



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: I Month of publication: January 2021

DOI: <https://doi.org/10.22214/ijraset.2021.32839>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

A Generalised Neyman Type - A Distribution

Binod Kumar Sah¹, A. Mishra²

¹Department of Statistics, R.R.M. Campus (T.U), Janakpur, Nepal.

²Department of Statistics, Patna University, Patna, India.

Abstract: The Neyman Type - A distribution was developed by Neyman (1939) to describe the distribution of larvae in experimental field plots. It is widely used in describing population under the influence of contagion and bacteriology. In this paper, a generalization of Neyman Type-A distribution has been obtained by mixing the classical Poisson distribution with the restricted generalised Poisson distribution of Consul and Jain (1973). Probability mass function of the proposed distribution has been derived. The first four moments about origin as well as about the mean have been obtained. Parameters of the proposed distribution have been discussed by the method of moments. It is expected that the obtained generalized distribution can explain the variation in the various discrete data sets, having variance greater than the mean, more closely and in a better way than the Neyman type -A distribution.

Keywords: Neyman type-A distribution, Generalised Poisson distribution, Moments, Compounding, Estimation of parameters.

I. INTRODUCTION

The discrete probability distributions form a basic and promising field of study in the domain of Statistics and have many important applications in a wide variety of disciplines. It has been observed that there has been always an increasing amount of attention towards discrete distributions. Among all the distribution, the Poisson distribution given by its probability function

$$P(x) = \frac{e^{-\lambda} \lambda^x}{x!}; x=0,1,2,\dots; \lambda>0 \quad \dots \quad \dots \quad \dots \quad (1.1)$$

is found to be very much useful in various fields and due to small value of probability of an event, it is known as the distribution for rare events. This distribution is of so much importance that Fisher once remarked "Among all the distributions, the Poisson series is of first importance" [Haight (1967)]. Poisson distribution is one of the most widely used discrete distributions having applications in industry, agriculture and ecology, biology and medicine, telephony, accidents, commerce, queuing theory, sociology and demography, traffic flow, military science and the like. The equality of the mean and variance is an important characteristic of this distribution. The mean and variance are $\mu'_1 = \mu_2 = \lambda$. This distribution is a power series distribution with infinite non-negative integers support. It also belongs to the exponential family of distributions.

The Neyman Type-A distribution (NAD) was developed by Neyman (1939) to describe the distribution of larvae in experimental field plots. It is widely used in describing population under the influence of contagion and bacteriology. Pielou (1957) also have investigated the use of NAD in ecology.

The NAD is the Poisson – stopped – summed – Poisson distribution. The probability generating function (pgf) of this distribution has been obtained as

$$G(z) = e^{\left[\lambda \left\{ e^{-\phi(z-1)} - 1 \right\} \right]} \quad \dots \quad \dots \quad (1.2)$$

The distribution is the both a Poisson- stopped sum of Poisson distribution and also a Poisson mixture of Poisson distributions. The probability mass function (pmf) of this distribution has been obtained as

$$P(x=x) = \frac{e^{-\lambda} \phi^x}{x!} \sum_{j=0}^{\infty} \frac{(\lambda e^{-\phi})^j}{j!} j^x; x=0,1,2, \quad \dots \quad \dots \quad (1.3)$$

The first four moments about origin of the NAD have been obtained as

$$\left. \begin{aligned} \mu'_1 &= \lambda \phi \\ \mu'_2 &= \phi^2 \lambda^2 + \phi^2 \lambda + \phi \lambda \\ \mu'_3 &= \phi^3 \lambda^3 + 3\phi^3 \lambda^2 + \phi^3 \lambda + 3\phi^2 \lambda^2 + 3\phi^2 \lambda + \phi \lambda \\ \mu'_4 &= \phi^4 \lambda^4 + 6\phi^4 \lambda^3 + 7\phi^4 \lambda^2 + \phi^4 \lambda + 6\phi^3 \lambda^3 + 18\phi^3 \lambda^2 \\ &\quad + 6\phi^3 \lambda + 7\phi^2 \lambda^2 + 7\phi^2 \lambda + \phi \lambda \end{aligned} \right\} \quad \dots \quad \dots \quad (1.4)$$

The first four central moments of the NAD have been obtained as

$$\left. \begin{aligned} \mu_2 &= \lambda \phi (1 + \phi) \\ \mu_3 &= \lambda \phi (1 + 3\phi + \phi^2) \\ \mu_4 &= \lambda \phi (1 + 7\phi + 6\phi^2 + \phi^3) + 3\lambda^2 \phi^2 (1 + \phi)^2 \end{aligned} \right\} \dots \dots (1.5)$$

Skewness and Kurtosis of this distribution can be studied by the following moments ratios.

