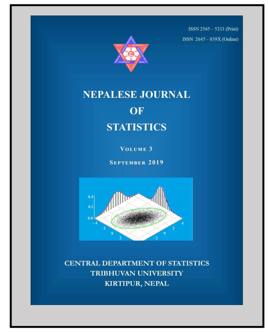
A Generalised Exponential-Lindley Mixture of Poisson Distribution

Binod Kumar Sah[1*] and A. Mishra[2]

Submitted: 09 May 2019; Accepted: 02 August 2019

Published online: 16 September 2019
DOI: https://doi.org/10.3126/njs.v3i0.25575



ABSTRACT

Background: The exponential and the Lindley (1958) distributions occupy central places among the class of continuous probability distributions and play important roles in statistical theory. A Generalised Exponential-Lindley Distribution (GELD) was given by Mishra and Sah (2015) of which, both the exponential and the Lindley distributions are the particular cases. Mixtures of distributions form an important class of distributions in the domain of probability distributions. A mixture distribution arises when some or all the parameters in a probability function vary according to certain probability law. In this paper, a Generalised Exponential- Lindley Mixture of Poisson Distribution (GELMPD) has been obtained by mixing Poisson distribution with the GELD.

Materials and Methods: It is based on the concept of the generalisations of some continuous mixtures of Poisson distribution.

Results: The Probability mass of function of generalized exponential-Lindley mixture of Poisson distribution has been obtained by mixing Poisson distribution with GELD. The first four moments about origin of this distribution have been obtained. The estimation of its parameters has been discussed using method of moments and also as maximum likelihood method. This distribution has been fitted to a number of discrete data-sets which are negative binomial in nature and it has been observed that the distribution gives a better fit than the Poisson–Lindley Distribution (PLD) of Sankaran (1970).

Conclusion: P-value of the GELMPD is found greater than that in case of PLD. Hence, it is expected to be a better alternative to the PLD of Sankaran for similar type of discrete data-set which is negative binomial in nature.

Keywords: Estimation of parameters, generalised exponential-Lindley distribution, goodness of fit, mixing, moments, Poisson-Lindley distribution.

Address correspondence to the author: Department of Statistics, R.R.M. Campus, Janakpur, Tribhuban University, Nepal. Email: sah.binod01@gmail.com[1*] (Corresponding author); Devendra Mishra Institute of Statistics, Patna, India, Email:mishraamar@rediffmail.com[2]

INTRODUCTION

The exponential and the Lindley (1958) distributions occupy central places among the class of continuous probability distributions and both play important roles in the statistical theory. A Generalised Exponential-Lindley Distribution (GELD) has been obtained by Mishra and Sah (2015) of which, both the exponential and the Lindley distributions are the particular cases. A GELD with two parameters α and ϕ is defined by its probability density function (pdf)

$$f(x;\alpha,\phi) = \frac{\phi^2 (1+\alpha x) e^{-\phi x}}{(\phi+\alpha)} \; ; x > 0, \alpha \geq 0, \phi > 0 \ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

It can easily be seen that at $\alpha = 0$, the distribution reduces to the exponential distribution and at $\alpha = 1$, it reduces to the Lindley distribution given by its pdf

$$f(x) = \frac{\phi^2 (1+x) e^{-\phi x}}{(\phi+1)} \quad ; x > 0 \; ; \phi > 0 \; \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(2)}$$

The cumulative distribution function of the GELD (1) is given by

$$F(x) = 1 - \frac{\phi+\alpha+\phi\alpha x}{\phi+\alpha} e^{-\phi x} \quad x > 0, \alpha > \phi > 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(3)}$$

The equations (1) to (3) were obtained by Mishra and Sah (2015). The r^{th} moment about origin of the generalized exponential-Lindley distribution has been obtained as

$$\mu'_r = \frac{\phi^2}{\phi+\alpha} \int_0^\infty x^r (1+\alpha x) e^{-\phi x} dx \; ; r = 1,2,3,4. = \frac{\Gamma(r+1)[\phi+(r+1)\alpha]}{\phi^r (\phi+\alpha)}$$

