On a Generalised Exponential-Lindley Mixture of Generalised Poisson Distribution

Binod Kumar Sah\textsuperscript{1*} and A. Mishra\textsuperscript{2}

Submitted: 13 May 2020; Accepted: 06 December 2020
Published online: 18 December 2020
DOI: https://doi.org/10.3126/njs.v4i0.33449

ABSTRACT

Background: A mixture distribution arises when some or all parameters in a mixing distribution vary according to the nature of original distribution. A generalised exponential-Lindley distribution (GELD) was obtained by Mishra and Sah (2015). In this paper, generalized exponential-Lindley mixture of generalised Poisson distribution (GELMGPD) has been obtained by mixing generalized Poisson distribution (GPD) of Consul and Jain’s (1973) with GELD. In the proposed distribution, GELD is the original distribution and GPD is a mixing distribution. Generalised exponential-Lindley mixture of Poisson distribution (GELMPD) was obtained by Sah and Mishra (2019). It is a particular case of GELMGPD.

Materials and Methods: GELMGPD is a compound distribution obtained by using the theoretical concept of some continuous mixtures of generalised Poisson distribution of Consul and Jain (1973). In this mixing process, GELD plays a role of original distribution and GPD is considered as mixing distribution.

Results: Probability mass function (pmf) and the first four moments about origin of the generalised exponential-Lindley mixture of generalised Poisson distribution have been obtained. The method of moments has been discussed to estimate parameters of the GELMGPD. This distribution has been fitted to a number of discrete data-sets which are negative binomial in nature. P-value of this distribution has been compared to the PLD of Sankaran (1970) and GELMPD of Sah and Mishra (2019) for similar type of data-sets.

Conclusion: It is found that P-value of GELMGPD is greater than that in each case of PLD and GELMPD. Hence, it is expected to be a better alternative to the PLD of Sankaran and GELMPD of Sah and Mishra for similar types of discrete data-sets which are negative binomial in nature. It is also observed that GELMGPD gives much more significant result when the value of $\theta$ is negative.

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

Address correspondence to the author: Department of Statistics, R.R.M. Campus, Janakpur, Tribhuvan University, Nepal. Email: sah.binod01@gmail.com\textsuperscript{1*} (Corresponding author); Department of Statistics, Patna University, Patna, India, Email: mishraamar@rediffmail.com\textsuperscript{2}
INTRODUCTION

The exponential distribution is the probability distribution of the events in a Poisson point process i.e. a process in which events occur continuously and independently at a constant rate. It is a particular case of gamma distribution. The Lindley distribution was originally proposed by Lindley (1958) in the context of Bayesian Statistics, as a counter example of fiducial Statistics. The exponential and Lindley distributions both are the particular cases of generalised exponential-Lindley distribution (GELD) of Mishra and Sah (2016) and its probability density function (pdf) is given by:

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

It can observe that at $\alpha = 0$, the distribution is reduced to the exponential distribution and at $\alpha = 1$, it reduces to the Lindley distribution having pdf:

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

The one parameter Poisson-Lindley distribution (PLD) was obtained by Sankaran (1970) and its probability mass function is given by:

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

Ghitany et al (2009) discussed the estimation methods for the one parameter Poisson-Lindley distribution (3) and its applications. Probability mass function of GELMPD has been obtained by mixing a Poisson distribution with the GELD (1) as:

$$P(x; \phi, \alpha) = \int_{0}^{\infty} \frac{e^{-\lambda}}{x!} \cdot \frac{\phi^2}{\phi + \alpha} \cdot (1 + \alpha \lambda) e^{-\phi \lambda} d\lambda$$

$$= \frac{\phi^2}{(\phi + \alpha)} \left[ \frac{1 + \phi + \alpha + \alpha x}{(1 + \phi)^{x+2}} \right] ; x > 0, \phi > 0 \text{ and } \phi > \alpha \hspace{1cm} \ldots \ldots \hspace{1cm} (4)$$

The expression (4) is the probability mass function of GELMPD. It is reduced to the Poisson-Lindley distribution of Sankaran (1970) at $\alpha = 1$. The first four moments about origin of the GELMPD have been obtained as

$$\mu'_1 = \frac{(\phi + 2\alpha)}{\phi(\phi + \alpha)} \hspace{1cm} \ldots \ldots \hspace{1cm} (5)$$

$$\mu'_2 = \frac{(\phi + 2\alpha)}{\phi(\phi + \alpha)} + \frac{2(\phi + 3\alpha)}{\phi^2(\phi + \alpha)} \hspace{1cm} \ldots \ldots \hspace{1cm} (6)$$

$$\mu'_3 = \frac{(\phi + 2\alpha)}{\phi(\phi + \alpha)} + \frac{6(\phi + 3\alpha)}{\phi^2(\phi + \alpha)} + \frac{6(\phi + 4\alpha)}{\phi^3(\phi + \alpha)} \hspace{1cm} \ldots \ldots \hspace{1cm} (7)$$
The equations (4) to (8) were obtained by Sah and Mishra (2019).

