Jour. Ind. Soc. Ag. Statistics
54(3), 2001 : 317-323

A Study on the Inflated Poisson Lindley Distribution

M. Borah and A. Deka Nath[1]
Tezpur University, Tezpur
(Received : June, 2000)

SUMMARY

The Poisson Lindley distribution has been further studied with some inflation of probability at zero. Some properties of this Inflated Poisson Lindley (IPL) distribution are discussed. The recurrence relations are obtained without derivatives, so that they will be easy to handle on computer for computation of higher order probabilities, moments, etc. The parameters of this distribution have been estimated by three methods. Examples are given for fitting of this distribution to real data, and the fit is compared with that obtained by using other distributions.

Key words : Poisson-Lindley distribution, Inflated distribution, Recurrence relation, Raw moments, Skewness, Kurtosis, Parameter estimation.

1. Introduction

Poisson Lindley distribution is a generalized poisson distribution (see Consul [5]) originally due to Lindley [10] with probability mass function

$$P_x(\phi) = \frac{\phi^2(\phi + 2 + x)}{(\phi + 1)^{x+3}} \qquad x = 0, 1, 2, \ldots \qquad (1.1)$$

Sankaran [12] further investigated this distribution with application to errors and accidents. In both the examples, single parameter Poisson Lindley distribution gives a better fit than Poisson distribution. It is a special case of Bhattacharya's [2] more complicated mixed poisson distribution. Some mixture of Poisson Lindley distributions derived by using Gurland's generalization [7] were studied by Borah and Deka Nath [4], where certain properties of Poisson-Poisson-Lindley and Poisson-Lindley-Poisson distributions were investigated.

A random variable X is said to have the discrete inflated distribution if its probability function is given by

[1] Department of Statistics, Darrang College, Tezpur, Assam

$$P(X=x)=\begin{cases}\omega+(1-\omega)p_0 & x=0\\ (1-\omega)p_x & x=1,2,3,\ldots\end{cases} \tag{1.2}$$

where ω is a parameter assuming arbitrary values in the interval (0, 1). It is also possible to take $\omega < 0$, provided $\omega + (1-\omega)p_0 \geq 0$ (See Johnson *et al.* [9]).

The discrete inflated distribution was first investigated by Singh [15]. He studied inflated poisson distribution to serve the probabilistic description of an experiment with a slight inflation at a point, say zero. Later Singh ([13], 14]) pointed out that there exists analogous situations in binomial distribution, i.e. distinct increase of frequency of observed event at point zero as well as respective decrease of its value at the remaining points. Pandey [11] studied the generalized inflated poisson distribution. Gerstenkorn [6] established the recurrence relation for the moments for the inflated negative binomial, poisson and geometric distribution.

In this paper, an Inflated Poisson-Lindley (IPL) distribution is discussed to serve the probabilistic description of an experiment with a slight inflation of probability at zero. The recurrence relations for moments and probabilities for IPL distribution are obtained. For fitting of the IPL distribution, three well-known data sets are considered for an empirical comparison and it is observed that this distribution gives better fit in all the cases.

2. *Recurrence Relation for Probabilities*

The probability generating function (p.g.f.), G(t) of IPL distribution may be written as

$$G(t)=\omega+(1-\omega)g(t) \tag{2.1}$$

where $g(t)=\dfrac{\phi^2(\phi+2-t)}{\{(\phi+1)(\phi+1-t)^2\}}$ is the p.g.f. of Poisson Lindley (PL) distribution $0<\omega<1, \phi>0$ (see Sankaran [12]). Differentiating (2.1) w.r.t. 't' and equating the coefficients of t^r from both sides, we have

$$P_r=\frac{(\phi+2+r)}{(\phi+1)(\phi+1+r)}P_{r-1} \qquad r>1 \tag{2.2}$$

where $P_0=\omega+(1-\omega)\phi^2(\phi+2)/(\phi+1)^3$ and $P_1=\dfrac{(1-\omega)\phi^2(\phi+3)}{(\phi+1)^4}$