The one parameter Poisson-Lindley Distribution (PLD) given by the following probability mass function

$$P(X) = \frac{\phi^2 (x+\phi+2)}{(\phi+1)^{x+3}} \quad ; x = 0,1,2,\ldots \quad ; \phi > 0 \; \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(4)}$$

introduced by Sankaran (1970) to model count data. This distribution arises from a Poisson distribution when its parameter follows a Lindley (1958) distribution with its probability density function given in (2). The first four moments about origin of the PLD (4) have been obtained as

$$\mu'_1 = \frac{(\phi+2)}{\phi(\phi+1)} \; \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(5)}$$

$$\mu'_2 = \frac{(\phi+2)}{\phi(\phi+1)} + \frac{2(\phi+3)}{\phi^2(\phi+1)} \; \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(6)}$$

$$\mu'_3 = \frac{(\phi+2)}{\phi(\phi+1)} + \frac{6(\phi+3)}{\phi^2(\phi+1)} + \frac{6(\phi+4)}{\phi^3(\phi+1)} \; \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(7)}$$

$$\mu'_4 = \frac{(\phi+2)}{\phi(\phi+1)} + \frac{14(\phi+3)}{\phi^2(\phi+1)} + \frac{36(\phi+4)}{\phi^3(\phi+1)} + \frac{24(\phi+5)}{\phi^4(\phi+1)} \; \ldots\ldots\ldots\ldots\ldots\ldots \text{(8)}$$

The equations (4) to (8) were obtained by Sankaran (1970). Ghitany and Al-Mulairi (2009) discussed the estimation methods for the one parameter PLD (4) and its applications. In this paper, a Generalised Exponential- Lindley Mixture of Poisson Distribution (GELMPD) has been obtained by mixing a Poisson distribution with a GELD (1). The first four moments about origin have been obtained and the estimation of its parameters has been discussed. The distribution has been fitted to some discrete data-sets and it has been observed that it is more flexible than the Sankaran's PLD for analyzing different types of count data.

MATERIALS AND METHODS

It is based on the concept of some continuous mixture of Poisson distribution. Probability mass function of the GELMPD has been obtained by mixing Poisson distribution with GELD (1) such that it follows basic properties of probability distributions. The first four moments about origin of GELMPD has been obtained. The parameters of the proposed distribution have been obtained by using the first two moments about origin as well as Maximum Likelihood method. The distribution has been fitted to some discrete data-sets which are negative binomial in nature and it has been observed that it is more flexible than the Sankaran's PLD for analyzing different types of count data.

RESULTS

A GELMPD

Suppose that the parameter λ of a Poisson distribution follows GELD (1). Thus, the GELMPD can be obtained as

$$P(x;\phi,\alpha) = \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}.\frac{\phi^2}{\phi+\alpha}.(1+\alpha\lambda)e^{-\phi\lambda}d\lambda \quad \ldots\ldots\ldots\ldots\ldots (9)$$

$$= \frac{\phi^2}{(\phi+\alpha)x!}\left[\int_0^\infty \lambda^x e^{-(1+\phi)\lambda}d\lambda + \alpha\int_0^\infty \lambda^{x+1}e^{-(1+\phi)\lambda}d\lambda\right]$$

$$= \frac{\phi^2}{(\phi+\alpha)x!}\left[\frac{\Gamma(x+1)}{(1+\phi)^{x+1}} + \alpha\frac{\Gamma(x+2)}{(1+\phi)^{x+2}}\right]$$

$$= \frac{\phi^2}{(\phi+\alpha)}\left[\frac{1+\phi+\alpha+\alpha x}{(1+\phi)^{x+2}}\right] \; ; x > 0,\ \phi > 0 \text{ and } \phi > \alpha \quad \ldots (10)$$

Expression (9) has been obtained on the basis of continuous mixture of Poisson distribution. The parameter λ of the Poisson distribution is considered as a continuous variable which follows GELD. Expression (10) is the probability mass function of GELMPD. It reduces to the PLD of Shankaran (1970) at α = 1.