Consul and Jain (1973) obtained a two-parameter generalised Poisson distribution (GPD) given by its probability function

\[
P(x) = \frac{\lambda (\lambda + x \theta)^{x-1} e^{-(\lambda + x \theta)}}{x!} \quad \text{where} \quad x = 0, 1, 2, \ldots ; \lambda > 0; |\theta| < 1 \tag{9}
\]

In this paper, a generalised exponential- Lindley mixture of generalised Poisson distribution (GELMGPD) has been obtained by mixing generalised Poisson distribution (9) with the GELD (1). Hence, the title of the article is named as “On a Generalised Exponential- Lindley Mixture of Generalised Poisson Distribution”. Probability mass of function (pmf) and the first four moments about origin of the generalised exponential-Lindley mixture of generalised Poisson distribution has been obtained. The method of moments has been discussed to estimate its parameters. This distribution has been fitted to a number of discrete data-sets which are negative binomial in nature.

**MATERIALS AND METHODS**

GELMGPD is a compound distribution obtained by using the theoretical concept of some continuous mixtures of generalised Poisson distribution of Consul and Jain (1973). In this mixing process, GELD plays a role of original distribution and GPD is considered as mixing distribution. The parameter \( \lambda \) of the GPD (9) follows GELD (1) which is continuous in nature. In this process, the mixing distribution is discrete, so the resultant distribution is also discrete in nature. After constructing probability mass function of the GELMGPD, the first four moments about origin have been obtained. The method of moments has been discussed to estimate its parameters. This distribution has been fitted to a number of discrete data-sets which are negative binomial in nature. P-value of this distribution has been compared to the PLD of Sankaran (1970) and GELMPD of Sah and Mishra (2019) for similar type of data-sets and finally, conclusions have been drawn.

**RESULTS**

*A generalised exponential-Lindley mixture of generalised Poisson distribution (GELMGPD)*

The GPD (9) have two parameters \( \lambda \) and \( \theta \). Let us suppose that the parameter \( \lambda \) in the GPD (9) is a random variable which follows the GELD (1) with parameters \( \phi \) and \( \alpha \). Thus, the GELMGPD can be obtained as:
\[ P(x; \phi, \alpha, \theta) = \frac{\phi^2 e^{-\theta x}}{\Gamma(x+1)(\phi + \alpha)} \int_0^\infty \lambda^x \left(1 + \frac{\theta x}{\lambda}\right)^{x-1} \frac{(1 + \alpha \lambda)e^{-\lambda(1+\phi)}}{(\phi + \alpha)} d\lambda \ldots \ldots \ (10) \]

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

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

\[ = \frac{\phi^2 e^{-\theta x}}{\Gamma(x+1)(\phi + \alpha)} \left[ \sum_{i=0}^{x-1} \frac{\lambda^i (\lambda x)^1}{i!} \left(1 + \frac{\theta x}{\lambda}\right)^{x-1} e^{-\lambda(1+\phi)} d\lambda \right] \]

\[ = \frac{\phi^2 e^{-\theta x}}{\Gamma(x+1)(\phi + \alpha)} \left[ \sum_{i=0}^{x-1} \frac{\lambda^i (\lambda x)^1}{i!} \left(1 + \frac{\theta x}{\lambda}\right)^{x-1} e^{-\lambda(1+\phi)} d\lambda \right] \]

\[ = \frac{\phi^2 e^{-\theta x}}{\Gamma(x+1)(\phi + \alpha)} \left[ \sum_{i=0}^{x-1} \frac{\lambda^i (\lambda x)^1}{i!} \left(1 + \frac{\theta x}{\lambda}\right)^{x-1} e^{-\lambda(1+\phi)} d\lambda \right] \]

\[ = \frac{\phi^2 e^{-\theta x}}{\Gamma(x+1)(\phi + \alpha)} \left[ \sum_{i=0}^{x-1} \frac{\lambda^i (\lambda x)^1}{i!} \left(1 + \frac{\theta x}{\lambda}\right)^{x-1} e^{-\lambda(1+\phi)} d\lambda \right] \]

\[ = \frac{\phi^2 e^{-\theta x}}{\Gamma(x+1)(\phi + \alpha)} \left[ \sum_{i=0}^{x-1} \frac{\lambda^i (\lambda x)^1}{i!} \left(1 + \frac{\theta x}{\lambda}\right)^{x-1} e^{-\lambda(1+\phi)} d\lambda \right] \]