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(1 + 3\phi + \phi^2)^2}{\lambda \phi (1 + \phi)^3} \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{(1 + 7\phi + 6\phi^2 + \phi^3)}{\lambda \phi (1 + \phi)^2} \dots \dots (1.6)$$

Consul and Jain (1973) presented a new generalisation of Poisson distribution with two parameters λ and θ_1 given by its probability function

$$P(x) = \frac{\lambda(\lambda + x\theta_1)^{x-1} e^{-(\lambda + x\theta_1)}}{x!} \dots \dots$$

Where $x = 0, 1, 2, \dots; \lambda > 0; |\theta_1| < 1$

In many applied problems it is known in advance that the second parameter θ_1 in the GPD (1.7) is linearly proportional to the parameter λ . Taking $\theta_1 = \theta\lambda$ in (1.7) we get the restricted form of the GPD as

$$P(x) = \frac{\lambda^x (1 + x\theta)^{x-1} e^{-\lambda(1+x\theta)}}{x!} ; x = 0, 1, 2, \dots \dots \dots (1.8)$$

It can easily be seen that the classical Poisson distribution is a particular case of it and can be obtained by putting $\theta = 0$.

In the present paper a generalization of NAD has been obtained by mixing the classical Poisson distribution with the restricted generalised Poisson distribution (RGPD) of Consul and Jain (1973). The various aspects of the resultant distribution such as its moments, estimation of parameters have also been obtained. The details accounts of this distribution have been given by B.K.Sah (2013) in Ph.D. thesis entitled "Generalisations of countable and continuous mixtures of Poisson distribution and their applications".

II. MATERIAL AND METHODS

It is based on the concept of some countable mixtures of Poisson distribution. Probability mass function of the generalised Neyman Type-A distribution has been obtained by mixing generalised Poisson distribution with Poisson distribution (1) such that it follows basic properties of probability distributions. The first four moments about origin of DNAD has been obtained. The parameters of the proposed distribution have been obtained by using the first three moments about origin of the GNAD. It is expected that it is more flexible than the NAD (1.4) due to additional parameter θ for analyzing different types of count data.

III. GENERALISED NEYMAN TYPE- A DISTRIBUTION (GNAD)

The restricted GPD mixture of the classical Poisson distribution can symbolically be shown as

$$\text{Poisson (m)} \wedge_{\lambda/\phi=j} \text{Restricted GPD } (\lambda, \theta) \dots \dots (3.1)$$

The resultant distribution (3.1) can also be called generalized Neyman Type - A distribution (GNAD). Probability mass function of the GNAD (3.1) can be obtained as follows.

$$\begin{aligned} P[x; \phi, \lambda, \theta] &= \sum_{j=0}^{\infty} \frac{e^{-m} m^x}{x!} \lambda^j (1 + \theta j)^{j-1} e^{-\lambda(1+\theta j)} \\ &= \sum_{j=0}^{\infty} \frac{e^{-\phi j} (\phi j)^x}{x!} \cdot \frac{\lambda^j (1 + \theta j)^{j-1} e^{-\lambda(1+\theta j)}}{j!} \dots \dots (3.2) \end{aligned}$$

$$= \frac{e^{-\lambda \phi^x}}{x!} \sum_{j=0}^{\infty} \frac{[\lambda e^{-(\phi + \lambda \theta)}]^j}{j!} j^x (1 + \theta j)^{j-1} \dots \dots (3.3)$$

The relation (3.3) is the generalised Neyman Type - A distribution. It can easily be seen that at $\theta = 0$, this distribution is reduced to the NAD (1.4).