Moments of GELMPD

The r[th] moment about origin of the GELMPD (10) can be obtained as

$$\mu_r' = E\left[E(X^r / \lambda)\right]$$

$$\mu_r' = \int_0^\infty (\sum x^r \frac{e^{-\lambda}\lambda^x}{\Gamma(x+1)}) \frac{\phi^2}{(\phi+\alpha)}(1+\alpha\lambda)e^{-\phi\lambda}d\lambda \quad \text{............ (11)}$$

Obviously, the expression under bracket is the r[th] moment about origin of the Poisson distribution. Taking $r = 1$ in (11) and using the mean of the Poisson distribution, the mean of the GELMPD is obtained as

$$\mu_1' = \frac{\phi^2}{\phi+\alpha}\int_0^\infty \lambda(1+\alpha\lambda)e^{-\phi\lambda}d\lambda$$

$$= \frac{(\phi+2\alpha)}{\phi(\phi+\alpha)} \quad \text{.. (12)}$$

The expression (12) reduces to the first moment about origin of the PLD at α = 1. The second moment about origin of the GELMPD can be obtained as

$$\mu_2' = \frac{\phi^2}{\phi+\alpha}\int_0^\infty (\lambda^2+\lambda)(1+\alpha\lambda)e^{-\phi\lambda}d\lambda$$

$$= \frac{(\phi+2\alpha)}{\phi(\phi+\alpha)} + \frac{2(\phi+3\alpha)}{\phi^2(\phi+\alpha)} \quad \text{.......................... (13)}$$

and the variance as

$$\mu_2 = \mu_2' - \mu_1'^2$$

$$= \frac{1}{\phi^2(\phi+\alpha)}\left[\frac{(6\alpha+2\alpha\phi+2\phi+\phi^2)(\phi+\alpha)-(\phi^2+4\alpha\phi+4\phi^2)}{(\phi+\alpha)}\right]$$

$$= \frac{1}{\phi^2}\left[\frac{\phi^3+(3\phi^2\alpha+\phi^2)+(4\alpha\phi+2\alpha^2\phi)+2\alpha^2}{(\phi+\alpha)^2}\right] \quad \text{.................. (14)}$$

The expression (14) reduces to the variance of PLD at α = 1. Similarly, taking $r = 3$ and $r = 4$ in (11) and using the respective moments about origin of the Poisson distribution, the third and fourth moments about origin the GELMPD are obtained as

$$\mu_3' = \frac{\phi^2}{\phi+\alpha}\int_0^\infty (\lambda^3+3\lambda^2+\lambda)(1+\alpha\lambda)e^{-\phi\lambda}d\lambda\,.$$

After a little simplification, we get

 ISSN: 2565-5213 (Print); 2465-839X (Online)

$$\mu_3' = \frac{(\phi+2\alpha)}{\phi(\phi+\alpha)} + \frac{6(\phi+3\alpha)}{\phi^2(\phi+\alpha)} + \frac{6(\phi+4\alpha)}{\phi^3(\phi+\alpha)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (15)$$

$$\mu_4' = \frac{\phi^2}{\phi+\alpha}\int_0^\infty (\lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda)(1+\alpha\lambda)e^{-\phi\lambda}d\lambda$$

After a little more simplification, we get

$$\mu_4' = \frac{(\phi+2\alpha)}{\phi(\phi+\alpha)} + \frac{14(\phi+3\alpha)}{\phi^2(\phi+\alpha)} + \frac{36(\phi+4\alpha)}{\phi^3(\phi+\alpha)} + \frac{24(\phi+5\alpha)}{\phi^4(\phi+\alpha} \quad \ldots\ldots\ldots (16)$$

It can be observed that at α = 1 these moments reduce to the respective moments of the one-parameter PLD.