After a little simplification and arrangement of terms, we get:

\[ P(x; \alpha, \phi, \theta) = \frac{\phi^2 e^{-\theta x}}{(\phi + \alpha)(1+\phi)^{x+2}} \left[ (1 + \phi + \alpha(1+x)) + \sum_{i=1}^{x-1} \frac{\theta i^x (x-i)^{i-1} (1+\phi + \alpha(x-i+1))}{i!(1+\phi)^i} \right] \ldots \ldots \ (11) \]
where \( x \geq 0, |\theta| < 1, \phi > 0, \) and \( \phi > \alpha. \)

The expression (11) is the pmf of GELMGPD. We can observe that at \( \theta = 0, \) it is reduced to the GELMPD (4). It can also be noted that \( P(X=0) \) in the GELMGPD becomes independent of \( \theta, \) the additional parameter introduced in the distribution, and its role starts from \( P(X=1) \) onwards. It can also be seen that at \( \alpha = 1 \) that it is reduced to the generalised Poisson-Lindley distribution (GPLD) of B. K. Sah (2012).

**Moments of GELMGPD**

The \( r \)th moment about origin of the GELMGPD (11) can be obtained as:

\[
\mu_r' = E \left[ E( X^r / \lambda) \right] = \ldots \ldots \quad (12)
\]

\[
\mu_r' = \int_0^\infty \sum_{\lambda=0}^{\infty} x^r \lambda (x+\lambda)^{\lambda-1} e^{-\lambda(\lambda+x)} \frac{\phi^2 (1+\alpha \lambda) e^{-\phi \lambda}}{(\phi+\alpha)} d\lambda = \ldots \ldots \quad (13)
\]

It is clear that the expression under bracket in (13) is the \( r \)th moment about origin of the GPD (9). Taking \( r = 1 \) in (13) and using the mean of the GPD, the mean of the GELMGPD (11) can be obtained as:

\[
\mu_1' = \frac{\phi^2}{(\phi+\alpha)(1-\theta)} \int_0^\infty \frac{\lambda + \alpha \lambda^2}{(1-\theta)^3} e^{-\phi \lambda} d\lambda = \frac{(\phi+2\alpha)}{\phi(\phi+\alpha)(1-\theta)} = \ldots \ldots \quad (14)
\]

Again, taking \( r = 2 \) in (13) and using the second moment about origin of the GPD, the second moment about origin of the GELMGPD (11) can be obtained as:

\[
\mu_2' = \frac{\phi^2}{(\phi+\alpha)} \int_0^\infty \frac{\lambda^2}{(1-\theta)^5} + \frac{\lambda^2}{(1-\theta)^3} (1+\alpha \lambda) e^{-\phi \lambda} d\lambda
\]

which after a little simplification gives:

\[
= \frac{(\phi+2\alpha)}{\phi(\phi+\alpha)(1-\theta)^3} + \frac{(2\phi+6\alpha)}{\phi^2(\phi+\alpha)(1-\theta)^2} = \ldots \ldots \quad (15)
\]

Taking \( r = 3 \) in (13) and using the third moment about origin of the GPD, the third moment about origin of the GELMPD can be obtained as:

\[
\mu_3' = \frac{\phi^2}{(\phi+\alpha)} \int_0^\infty \frac{\lambda(1+2\theta)}{(1-\theta)^5} + \frac{3\lambda^2}{(1-\theta)^4} + \frac{\lambda^3}{(1-\theta)^3} (1+\alpha \lambda) e^{-\phi \lambda} d\lambda
\]

which after a little simplification gives:

\[
= \frac{(1+2\theta)(\phi+2\alpha)}{\phi(\phi+\alpha)(1-\theta)^5} + \frac{3(2\phi+6\alpha)}{\phi^2(\phi+\alpha)(1-\theta)^4} + \frac{6(\phi+4\alpha)}{\phi^3(\phi+\alpha)(1-\theta)^3} = \ldots \ldots \quad (16)
\]