IV. MOMENTS OF GNAD

The r^{th} moment about origin of the generalised form (3.3) is obtained as

$$\begin{aligned}\mu'_r &= E \left[E \left(x^r / \frac{m}{\phi} = j \right) \right] \\ &= \sum_{j=0}^{\infty} \left[\sum_{x=0}^{\infty} \frac{x^r e^{-\phi j} (\phi j)^x}{x!} \right] \cdot \frac{\lambda^j (1+j\theta)^{j-1} e^{-\lambda(1+j\theta)}}{j!} \quad \dots \quad (4.1)\end{aligned}$$

Taking $r=1$ in (4.1), the first moment about origin of the GNAD can be obtained as follows.

$$\begin{aligned}\mu'_1 &= \sum_{j=0}^{\infty} \left[\sum_{x=0}^{\infty} x \frac{e^{-\phi j} (\phi j)^x}{x!} \right] \cdot \frac{\lambda^j (1+j\theta)^{j-1} e^{-\lambda(1+j\theta)}}{j!} \quad \dots \quad (4.2) \\ &= \phi \left[\sum_{j=0}^{\infty} j \frac{\lambda^j (1+j\theta)^{j-1} e^{-\lambda(1+j\theta)}}{j!} \right]\end{aligned}$$

The expression under bracket is the mean of the restricted GPD.

$$\text{i.e.} \quad \mu'_1 = \phi \cdot \frac{\lambda}{(1-\theta\lambda)} \quad \dots \quad (4.3)$$

Taking $r=2$ in (4.1), the second moment about origin of the GNAD can be obtained as follows.

$$\begin{aligned}\mu'_2 &= \sum_{j=0}^{\infty} \left[\sum_{x=0}^{\infty} \frac{x^2 e^{-\phi j} (\phi j)^x}{x!} \right] \text{Restricted GPD}(\lambda, \theta) \quad \dots \quad (4.4) \\ &= \sum_{j=0}^{\infty} (\phi^2 j^2 + \phi j) \text{RGPD}(\lambda, \theta) \\ &= \phi^2 \left[\sum_{j=0}^{\infty} j^2 \text{RGPD}(\lambda, \theta) \right] + \phi \left[\sum_{j=0}^{\infty} j \text{RGPD}(\lambda, \theta) \right]\end{aligned}$$

The expressions under bracket are respectively second and first moments about origin of the restricted GPD. Hence, the second moment about origin is given by

$$= \phi^2 \left[\frac{\lambda}{(1-\theta\lambda)^3} + \frac{\lambda^2}{(1-\theta\lambda)^2} \right] + \frac{\phi\lambda}{(1-\theta\lambda)} \quad \dots \quad (4.5)$$

Taking $r=3$ in (4.1), the third moment about origin of the GNAD can be obtained as follows.

$$\begin{aligned}\mu'_3 &= \sum_{j=0}^{\infty} \left[\sum_{x=0}^{\infty} \frac{x^3 e^{-\phi j} (\phi j)^x}{x!} \right] \text{RGPD}(\lambda, \theta) \quad \dots \quad (4.6) \\ &= \sum_{j=0}^{\infty} (\phi^3 j^3 + 3\phi^2 j^2 + \phi j) \text{RGPD}(\lambda, \theta)\end{aligned}$$

The expressions under bracket are respectively third, second and first moments about origin of the restricted GPD.

$$= \phi^3 \left[\frac{\lambda(1+2\theta\lambda)}{(1-\theta\lambda)^5} + \frac{3\lambda^2}{(1-\theta\lambda)^4} + \frac{\lambda^3}{(1-\theta\lambda)^3} \right] + 3\phi^2 \left[\frac{\lambda}{(1-\theta\lambda)^3} + \frac{\lambda^2}{(1-\theta\lambda)^2} \right] + \frac{\phi\lambda}{(1-\theta\lambda)} \quad \dots \quad (4.7)$$

Similarly, taking $r=4$ in (4.1), the fourth moment about origin of the GNAD can be obtained as follows.