Estimation of parameters

Method of moments

Here, the method of moments has been used to estimate the parameters of the GELMPD. The parameter α of this distribution can be obtained by solving the expression (12) of $\mu_1^/$ as

$$\phi^2\mu_1^/ + \phi\alpha\mu_1^/ = \phi + 2\alpha$$

$$\text{i.e. } \alpha = \frac{\phi - \phi^2\mu_1^/}{\mu_1^/\phi - 2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (17)$$

Substituting the value of α in the expression (4.3) of $\mu_2^/$,we get the estimator ϕ by solving the following quadratic equation.

$$(\mu_2^/ - \mu_1^/)\phi^2 - 4\mu_1^/\phi + 2 = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (18)$$

$$\text{i.e. } \phi = \frac{2\mu_1^/ \pm \sqrt{(2(2\mu_1^{/2} - \mu_2^/ + \mu_1^/))}}{(\mu_2^/ - \mu_1^/)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (19)$$

Replacing the population moments by the respective sample moments and using the expression (19), we can obtain the estimated value of ϕ .

Maximum likelihood method

Let x_1 ,x_2 , ... ,x_n be the value of observations of a random sample of size n from the two-parameter GELMPD(10) and let f_x be the observed frequency in the sample corresponding to X=x (x=1,2,...,k) such that $\sum_{x=1}^{k} f_x = n$,where k is the largest observed value having non-zero frequency. Let the likelihood function L of the GELMPD is given by

$$L=\left(\frac{\phi^2}{\phi+\alpha}\right)^n.(1+\phi)^{-\sum_{x=1}^{k}(x+2)f_x}\prod_{x=1}^{k}(1+\phi+\alpha+\alpha x)^{f_x} \quad ......\ (20)$$

and the log likelihood function is obtained as

$$\log\ L=n\log\left(\frac{\phi^2}{\phi+\alpha}\right)-\sum_{x=1}^{k}f_x(x+2)\log(1+\phi)+\sum_{x=1}^{k}f_x\log(1+\phi+\alpha+\alpha x) \quad ......\ (21)$$

The two likelihood equations are thus obtained as

$$\frac{\partial \log L}{\partial\phi}=\frac{2n}{\phi}-\frac{n}{(\phi+\alpha)}-\frac{\sum_{x=1}^{k}(x+2)f_x}{(1+\phi)}+\sum_{x=1}^{k}(1+\phi+\alpha+\alpha x)^{-1}f_x=0 \quad ....\ (22)$$

$$\frac{\partial \log L}{\partial\alpha}=-\frac{n}{(\phi+\alpha)}+\sum_{x=1}^{k}(1+\phi+\alpha+\alpha x)^{-1}(1+x)f_x=0 \quad ....\ (23)$$

Two equations (22) and (23) do not seem to be solved directly. However, Fisher's scoring method can be applied to solve these equations. By applying this method, the following equations for $\hat{\phi}$ and $\hat{\alpha}$ can be solved.

$$\begin{bmatrix}\frac{\partial^2\log L}{\partial\phi^2} & \frac{\partial^2\log L}{\partial\phi\partial\alpha}\\ \frac{\partial^2\log L}{\partial\phi\partial\alpha} & \frac{\partial^2\log L}{\partial\alpha^2}\end{bmatrix}_{\substack{\hat{\phi}=\phi_0\\ \alpha=\alpha_0}}\begin{bmatrix}\hat{\phi}-\phi_0\\ \alpha-\alpha_0\end{bmatrix}=\begin{bmatrix}\frac{-\partial\log L}{\partial\phi}\\ \frac{-\partial\log L}{\partial\alpha}\end{bmatrix} \quad ....(24)$$

where ϕ_0 and α_0 are the initial values of ϕ and α respectively. These equations are solved iteratively till sufficiently close estimates of $\hat{\phi}$ and $\hat{\alpha}$ are obtained.

