial}{\partial\alpha}\log p_x = 0 \quad (3.1.7.8)$$

$$\sum_{x=1}^{k}(n_x - p_x)\frac{\partial}{\partial\beta}\log p_x = 0 \quad (3.1.7.9)$$

Substituting the corresponding expressions of the derivatives to $(3.1.7.8)$ and $(3.1.7.9)$, we get

$$\sum_{x=1}^{k}(n_x - p_x)\left[\frac{N}{(1-\alpha)\log(1-\alpha)}+\frac{N\bar{x}}{\alpha}-\frac{(\beta-1)N\bar{x}}{(1-\alpha)}\right] = 0 \quad (3.1.7.10)$$

$$\sum_{x=1}^{k}(n_x - p_x)\left[N\bar{x}\log(1-\alpha)+\sum_{x=1}^{k}\sum_{j=1}^{x-1}\frac{xf_x}{(\beta x - j)}\right] = 0 \quad (3.1.7.11)$$

which has referred to Kemp (1986) as an equation from minimum discrimination information and ML estimation and called as weighted discrepancies estimation method.

NUMERICAL IILUSTRATION

The GLSD was fitted to count of red mites on apple leaves.

Table 1: Comparison of Observed frequencies with Expected frequencies of GLSD for Count of red mites on Apple Leaves

No. of mites/leaf	Observed Frequency	Expected Frequency of GLSD			
		Methods of Estimation			
		ML	Moments	MC	WD
3	39	44.10	41.50	42.20	41.20
4	22	18.30	19.50	19.40	19.40
5	17	15.20	15.70	15.25	16.10
6	15	12.70	11.90	12.85	12.50
7	9	8.30	9.90	9.70	9.70
8	6	5.40	5.20	5.00	5.40
9	3	4.10	4.00	3.50	4.20
10	1	1.90	2.20	2.00	1.80
≥11	0	2.00	2.10	2.10	1.70
Total	112	112.00	112.00	112.00	112.00
Parameter (Estimates) $\hat{\alpha}$		0.781	0.715	0.725	0.703
$\hat{\beta}$		0.767	0.742	0.761	0.735
χ^2		3.591	3.185	3.218	3.062
p-value		0.892	0.922	0.920	0.930

In the above table the estimates of parameters $\hat{\alpha}$ and $\hat{\beta}$ in different methods of estimation are ML (0.781, 0.767), Moments (0.715, 0.742), MC (0.725, 0.761) and WD (0.703, 0.735). The p-value of all the methods are non-significant and in agreement with the observed frequencies. We observed that all the estimation techniques, the ML, the moment, the MC and the WD method perform well in estimating the GLSD parameters.

REFERENCES

Consul, P. C., & Shenton, L. R. (1972). Use of Lagrangian expression for generating generalized probability distributions. *SIAM J. Appl. Math*, *23*(2), 239-249.

Gupta, R. C. (1974). Modified power series distribution and some of its applications, *Sankhya. Ser. B 35*, 288-298.

Hanseen, B. B., & Willekens, E. (1990). The generalized logarithmic series distribution, *Statistics and Probability Letters 9*, 311-316.

Jani, P. N. (1977). Minimum variance unbiased estimate for some left truncated modified power series distribution. *Sankhya, Series B 3*(39), 258-278.

Jain, G. C., & Gupta, R. P. (1973). A logarithmic type distribution. *Trabjos Estadist*, *24,* 99 -105.

Jani, P. N., & Shah, S. M. (1979). On fitting of the generalized logarithmic series distribution. *Journal of the Indian Statistical Association 30*(3), 1-10.

Kemp, A. W. (1986). Weighted discrepancies and maximum likelihood estimation for discrete distributions. *Communication in Statistics - Theory and Methods 15*(3), 783-803.

Mishra, A. (1979). Generalization of some discrete distributions. *J Bihar. Math. Soc 11*, 12-22.

Rao, B. R. (1981). Correlation between the numbers of two types of children in a family with the mpsd for the family size. *Communications in Statistics – Theory and Methods 10*(3), 249-254.

Wani, F. J., Raja, T. A., Maqbool, S., Khan, I., Bhat, M. A., & Jeelani, M. I. (2016). A comparison of generalized logarithmic series distribution, generalized poisson distribution and generalized negative binomial distribution for counts of red mites on apple leaves in kashmir valley. *Int. J. Agricult. Stat. Sci 12*(1), 117-120.