$$\begin{aligned}\mu'_4 &= \sum_{j=0}^{\infty} \left[\sum_{x=0}^{\infty} \frac{x^4 e^{-\phi j} (\phi j)^x}{x!} \right] \cdot \text{RGPD}(\lambda, \theta) \quad \dots \quad (4.8) \\ &= \sum_{j=0}^{\infty} (\phi^4 j^4 + 6\phi^3 j^3 + 7\phi^2 j^2 + \phi j) \text{RGPD}(\lambda, \theta)\end{aligned}$$

The expressions under bracket are respectively the fourth, third, second and first moment about origin of the restricted GPD

$$\begin{aligned}
 &= \phi^4 \left[\frac{\lambda(1+8\theta\lambda+6\theta^2\lambda^2)}{(1-\theta\lambda)^7} + \frac{\lambda^2(7+8\theta\lambda)}{(1-\theta\lambda)^6} + \frac{6\lambda^3}{(1-\theta\lambda)^5} + \right. \\
 &\quad \left. \frac{\lambda^4}{(1-\theta\lambda)^4} \right] + 6\phi^3 \left[\frac{\lambda(1+2\theta\lambda)}{(1-\theta\lambda)^5} + \frac{3\lambda^2}{(1-\theta\lambda)^4} + \frac{\lambda^3}{(1-\theta\lambda)^3} \right] + \\
 &\quad 7\phi^2 \left[\frac{\lambda}{(1-\theta\lambda)^3} + \frac{\lambda^2}{(1-\theta\lambda)^2} \right] + \frac{\lambda\theta}{(1-\theta\lambda)} \quad \dots \quad \dots \quad (4.9)
 \end{aligned}$$

It can be verified that all these moments reduce to the respective moments of the NAD at $\theta = 0$.

The first four moments about mean of the GNAD can be obtained as

$$\begin{aligned}
 \mu_2 &= \mu'_2 - (\mu'_1)^2 \\
 &= \phi^2 \left[\frac{\lambda}{(1-\theta\lambda)^3} + \frac{\lambda^2}{(1-\theta\lambda)^2} \right] + \frac{\phi\lambda}{(1-\phi\lambda)} - \left[\frac{\phi\lambda}{(1-\phi\lambda)} \right]^2 \\
 &= \phi\lambda \left[\frac{\phi}{(1-\theta\lambda)^3} + \frac{1}{(1-\theta\lambda)} \right] \quad \dots \quad \dots \quad (4.10)
 \end{aligned}$$

The expression (4.10) is the variance of GNAD (3.3). At $\theta = 0$, the expression (4.10) is reduced to the variance of NAD (1.3).

The third moment about the mean of GNAD can be obtained as

$$\begin{aligned}
 \mu_3 &= \mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3 \\
 &= \phi \left[\frac{\lambda(1+2\theta\lambda)}{(1-\theta\lambda)^5} + \frac{3\lambda^2}{(1-\theta\lambda)^4} + \frac{\lambda^3}{(1-\theta\lambda)^3} \right] + 3\phi^2 \left[\frac{\lambda}{(1-\theta\lambda)^3} + \frac{\lambda^2}{(1-\theta\lambda)^2} \right] + \frac{\phi\lambda}{(1-\theta\lambda)} \\
 &\quad - 3 \left[\phi^2 \left[\frac{\lambda}{(1-\theta\lambda)^3} + \frac{\lambda^2}{(1-\theta\lambda)^2} \right] + \frac{\phi\lambda}{(1-\theta\lambda)} \right] \left[\frac{\phi\lambda}{(1-\theta\lambda)} \right]^2 + 2 \left[\frac{\phi\lambda}{(1-\theta\lambda)} \right]^3 \\
 &= \frac{\phi^3\lambda(1+2\theta\lambda)}{(1-\theta\lambda)^5} + \frac{3\phi^2\lambda}{(1-\theta\lambda)^3} + \frac{\phi\lambda}{(1-\theta\lambda)} \quad \dots \quad \dots \quad (4.11)
 \end{aligned}$$

It can be seen that at $\theta = 0$, the expression (4.11) is reduced to the third central moment of the NAD (1.3).