Goodness of fit

The GELMPD has been fitted to a number of discrete data-sets which were of the nature of negative binomial distribution and in all of such cases this has been found to provide a very close fits to the two data-sets have been given here. These data-sets are (1) from Kemp and Kemp (1965) relating to the distribution of mistakes in copying group of random digits (Table 1) and (2) Beall (1940) relating to the distribution of Pyrautanablilalis in 1937 (Table 2). The first data has been used by Sankaran (1970) in the paper entitled 'The Discrete Poisson-Lindley Distribution' and the second data has been used by M. Borah and A. DekaNath in a study entitled 'A study on the Inflated Poisson –Lindley distribution'. All these data-sets have been used by B. K. Sah (2012) in the Ph.D. thesis entitled "Generalisations of Some Countable and Continuous Mixtures of Poisson Distribution and Their Applications" for the study of generalization of PLD.

Table1. Distribution of mistakes in copying groups of random digits.

Number of errors per group	Observed frequency	Expected frequency of PLD	Expected frequency of GELMPD
0	35	33.0	33.0
1	11	15.3	15.1
2	8	6.8	6.8
3	4	2.9	3.0
4	2	2.0	2.1
Total	60	60	60
μ_1'	0.7833		
μ_2'	1.8500		
$\hat{\phi}$		1.7434	0.9375001
$\hat{\alpha}$		-	-0.1967592
d.f.		2	2
χ^2		2.21	1.65
P-value		0.57	0.75

Table 2. Distribution of Pyraustanablilalis in 1937.

Number of insects	Observed frequency	Expected frequency of PLD	Expected frequency of GELMPD
0	33	31.5	31.9
1	12	14.2	13.8
2	6	6.1	5.9
3	3	2.5	2.5
4	1	1.0	1.1
5	1	0.7	0.8
Total	56	56	56
μ_1'	0.7500		
μ_2'	1.8571		
$\hat{\phi}$		1.8081	0.9375001
$\hat{\alpha}$		-	-0.1191623
d.f.		2	1
χ^2		0.53	0.29
P-value		0.83	0.85

CONCLUSION

The expected frequencies according to the one parameter PLD have also been given in these tables for ready comparison with those obtained by the GELMPD. It has been observed that the GELMPD gives better fit to most of the discrete data-sets which are negative binomial in nature than the one parameter PLD of Sankaran (1970). Hence, it is expected to be a better alternative to the PLD of Sankaran for similar type of discrete data-set which is negative binomial in nature.

CONFLICT OF INTEREST

The author declared that there is no conflict of interest.

ACKNOWLEDGEMENTS

The authors express gratitude to the referee(s) for valuable comments and suggestions which improved the quality of the research article.

REFERENCES

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

Borah, M., & DekaNath, A. (2001). A study on the Inflated Poisson-Lindley distribution, *Journal of the Indian society of Agricultural Statistics*, *54*(2), 317-323.

Ghitany, M. E. & Al-Mutairi, D. K. (2009). Estimation methods for the discrete Poisson-Lindley distribution. *J. Stat. Compt. Simul*. *79* (1); 1-9.
doi: 10.1080/00949650701550259

Kemp, C. D., & Kemp, A. W. (1965). Some properties of the Hermite distribution. *Biometrika*, *52*, 381-394.

Lindely, D.V. (1958). Fiducial distributions and Bayes theorem, *Journal of Royal Statistical Society*, Ser. B, *20*, 102-107.

Mishra, A., & Sah, B. K. (2016). A Generalised Exponential- Lindley Distribution, *Recent Advances in Mathematics, Statistics and Computer Science,* 229-238.

Sankaran, M. (1970). The discrete Poisson-Lindley distribution, *Biometrics*, *26*,145-149.

Sah, B. K. (2012). Generalisations of some countable and continuous mixtures of Poisson distribution and their applications, (Doctoral thesis, Patna University, Patna India).

Reference to this paper should be made as follows:
Sah, B. K., & Mishra, A. (2019). A generalised exponential-Lindley mixture of Poisson distribution. *Nep. J. Stat., 3,* 11-20.