3. Recurrence Relation for Moments

The raw moments recurrence relation for IPL distribution may similarly be written as

$$\mu_r' = \frac{(1-\omega)\{(\phi+3)-2^r\}}{\phi(\phi+1)} + \sum_{j=0}^{r-1} \frac{\left(3\alpha - 3\times 2^{j+1}\alpha^2 + 2^{j+1}\alpha^3\right)}{(1-\alpha)^3}\binom{r}{j+1}\mu_{r-j}',\ r>1 \quad (3.1)$$

where $\alpha = \dfrac{1}{(\phi+1)}$

$$\mu_1' = \frac{(1-\omega)(\phi+2)}{\{\phi(\phi+1)\}} \text{ and } \mu_2' = \frac{(1-\omega)(\phi^2+4\phi+6)}{\{\phi^2(\phi+1)\}}$$

Thus the variance may be obtained as

$$\mu_2 = \frac{(1-\omega)\{\phi^3+4\phi^2+6\phi+2+\omega(\phi+2)^2\}}{\{\phi^2(\phi+1)^2\}} \quad (3.2)$$

Putting $\omega = 0$ in (3.2) the variance of PL distribution may be obtained (see Borah *et al.* [4]). The expression for the coefficient of skewness and kurtosis can be written in terms of ϕ and ω

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{P}{Q} \quad (3.3)$$

where $P = \{\phi^5 + 7\phi^4 + 22\phi^3 + 32\phi^2 + 18\phi + 4 + \omega(3\phi^4 + 17\phi^3 + 36\phi^2 + 30\phi + 4)$

$$+\omega^2(\phi^3 + 6\phi^2 + 12\phi + 18)\}$$

and $Q = \sqrt{(1-\omega)\{\phi^3 + 4\phi^2 + 6\phi + 2 + \omega(\phi+2)^2\}^3}$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{A + \omega B + 3\omega^2 C + 3\omega^3 D}{(1-\omega)\{\phi^3 + 4\phi^2 + 6\phi + 2 + \omega(\phi+2)^2\}^2} \quad (3.4)$$

where $A = \phi^7 + 2\phi^6 + 73\phi^5 + 174\phi^4 + 256\phi^3 + 152\phi^2 - 24\phi + 12$

$B = 7\phi^6 + 54\phi^5 + 181\phi^4 + 312\phi^3 + 34\phi^2 + 264\phi + 12$

$C = 4\phi^5 + 30\phi^4 + 62\phi^3 + 112\phi^2 + 32\phi$ and

$D = 2\phi^4 + 16\phi^3 + 52\phi^2 + 64\phi + 32$

It is clear from the above expression of γ_1 that for any given value of $\phi > 0$ and ω closes to unity, the skewness is infinitely large and it becomes smaller and smaller as the value of ω decreases. The IPL distribution is easily seen to be leptokurtic as the value of γ_2 is positive for all values of $\phi > 0$ and $0 < \omega < 1$ though there is a factor '-24ϕ' in the numerator of (3.4).

4. Estimation of Parameters

The estimation of parameters of inflated distributions other than ω can be carried out by ignoring the observed frequency in the zero class, and then using a technique appropriate to the original distribution truncated by omission of zero class. After the other parameters have been estimated, parameter ω can then be estimated by equating the observed and expected frequencies in the zero class (See Johnson *et al.* [9]). Three methods for estimating parameters of IPL distribution, i.e. method of maximum likelihood, method of moments and ratio of first two frequencies with mean are discussed in this section.

(a) Method of Maximum Likelihood (ML): Since IPL distribution is a zero modified distribution, one of the ML equations is (see Johnson *et al.* [9])

$$\hat{\omega} + \frac{(1-\hat{\omega})\hat{\phi}^2(\hat{\phi}+2)}{(\hat{\phi}+1)^3} = \frac{n_0}{N} \tag{4.1}$$

where $\frac{n_0}{N}$ is the observed proportion of zeros. It is also a power series distribution so the other ML equation will be

$$\bar{x} = \frac{(1-\hat{\omega})(\hat{\phi}+2)}{\{\hat{\phi}(\hat{\phi}+1)\}} \tag{4.2}$$

Eliminating $\hat{\omega}$ from equation (4.1) and (4.2), we have

$$\frac{\hat{\phi}(\hat{\phi}+1)}{(\hat{\phi}+2)}\bar{x} - \frac{\hat{\phi}^2}{(\hat{\phi}+1)^2} = 1 - \frac{n_0}{N} \tag{4.3}$$

$\hat{\phi}$ can be estimated from equation (4.3) by using Newton Raphson method and then $\hat{\omega}$ may be estimated from equation (4.1).

