Visit us - www.researchjournal.co.in ■ DOI : 10.15740/HAS/IRJAES/8.1/60-64
International Research Journal of Agricultural Economics and Statistics
Volume 8 | Issue 1 | March, 2017 | 60-64 ■ e ISSN-2231-6434

Research Paper

On generalized logarithmic series distribution and its application to leaf spot grades in mulberry

■ FEHIM JEELANI WANI, T.A. RAJA, SHOWKET MAQBOOL, IMRAN KHAN AND M. IQBAL JEELANI BHAT

See end of the paper for authors' affiliations

Correspondence to :
FEHIM JEELANI WANI
Division of Agricultural Statistics, Faculty of Horticulture, Shalimar, Sher-e-Kashmir University of Agricultural Sciences and Technology (K), SRINAGAR (J&K) INDIA
Email : faheemwani149@gmail.com

Paper History :
Received : *06.01.2017;*
Revised : *17.01.2017;*
Accepted : *27.01.2017*

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 different leaf spot grades in four varieties of mulberry namely Ichinose, Gosherami, Rokokuvoso and Kokuso-20 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.

KEY WORDS : Generalized logarithmic series distribution, Leaf spot grades, Mulberry

HOW TO CITE THIS PAPER : Wani, Fehim Jeelani, Raja, T.A., Maqbool, Showket, Khan, Imran and Bhat, M. Iqbal Jeelani (2017). On generalized logarithmic series distribution and its application to leaf spot grades in mulberry. *Internat. Res. J. Agric. Eco. & Stat.*, **8** (1) : 60-64, **DOI : 10.15740/HAS/IRJAES/8.1/60-64.**

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:

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

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

Since GLSD is generalization of logarithmic series distribution (LSD). GLSD will reduce to logarithmic series distribution (LSD) when taking β=1. 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 Hanseen 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, whereas 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.

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

Put β=1

$$p(X=x) = -\frac{1}{\log(1-\alpha)}\frac{\sqrt{x}}{x!\sqrt{1}}\alpha^x(1-\alpha)^0$$

$$= -\frac{1}{\log(1-\alpha)}\frac{(x-1)!}{x(x-1)!}\alpha^x$$

So one parameter LSD is a particular case of GLSD

$$= -\frac{1}{x\log(1-\alpha)}\alpha^x$$

MATERIALS AND METHODS :

For the present study mulberry data set was utilized. Four different varieties of mulberry namely Ichinose, Goshoerami, Rokokuyoso and Kokuso having different leaf spot disease intensity were chosen for the study. Three trees of each variety were selected at random, from each tree, three branches were selected at random and then from each branch the spots of the disease were recorded from all the leaves. The leaves with no spot were referred to as disease free and named as grade zero (0 grade). The leaves having 1 to 5, 6 to 10, 11 to 15, 16 to 20 and more than 20 spots were total graded as 1, 2, 3, 4 and 5 grade, respectively. The number of leaves in three branches per tree falling in different grades was recorded. Similarly for all the four varieties, the number of leaves in three branches falling under different grades were recorded. All the leaves of the three randomly selected branches per plant were observed for different leaf spot grading.

Estimation of parameters :

Maximum likelihood estimates :

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

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

The two likelihood functions can be obtained as

$$\frac{\partial \text{Log } L}{\partial\alpha} = -\frac{N\theta}{(1-\alpha)} + \frac{N\bar{x}}{\alpha} - \frac{(\beta-1)N\bar{x}}{(1-\alpha)} = 0 \quad (2)$$

$$\frac{\partial \text{ Log } L}{\partial\beta} = \frac{-N\bar{x}}{\theta} + \sum_{x=1}^{k}\sum_{j=1}^{x-1}\frac{xf_x}{\beta x - j} = 0 \quad (3)$$

where $\bar{x}$ is the sample mean. From equation (3), we get

$$\beta = \frac{1}{\alpha} - \frac{\theta}{\bar{x}} \quad (4)$$

Putting this equ. in (3), we get

$$\varphi(\alpha) = \frac{N\bar{x}}{\theta} + \sum_{x=1}^{k}\sum_{j=1}^{x-1}\frac{xf_x}{\left(\frac{1}{\alpha} - \frac{\theta}{\bar{x}}\right)x - j} = 0 \quad (5)$$

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

Estimates from moments :

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

$$\mu_1 = \frac{\alpha\theta}{(1-\beta\theta)} \quad (6)$$

$$\mu_2 = \frac{\alpha\theta\left(1-\theta-\alpha\theta+\alpha\beta\theta^2\right)}{(1-\beta\theta)^3} \quad (7)$$

$$\mu_{r+1} = \frac{\theta(1-\theta)}{1-\beta\theta}\frac{\partial\mu_r}{\partial\theta} + r.\mu_3\mu_{r-1} \quad (8)$$