Finally, taking \( r = 4 \) in (13) and using the fourth moment about origin of the GPD, the fourth moment about origin of the GELMGPD can be obtained as:
\[ \mu'_1 = \frac{\phi^2}{(\phi + \alpha)^2} \int_0^\infty \left[ \frac{1 + 8\theta + 6\theta^2}{(1 - \theta)^3} \lambda^2 + \frac{7 + 8\theta}{(1 - \theta)^6} \lambda^3 + \frac{6\lambda^3}{(1 - \theta)^5} + \frac{\lambda^4}{(1 - \theta)^4} \right] (1 + \alpha \lambda) e^{-\phi \lambda} d\lambda \]

which after a little simplification gives:

\[ = \frac{(1 + 8\theta + 6\theta^2)(\phi + 2\alpha)}{\phi(\phi + \alpha)(1 - \theta)^3} + \frac{(7 + 8\theta)(2\phi + 6\alpha)}{\phi^2(\phi + \alpha)(1 - \theta)^6} + \frac{36(\phi + 4\alpha)}{\phi^3(\phi + \alpha)(1 - \theta)^5} + \frac{24(\phi + 5\alpha)}{\phi^4(\phi + \alpha)(1 - \theta)^4} \]  

(17)

It can easily be seen that at \( \theta = 0 \), these moments are reduced to the respective moments of the GELMPD.

**Estimation of parameters**

The GELMGPD (11) have three parameters \( \phi, \alpha \) and \( \theta \). Here, we have obtained the estimates of these parameters by using \( P(x=0) \) and, the first and second moments about origin of the GELMGPD.

\[ P(X = 0) = \frac{\phi^2}{(1 + \phi)^2} (1 + \phi + \alpha) = k \text{ (say)} \]  

(18)

or

\[ \alpha = \frac{\phi^2}{k(1 + \phi)^2 - \phi^2} - \phi \]  

(19)

Substituting the value of \( \alpha \) in the expression of \( \mu'_2 \) (15), we get:

\[ \mu_2 \left[ 2\phi - k(1 + \phi)^2 - \phi^2 \right] - [6\phi^2 - 4\phi \left( k(1 + \phi)^2 - \phi^2 \right) + \mu'_1 \phi^4] \mu'_1 = 0 \]  

(20)

The expression (20) is the polynomial equation in \( \phi \) which can be solved by the Newton-Raphson or Regula-Falsi method. The first and second population moments are replaced by their respective sample moments and putting the value of \( \phi \) and \( \alpha \) in expression (14) an estimate of \( \theta \) can be obtained by:

\[ (1 - \theta) = -\frac{(\phi + 2\alpha)}{\phi(\phi + \alpha)\mu'_1} \]  

(21)

**Goodness of fit**

To test goodness of fit, the GELMPGD was fitted to a number of discrete data-sets, having variance greater than the mean, to which earlier PLD (3) has been fitted by others and it was found that for almost all the considered data-sets, the GELMGPD (11) provide closer fits than the PLD (3) and GELMGPD (4). Here, three data-sets are used. These data-sets are (1) Number of red mites on apple leaves, available in Fisher at al (1943), (2) Beall (1940) relating to the distribution of Pyrautanabilalis in 1937 and (3) data relating to the Haemocytometer counts of yeast cell. The first data has been used by Shanker at al (2017) and the second and third data have been used by M. Borah and A. Deka Nath (2001). All these data-sets have been used by B. K. Sah (2012) for the study of generalization of Poisson-Lindley distribution.
Table 1. Number of red mites on the apple leaves.

| Number of mites per leaf | Observed frequency | Expected frequency of PLD | Expected frequency of GELMPD | Expected frequency of GELMGPD |
|--------------------------|--------------------|--------------------------|-----------------------------|-----------------------------|
| 0                        | 38                 | 35.8                     | 36.1                        | 38.0                        |
| 1                        | 17                 | 20.7                     | 20.2                        | 19.3                        |
| 2                        | 10                 | 11.4                     | 11.1                        | 10.3                        |
| 3                        | 9                  | 6.0                      | 6.0                         | 6.6                         |
| 4                        | 3                  | 3.1                      | 3.3                         | 3.1                         |
| 5                        | 2                  | 1.6                      | 1.7                         | 1.7                         |
| 6                        | 1                  | 0.8                      | 0.9                         | 0.9                         |
| 7+                       | 0                  | 0.6                      | 0.7                         | 0.3                         |
| Total                    | 80                 | 80                       | 80                          | 80                          |