(b) Methods of Moments : The parameters may be obtained from the moments as

$$\hat{\phi} = \frac{\left\{(2\mu_1' - \mu_2') + \sqrt{(\mu_2' - 2\mu_1' + 2\mu_1'\mu_2')}\right\}}{(\mu_2' - \mu_1')} \tag{4.4}$$

$$\hat{\omega} = 1 - \frac{\hat{\phi}(\hat{\phi}+1)\mu_1'}{(\hat{\phi}+2)} \tag{4.5}$$

where μ_1' and μ_2' denote mean and second order raw moments respectively.

(c) Ratio of First Two Frequencies and Mean: Eliminating ω between first two frequencies, we get

$$\hat{\phi} = \left(\frac{n_1}{2n_2} - 2\right) + \sqrt{\left\{\left(2 - \frac{n_1}{2n_2}\right)^2 - \left(3 - \frac{4n_1}{n_2}\right)\right\}} \qquad (4.6)$$

where $\frac{n_1}{N} = \frac{(1-\omega)(\hat{\phi}+3)\hat{\phi}^2}{(\hat{\phi}+1)^4}$ and $\frac{n_2}{N} = \frac{(1-\omega)(\hat{\phi}+4)\hat{\phi}^2}{(\hat{\phi}+1)^5}$, are the first two relative frequencies and $\hat{\omega}$ may be estimated from equation (4.5).

5. Fitting of IPL Distribution to Data

For the fitting of IPL distribution, we consider two data sets of Beall [1] in Tables 1 and 2, for which generalized Poisson distribution (GPD) was fitted by Jain [8] (using MLE). In Table 3, we consider Student's historic data on Haemocytometer of yeast cells, for which Gegenbauer distribution was fitted by Borah [3], using method of moments. It is observed from Table 1, 2 and 3 that ML gives better result in all the cases. In case of Table 2 the method ratio of first two frequency with mean does not give better fit, as the computed χ^2 value is quite large, hence the result is not reported in this case. It is also clear from the values of the expected IPL frequencies that there is some improvement, however small it may be, in fitting of IPL distribution over the other distributions considered earlier.

Table 1. Fit of distribution on Pyrausta nublilalis in 1937 (data of Beall [1])

No. of Insects	Observed Frequency	IPL (Maximum Likelihood)	IPL (Method of Moments)	IPL (Ratio of Two Freq.)	GPD (Jain [8])
0	33	33.00	32.07	34.08	32.46
1	12	12.41	13.47	11.23	13.47
2	6	5.84	6.00	5.61	5.60
3	3	2.66	2.59	2.71	2.42
4	1	1.18	1.096	1.28	1.08
5	1	0.91	0.774	1.09	0.97
Total	56	56.00	56.00	56.00	56.00
		$\hat{\phi} = 1.588$	$\hat{\phi} = 1.719$	$\hat{\phi} = 1.449$	
Parameter estimates		$\hat{\omega} = 0.1406$	$\hat{\omega} = 0.0573$	$\hat{\omega} = 0.228$	
		$\chi^2 = 0.029$	$\chi^2 = 0.215$	$\chi^2 = 0.096$	$\chi^2 = 0.25$

Table 2. Fit of distribution of Corn Borer (data of Beall [1])