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 eq. (6) and (7), we have

$$\alpha^2\theta^2 - Q(1-\theta) = 0 \quad (9)$$

where,

$$Q = \mu^3 / \left(\mu^2 + \mu_2\right) \quad (10)$$

The equ. (9) 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 (9) as

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

which gives, for θ>0

$$\theta = 2\left(\sqrt{Q'(Q'+1)}\right) - Q' \quad (11)$$

where, $Q' = 3(Q-1)$ and Q is given by equ. (10)

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

Using equ. (6) and replacing μ by $\bar{X}$, we get

$$\beta^* = 1/\theta^* - \alpha^*/\bar{X} \tag{12}$$

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

$$\sum_{x=1}^{k} (n_x - p_x)\left(1 + \frac{n_x}{p_x}\right)\frac{\partial}{\partial\alpha}\log p_x = 0 \tag{13}$$

$$\sum_{x=1}^{k} (n_x - p_x)\left(1 + \frac{n_x}{p_x}\right)\frac{\partial}{\partial\beta}\log p_x = 0 \tag{14}$$

Substituting the corresponding derivatives to (13) and (14) we get

$$\sum_{x=1}^{k} (n_x - p_x)\left(1 + \frac{n_x}{p_x}\right)\left[\frac{N}{(1-\alpha)\log(1-\alpha)} + \frac{N\bar{x}}{\alpha} - \frac{(\beta-1)N\bar{x}}{(1-\alpha)}\right] \tag{15}$$

$$\sum_{x=1}^{k} (n_x - p_x)\left(1 + \frac{n_x}{p_x}\right)\left[N\bar{x}\log(1-\alpha) + \sum_{x=1}^{k}\sum_{j=1}^{x-1}\frac{x n_x}{(\beta x - j)}\right] \tag{16}$$

The resulting equ. (15) and (16) are known as minimum chi-square equations.

Weighted discrepancies (WD) method :

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

The corresponding relative frequencies are given by

$$n_x = \frac{f_x}{N}; \quad x = 0,1,2,\ldots k \tag{17}$$

The log likelihood function can be written as

$$\log L = \sum_{x=1}^{k} N n_x \log p_x(\alpha,\beta) \tag{18}$$

The likelihood equations are

$$\sum_{x=1}^{k} n_x \frac{\partial}{\partial\alpha}\log p_x = 0 \tag{19}$$

$$\sum_{x=1}^{k} n_x \frac{\partial}{\partial\beta}\log p_x = 0 \tag{20}$$

Again as $\sum_{x=1}^{k} p_x = 1$ (21)

$$\sum_{x=1}^{k} p_x \frac{\partial}{\partial\alpha}\log p_x = 0 \tag{22}$$

$$\sum_{x=1}^{k} p_x \frac{\partial}{\partial\beta}\log p_x = 0 \tag{23}$$

Subtracting (22) from (19) and (23) from (26), we get

$$\sum_{x=1}^{k} (n_x - p_x)\frac{\partial}{\partial\alpha}\log p_x = 0 \tag{24}$$

$$\sum_{x=1}^{k} (n_x - p_x)\frac{\partial}{\partial\beta}\log p_x = 0 \tag{25}$$

Substituting the corresponding expressions of the derivatives to (24) and (25), we get

$$\sum_{x=1}^{k} (n_x - p_x)\left[\frac{N}{(1-\alpha)\log(1-\alpha)} + \frac{N\bar{x}}{\alpha} - \frac{(\beta-1)N\bar{x}}{(1-\alpha)}\right] = 0 \tag{26}$$

Table 1 : Comparison of observed frequencies with expected frequencies of GLSD for leaf spot grades in Ichinose variety of mulberry

Leaf spot grades	Observed frequency	Expected frequency of GLSD			
		Methods of estimation			
		ML	Moments	MC	WD
0	61	64.70	62.80	64.00	62.50
1	18	16.10	15.30	13.15	15.50
2	11	9.10	10.55	10.45	10.50
3	3	2.00	3.30	3.50	3.40
4	1	1.10	1.10	1.70	1.20
5	0	1.00	0.95	1.20	0.90
Total	94	94.00	94.00	94.00	94.00
Parameters (Estimate)					
$\hat{\alpha}$		0.871	0.867	0.862	0.853
$\hat{\beta}$		0.961	0.921	0.932	0.915
χ^2		1.645	1.305	2.402	1.145
p-value		0.896	0.934	0.791	0.950

Table 2 : Comparison of observed frequencies with expected frequencies of GLSD for leaf spot grades in Gosherami variety of mulberry

Leaf spot grades	Observed frequency	Expected frequency of GLSD			
		Methods of estimation			
		ML	Moments	MC	WD
0	58	62.10	61.20	62.00	61.20
1	17	15.00	15.40	15.10	15.45
2	8	5.80	6.10	6.00	6.20
3	3	2.70	2.75	2.50	2.60
4	2	1.40	1.50	1.30	1.80
5	0	1.00	1.05	1.10	0.75
Total	88	88.00	88.00	88.00	88.00
Parameters (Estimate)					
$\hat{\alpha}$		0.861	0.845	0.855	0.832
$\hat{\beta}$		0.952	0.931	0.940	0.925
χ^2		1.877	1.486	1.826	1.419
p-value		0.866	0.897	0.863	0.915