Similarly, the fourth moment about the mean can be obtained as follows

$$\begin{aligned}
 \mu_4 &= \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2(\mu'_1)^2 - 3(\mu'_1)^4 \\
 &= \frac{\phi^4\lambda(1+8\theta\lambda+6\theta^2\lambda^2)}{(1-\theta\lambda)^7} + \frac{\phi^4\lambda^2(7+8\theta\lambda)}{(1-\theta\lambda)^6} + \frac{6\phi^4\lambda^3}{(1-\theta\lambda)^5} + \frac{\phi^4\lambda^4}{(1-\theta\lambda)^4} \\
 &\quad + \frac{6\phi^3\lambda(1+2\theta\lambda)}{(1-\theta\lambda)^5} + \frac{18\phi^3\lambda^2}{(1-\theta\lambda)^4} + \frac{6\phi^3\lambda^3}{(1-\theta\lambda)^3} + \frac{7\phi^2\lambda}{(1-\theta\lambda)^3} + \frac{7\phi^2\lambda^2}{(1-\theta\lambda)^2} + \frac{\phi\lambda}{(1-\theta\lambda)}
 \end{aligned}$$

$$\begin{aligned}
& -4 \left[\frac{\phi^4 \lambda^2 (1+2\theta\lambda)}{(1-\theta\lambda)^6} + \frac{3\phi^4 \lambda^3}{(1-\theta\lambda)^5} + \frac{\phi^4 \lambda^4}{(1-\theta\lambda)^4} + \frac{3\phi^3 \lambda^2}{(1-\theta\lambda)^4} + \frac{3\phi^3 \lambda^3}{(1-\theta\lambda)^3} + \frac{\phi^2 \lambda^2}{(1-\theta\lambda)^2} \right] \\
& + \frac{6\phi^4 \lambda^3}{(1-\theta\lambda)^5} + \frac{6\phi^4 \lambda^4}{(1-\theta\lambda)^4} + \frac{6\phi^3 \lambda^3}{(1-\theta\lambda)^3} - \frac{3\phi^4 \lambda^4}{(1-\theta\lambda)^4} \\
& = \phi\lambda \left[\frac{1}{(1-\theta\lambda)} + \frac{7\phi}{(1-\theta\lambda)^3} + \frac{6\phi^2 (1+2\theta\lambda)}{(1-\theta\lambda)^5} + \frac{\phi^3 (1+8\theta\lambda+6\theta^2\lambda^2)}{(1-\theta\lambda)^7} \right] + 3\phi^2 \lambda^2 \left[\frac{1}{(1-\theta\lambda)^2} + \frac{2\phi}{(1-\theta\lambda)^4} + \frac{\phi^2}{(1-\theta\lambda)^6} \right] \dots \quad (4.12)
\end{aligned}$$

The expression (4.12) is the fourth central moment of the GNAD (3.3) . At $\theta = 0$, it is reduced to the fourth central moment of the NAD (1.3).

V. ESTIMATION OF PARAMETERS OF GNAD

The generalised mixture distribution (3.3) consists of three parameters ϕ, λ and θ and so for estimating these parameters by the method of moments, the first three moments are required. Let

$$s = \frac{1}{1-\theta\lambda} \Rightarrow \theta\lambda = \frac{s-1}{s} \quad \dots \quad (5.1)$$

The mean and variance of the GNAD can also be written in terms of s as follows.

$$\mu_1' = \lambda\phi s \quad \dots \quad (5.2)$$

$$\mu_2 = \lambda\phi^2 s^3 + \lambda\phi s \quad \dots \quad (5.3)$$

Dividing the equation (5.3) by (5.2) , we get

$$\frac{\mu_2}{\mu_1'} = K_1 (\text{say}) = \phi s^2 + 1$$

$$\text{i.e. } \phi = (K_1 - 1)s^{-2} \quad \dots \quad (5.4)$$

The third central moment of the GNAD can also be written in terms of s as

$$\mu_3 = \phi^3 \lambda (1+2\theta\lambda) s^5 + \lambda\phi s \quad \dots \quad (5.5)$$

$$\mu_2 - \mu_1' = \lambda\phi^2 s^3 \quad \dots \quad (5.6)$$

$$\mu_3 - \mu_1' = \phi^3 \lambda (1+2\theta\lambda) s^5 \quad \dots \quad (5.7)$$