Corn Borer per Hill	Observed Frequency	IPL (Maximum Likelihood)	IPL (Method of Moments)	GPD (Jain [8])
0	43	42.99	44.99	43.91
1	35	32.12	30.39	32.00
2	17	19.45	18.81	19.11
3	11	11.31	11.19	10.88
4	5	6.40	6.47	6.12
5	4	3.55	3.66	3.44
6	1	1.94	2.04	1.94
7	2	1.05	1.12	1.10
8	2	1.19	1.30	1.50
Total	120	120	120	120
Parameter estimates		$\hat{\phi} = 1.0587$	$\hat{\phi} = 1.0715$	
		$\hat{\omega} = -0.5696$	$\hat{\omega} = -0.0087$	
		$\chi^2 = 0.577$	$\chi^2 = 0.995$	$\chi^2 = 0.87$

Table 3. Haemocytometer Counts of Yeast Cells

No. of Yeast cells per sq.	Observed Frequency	IPL (Maximum Likelihood)	IPL (Method of Moments)	IPL (Ratio of first Two Freq)	Gegenbauer (Borah [3])
0	213	213.00	210.46	204.00	214.15
1	128	127.00	131.14	139.18	123.00
2	37	40.91	40.76	40.23	44.88
3	18	12.82	12.39	11.39	13.36
4	3	3.95	3.71	3.18	3.55
5	1	1.20	1.09	0.88	0.86
6	0	0.53	0.45	0.34	0.20
Total	400	400.00	400.00	400.00	400.00
Parameter estimates		$\hat{\phi} = 2.669$	$\hat{\phi} = 2.774$	$\hat{\phi} = 3.0328$	
		$\hat{\omega} = -0.431$	$\hat{\omega} = -0.497$	$\hat{\omega} = -0.6586$	
		$\chi^2 = 1.037$	$\chi^2 = 1.53$	$\chi^2 = 3.93$	$\chi^2 = 2.8342$

ACKNOWLEDGEMENT

The authors are grateful to the referee for his valuable comments which improved the paper considerably.

REFERENCES

[1] Beall, G. (1940). The fit and significance of contagious distributions when applied to observations on larval insects. *Ecology,* **21**, 460-474.

[2] Bhattacharya, S.K. (1966). Confluent hypergeometric distributions of discrete and continuous type with applications to accident proneness. *Bulletin of the Cal. Stat. Assoc.,* **15**, 20-31.

[3] Borah, M. (1984). The Gegenbauer distribution revisited: Some recurrence relations for moments, cumulants, etc., estimation of parameters and its goodness of fit. *Jour. Ind. Soc. Agril. Stat.,* **36**, 72-78.

[4] Borah, M. and Deka Nath, A. (2001). Poisson Lindley and some of its mixture distributions. *Pure and Applied Mathematika Sciences,* **53**, 1–2, 1–8.

[5] Consul, P.C. (1989). *Generalized Poisson Distribution Properties and Applications.* Marcel Dekker, Inc., New York and Basel.

[6] Gerstenkorn, T. (1979). Moment recurrence relation for the inflated negative binomial, poisson and geometric distribution. *Demonstratio Mathematica,* **11(2)**, 389-411.

[7] Gurland, J. (1957). Some interrelations among compound and generalized distributions. *Biometrika,* **44**, 265-268.

[8] Jain, G.C. (1974). On power series distribution association with lagrange expansion. *Biometrische Zeitsehrift,* **17**, 85-97.

[9] Johnson, N.L, Kotz, S. and Kemp, A.W. (1992). *Univariate Discrete Distributions.* 2nd Edn, John Wiley & Sons Inc.

[10] Lindley, D.V. (1958). Fiducial distributions and Baye's theorem. *J. Roy. Statist. Soc.,* **B 20**, 102-107.

[11] Pandey, K.N. (1965). On generalized inflated poisson distribution. *Banaras Hindu Univ. J. Sci. Res.,* **15**, 157-162.

[12] Sankaran, M. (1970). The discrete poisson lindley distribution. *Biometrics,* **26**, 145-146.

[13] Singh, M.P. (1965). Inflated binomial distribution. *Banaras Hindu Univ. J. Sci. Res.,* **16(1)**, 87-90.

[14] Singh, M.P. (1966). A note on generalized inflated binomial distribution. *Sankhya,* **A28 (1)**, 99.

[15] Singh, S.N. (1963). A note on inflated poisson distribution. *J. Ind. Statist. Assoc.,* **1**, 140-144.