Table 2. Distribution of Pyraustanablilalis in 1937.

| Number of insects | Observed frequency | Expected frequency of PLD | Expected frequency of GELMPD | Expected frequency of GELMGPD |
|-------------------|--------------------|---------------------------|-----------------------------|-----------------------------|
| 0                 | 33                 | 31.5                      | 31.9                        | 33.0                        |
| 1                 | 12                 | 14.2                      | 13.8                        | 11.5                        |
| 2                 | 6                  | 6.1                       | 5.9                         | 5.9                         |
| 3                 | 3                  | 2.5                       | 2.5                         | 3.1                         |
| 4                 | 1                  | 1.0                       | 1.1                         | 1.5                         |
| 5                 | 1                  | 0.7                       | 0.8                         | 1.0                         |
| Total             | 56                 | 56                        | 56                          | 56 Contd..                  |
### Table 3. Haemocytometer counts of yeast cell.

| Number of Yeast Cell | Observed frequency | Expected frequency of PLD | Expected frequency of GELMPD | Expected frequency of GELMGPD |
|----------------------|--------------------|---------------------------|-----------------------------|------------------------------|
| 0                    | 213                | 234.4                     | 228.6                       | 213.0                        |
| 1                    | 128                | 99.3                      | 101.5                       | 128.5                        |
| 2                    | 37                 | 40.4                      | 43.6                        | 37.3                         |
| 3                    | 18                 | 16.0                      | 17.9                        | 16.1                         |
| 4                    | 3                  | 6.2                       | 6.8                         | 3.4                          |
| 5                    | 1                  | 3.7                       | 1.6                         | 1.7                          |
| Total                | 400                | 400                       | 400                         | 400                          |

**Contd.**

| Parameter | Value | Value | Value |
|-----------|-------|-------|-------|
| $\mu_1'$  | 0.7500|       |       |
| $\mu_2'$  | 1.8571|       |       |
| $^\wedge\phi$ | 1.8081| 0.9375001| 0.886171302|
| $^\wedge\alpha$ | - | -0.1191623 | -0.287240047 |
| $^\wedge\theta$ | 2 | 1 | 0.2169876 |
| d.f. | | | 1 |
| $\chi^2$ | 0.53 | 0.29 | 0.061 |
| P-value | 0.83 | 0.85 | 0.89 |
All these data are related to count data from ecological and biological sciences. The main reason for selecting these data is overdispersion. We can also observe that expected frequency obtained by GELMGPD is much closer to its respective observed frequency for most of the observations than PLD and GELMPD. It indicates that GELMGPD has the least variance among all the mentioned tables. In Table 3, the P-value of the PLD and GELMPD are significant because P-value is less .05 but that of GELMGPD is highly not significantly (0.85). It may create doubt on the obtained result. It happens due to the negative value of the additional parameter (θ) of the GPD (9). The negative value of θ reduces variance of the GPD and hence that of proposed mixture distribution.

CONCLUSION

It is found that P-value of the proposed distribution is greater than that in each case of PLD and GELMPD. Hence, it is expected to be a better alternative to the PLD of Sankaran and GELMPD of Sah and Mishra for similar types of discrete data-sets which are negative binomial in nature. It is also observed that GELMGPD gives much more significant result when the value of θ is negative for similar type of data-sets.

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., & Deka Nath, A. (2001). A study on the inflated Poisson-Lindley distribution. *Journal of the Indian society of Agricultural Statistics, 54*(2), 317-323.

Consul, P. C., & Jain, G. C. (1973). A generalisation of the Poisson distribution. *Technometrics, 15*, 791-799.

Fisher, R. A, Corpet A. S., & Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of animal population. *Journal of Animal Ecology, 12*(1), 42-58.

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

Lindley, 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.

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

Sah, B. K., & Mishra, A. (2019). A generalised exponential-Lindley mixture of Poisson distribution. *Nep. J. Stat.*, 3, 11-20.

doi: https://doi.org/10.3126/njs.v3i0.25575.

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

Shanker, R., Fessahye, H., Shanker, R., Lionida, T. A., & Sium. S. (2017). On Discrete Poisson-Shanker Distribution and its applications. *Biometrics & Biostatistics International Journal*, 5(1), 6-14.

doi: 00121. DOI:10.15406/bbij.2017.05.00121.

**Reference** to this paper should be made as follows:

Sah, B. K., & Mishra, A. (2020). On a generalised exponential-Lindley mixture of generalised Poisson distribution. *Nep. J. Stat.*, 4, 33-42.