Dividing the expression (5.7) by (5.6), we get

$$\frac{\mu_3 - \mu_1'}{\mu_2 - \mu_1'} = (1+2\theta\lambda)\phi s^2 = K_2 (\text{say})$$

$$\text{i.e. } \phi = \frac{K_2 s^{-2}}{(1+2\theta\lambda)} \quad \dots \quad (5.8)$$

Equating the expressions (5.8) and (5.4) , we get

$$(K_1 - 1)s^{-2} = \frac{K_2 s^{-2}}{(1+2\theta\lambda)}$$

$$\text{Or, } \theta\lambda = \frac{(K_3 - 1)}{2} \quad \dots \quad (5.9)$$

$$\text{And } (1-\theta\lambda) = \frac{(3-K_3)}{2} \quad \dots \quad (5.10)$$

Substituting the value of $(1-\theta\lambda)$ in the expression (5.1), we get

$$s = \frac{2}{3-K_3} \quad \dots \quad (5.11)$$

The population moments are replaced by the respective sample moments. To estimate value of s , we have to calculate the value of K_1 , K_2 and K_3 . Putting the value of s in (5.4), we can estimate the value of ϕ as

$$\phi = \frac{(K_1 - 1)(3 - K_3)^2}{4} \quad \dots \quad (5.12)$$

Substituting the estimated value of ϕ in the expression (5.2), we can estimate the value of λ as

$$\lambda = \frac{\mu_1'}{\phi s} \quad \dots \quad (5.13)$$

Substituting the estimated value of λ in the expression (5.9), value of θ can be obtained as follows.

$$\theta = \frac{(K_3 - 1)}{2\lambda} \quad \dots \quad (5.14)$$

Estimation of parameters provides to test goodness of fit of the discrete data sets having variance greater than its mean.

VI. CONCLUSION

The additional parameter θ of the GNAD (3.3) appears to play an important role in explaining the inequality between mean and variance. From expression (4.3) and (4.5), we have $\mu_2 > \mu_1'$ which is the characteristic of the negative binomial distribution. But for non-zero values of θ the lower limit for the variance becomes different from mean. For $0 < \theta < 1$, the lower limit of the variance is higher than the mean, obviously providing an improved lower limit of the variance. Similarly, for $-1 < \theta < 0$ the lower limit of the variance is lower than the mean. Thus, the parameter θ has capacity to stretch the lower limit of the variance up and down according to the observed distribution. Obviously for non-negative value of θ the inequality between mean and variance is explained in a better way than that explained by the Neyman Type-A distribution. Thus the parameter θ takes into account the inequality between mean and variance more closely than that taken by NAD and so it is expected that in all such cases where variance is greater than mean, the GNAD (3.3) would explain observed data more closely than the NAD.

VII. CONFLICT OF INTEREST

The author declared that there is no conflict of interest.

VIII. ACKNOWLEDGEMENT

The author expresses their gratitude to the referee for his valuable comments and suggestions which improved the quality of the research article. The hearty thank goes to Professor A. Mishra for his tremendous guidance and valuable suggestions.

REFERENCES

- [1] Consul, P.C., and Jain, G.C. (1973a). A generalisation of the Poisson distribution, *Technometrics*, 15, 791-799.
- [2] Haight, F.A. (1967). *Handbook of the Poisson distribution*, New York, Willey
- [3] Neyman, J. (1939). On a new class of "contagious" distributions applicable in entomology and bacteriology, *Annals of Mathematical Statistics*, 10, 35-57.
- [4] Pielou, E.C. (1957). The effect of quadrat size on the estimation of the parameters of Neyman's and Thomas's distributions, *Journal of Ecology*, 45, 31-47.
- [5] Sah, B.K. (2013). *Generalisations of Some Countable and Continuous Mixtures of Poisson Distribution and their Applications*, Doctoral Thesis, Patna University, Patna.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)