Table 3 : Comparison of observed frequencies with expected frequencies of GLSD for leaf spot grades in Rokokuvoso variety of mulberry

Leaf spot grades	Observed frequency	Expected frequency of GLSD			
		Methods of estimation			
		ML	Moments	MC	WD
0	68	73.10	70.50	73.85	71.10
1	19	15.30	16.20	13.70	16.10
2	12	11.00	11.80	10.90	11.65
3	5	4.10	4.80	4.60	4.60
4	2	2.30	2.10	2.70	2.25
5	1	1.20	1.60	1.25	1.30
Total	107	107.00	107.00	107.00	107.00
Parameters (Estimate)					
$\hat{\alpha}$		0.870	0.810	0.865	0.862
$\hat{\beta}$		0.955	0.915	0.950	0.941
χ^2		0.798	0.651	1.250	0.422
p-value		0.977	0.986	0.940	0.991

Table 4 : Comparison of observed frequencies with expected frequencies of GLSD for leaf spot grades in Kokuso-20 variety of mulberry

Leaf spot grades	Observed frequency	Expected frequency of GLSD			
		Methods of estimation			
		ML	Moments	MC	WD
0	65	67.00	67.10	66.10	66.10
1	17	18.10	18.30	17.70	18.10
2	10	7.35	7.80	8.00	8.40
3	4	3.10	2.60	3.40	3.10
4	2	2.35	1.95	2.60	2.10
5	1	1.10	1.25	1.20	1.20
Total	99	99.00	99.00	99.00	99.00
Parameters (Estimate)					
$\hat{\alpha}$		0.775	0.785	0.720	0.753
$\hat{\beta}$		0.915	0.852	0.920	0.849
χ^2		0.726	0.424	0.601	0.396
p-value		0.982	0.993	0.988	0.995

$$\sum_{x=1}^{k}(n_x - p_x)\left[N\bar{x}\log(1-\alpha)+\sum_{x=1}^{k}\sum_{j=1}^{x-1}\frac{xf_x}{(\beta x - j)}\right]=0 \quad (27)$$

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

Application :

The GLSD has been found useful in research areas where the logarithmic series distribution has been applied. Jani and Shah (1979) used the GLSD to model the count of European red mites on apple leaves. The different methods of estimation for estimating the parameters of generalized logarithmic series distribution were compared by chi-square goodness of fit test.

In the Table 1 the estimates of parameters $\hat{\alpha}$ and $\hat{\beta}$ in different methods of estimation are ML (0.871, 0.961), Moments (0.867, 0.921), MC (0.862, 0.932) and WD (0.853, 0.915). 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.

In the Table 2 the estimates of parameters $\hat{\alpha}$ and $\hat{\beta}$ in different methods of estimation are ML (0.861, 0.952), Moments (0.845, 0.931), MC (0.855, 0.940) and WD (0.832, 0.925). 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.

In the Table 3 the estimates of parameters $\hat{\alpha}$ and $\hat{\beta}$ in different methods of estimation are ML (0.870, 0.955), Moments (0.810, 0.915), MC (0.865, 0.950) and WD (0.862, 0.941). 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.

In the Table 4 the estimates of parameters $\hat{\alpha}$ and $\hat{\beta}$ in different methods of estimation are ML (0.775, 0.915), Moments (0.785, 0.852), MC (0.720, 0.920) and WD (0.753, 0.849). 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.

Authors' affiliations:

T.A. RAJA, Division of Animal Genetics and Breeding, Faculty of Veterinary Science and Animal Husbandry, Sher-e-Kashmir University of Agricultural Sciences and Technology (K), SUHUMA (J&K) INDIA

SHOWKET MAQBOOL, Division of Agricultural Economics and Statistics, Faculty of Agriculture, Sher-e-Kashmir University of Agricultural Sciences and Technology (K), WADURA (J&K) INDIA

IMRAN KHAN, Division of Agricultural Statistics, Faculty of Horticulture, Shalimar, Sher-e-Kashmir University of Agricultural Science and Technology (K), SRINAGAR (J&K) INDIA

M. IQBAL JEELANI BHAT, Division of Statistics and Computer Science, Sher-e-Kashmir University of Agricultural Sciences and Technology (K), CHATHA (J&K) INDIA

LITERATURE CITED :

Consul, P.C. and 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. and Willekens, E. (1990). The generalized logarithmic series distribution. *Statistics & Probability Letters*, **9** : 311-316.

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

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

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

Kemp, A.W. (1986). Weighted discrepancies and maximum likelihood estimation for discrete distributions. *Communication in Statistics - Theory & 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 & Methods*, **10**(3): 249-254.