COM-Poisson Neyman Type A Distribution and its Properties

S. Thilagarathinam
Department of Mathematics
Nehru Memorial College
Trichy, Tamil Nadu, India.

V. Saavithri
Department of Mathematics
Nehru Memorial College
Trichy, Tamil Nadu, India.

R. Seethalakshmi
Department of Mathematics
Sastra University
Thanjavur, Tamil Nadu, India.

Abstract

In this paper, COM-Poisson Neyman type A distribution is introduced. COM-Poisson Neyman type A distribution is a Compound COM-Poisson distribution with Poisson compounding distribution. It is a generalization of Neyman type A distribution. Its properties are also studied. This distribution is found to be a better distribution for natural disaster data.

Key Words: Poisson distribution, Neyman type A distribution, COM-Poisson Distribution, COM-Poisson Neyman type A distribution.

1 Introduction

The COM-Poisson distribution is a generalization of Poisson distribution. It is the generalization of some well known discrete distributions like negative binomial, Bernoulli and Geometric distributions. In 1962, Conway & Maxwell introduced this distribution in the context of queuing systems. In 2005, Galit Shmueli revived this distribution and used for fitting discrete data. The COM-Poisson consists of an extra parameter, which we denote by ν and which governs the rate of decay of successive ratios of probabilities such that

$$\frac{P(X = x - 1)}{P(X = x)} = \frac{x^{\nu}}{\lambda}$$

.

Neyman(1939) constructed a statistical model of the distribution of larvae in a unit area of a field by assuming that the variation in the number of clusters of eggs per unit area could be represented by a Poisson distribution with parameter λ, while the number of larvae developing per clusters of eggs are assumed to have independent Poisson distribution all with the same parameter π.

In 1949, Thomas introduced a two-parameter counting distribution distinct from the Neyman type-A only in that mother pulses appeared in the final process along with daughter pulses. Although she originally referred to this as the double-Poisson distribution, it is called the Thomas distribution.

In 1943, Fellr, derived Neyman's distributions as compound Poisson distributions. This interpretation makes them suitable for modelling heterogeneous populations and renders them examples of apparent contagion. Neyman distribution also arise as generalized Poisson if the number of larvae observed at any plot are assumed to be hatched from Poisson distributed egg masses have some other discrete distribution.

In this paper, COM-Poisson Neyman type A distribution is introduced. COM-Poisson Neyman type A distribution is a Compound COM-Poisson distribution with Poisson compounding distribution. It is a generalization of Neyman type A distribution. Its properties are also studied.

In section 2,3, COM-Poisson and Neyman Type A Distributions are defined. In section 4, COM-Poisson Neyman Type A Distribution is defined and its mean and variane are obtained. Parameter estimations are derived in section 5.

2 COM-Poisson Distribution

The probability density function of COM-Poisson distribution is

$$P(X = x) = \frac{\lambda^x}{(x!)^\nu} \frac{1}{Z(\lambda, \nu)}, \qquad x = 0, 1, 2, ...$$

where $Z(\lambda, \nu) = \sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^\nu}$ for $\lambda > 0$ and $\nu \geq 0$.

The probability generating function of COM-Poisson distribution is

$$P_X(s) = \frac{Z(\lambda s, \nu)}{Z(\lambda, \nu)}$$

.

3 Neyman Type A Distribution

The random variable X has a Neyman type A distribution with parameters λ and ν. The probability generating function of Neyman type A distribution is

$$G_X(s) = e^{(\lambda e^{\nu(s-1)}-1)}$$

The probability mass function of Neyman type A distribution is

$$Pr\{X = x\} = \frac{\nu^x e^{-\nu}}{x!} \sum_{j=0}^{\infty} \frac{(\lambda e^{-\nu})^j}{j!} j^x, x = 1, 2, ...$$

4 Com-Poisson Neyman Type A Distribution

The COM-Poisson Neyman type A distribution arises in a model formed by supposing that objects (which are to be countable) occur in clusters. Suppose there are Y independent random variables of the form X, and N denotes the sum of these random variables, namely

$$N = X_1 + X_2 + ...X_Y$$

Then, the COM-Poisson Neyman type A distribution model is derived by supposing that
(i) X denotes the number of objects with in a cluster where $X \sim Poi(\lambda)$.
(ii) Y denotes the number of clusters, where $Y \sim Com - Poi(\mu, \nu)$.

This random variable N, formed by compounding in this fashion gives rise to the Com-poisson Neyman type A distribution and its probability generating function can be derived easily.

The probability generating function of X is known to be

$$G_X(s) = e^{\lambda(s-1)} \tag{1}$$

The probability mass function of Y is

$$P(Y = y) = \frac{\mu^y}{(y!)^\nu} \frac{1}{Z(\mu, \nu)}, y = 0, 1, 2, ...$$

where

$$Z(\mu, \nu) = \sum_{j=0}^{\infty} \frac{\mu^j}{(j!)^\nu}, \quad \text{for} \quad \mu, \nu \leq 0$$

The probability generating function of Y is

$$
\begin{aligned}
G_Y(s) = E(s^y) &= \sum_{y=1}^{\infty} s^y P(Y=y) \\
&= \sum_{y=1}^{\infty} s^y \frac{\mu^y}{(y!)^\nu} \frac{1}{Z(\mu,\nu)} \\
&= \frac{1}{Z(\mu,\nu)} \sum_{y=1}^{\infty} \frac{s^y \mu^y}{(y!)^\nu} \\
&= \frac{1}{Z(\mu,\nu)} \sum_{y=1}^{\infty} \frac{(s\mu)^y}{(y!)^\nu} \qquad (2) \\
G_Y(s) &= \frac{Z(\mu s,\nu)}{Z(\mu,\nu)}
\end{aligned}
$$

Since X_i's are iid and independent of Y, the Probability generating function of the random variable N can be readily found as follows

$$
\begin{aligned}
G_N(s) = E(s^N) &= E\left(s^{X_1+X_2+\ldots+X_y}\right) \\
&= \sum_{y=0}^{\infty} E\left(s^{X_1+X_2+\ldots+X_y|Y=y}\right) P(Y=y) \\
&= \sum_{y=0}^{\infty} [E(s)]^y P(Y=y) \\
&= \sum_{y=0}^{\infty} [G_X(s)]^y P(Y=y) = G_Y(G_X(s)) \\
&= \frac{Z(\mu G_X(s),\nu)}{Z(\mu,\nu)} \\
&= \frac{Z(\mu e^{\lambda(s-1)},\nu)}{Z(\mu,\nu)}
\end{aligned}
$$

$$
G_N(s) = \frac{1}{Z(\mu,\nu)} \sum_{j=0}^{\infty} \frac{\left(\mu e^{\lambda(s-1)}\right)^j}{(j!)^\nu} \qquad (3)
$$

Now, since the probability generating function of N in (3) can be expressed as

$$
\frac{1}{Z(\mu,\nu)} \sum_{j=0}^{\infty} \frac{\mu^j \left(e^{\lambda(s-1)}\right)^j}{(j!)^\nu} = \frac{1}{Z(\mu,\nu)} \sum_{j=0}^{\infty} \frac{\mu^j e^{-\lambda j} e^{\lambda j s}}{(j!)^\nu}
$$

upon collecting the coefficient of s^m in the above series, we find an explicit expression for the Probability mass function of N as.

$$P(N=m) = \frac{1}{Z(\mu,\nu)}\left[\sum_{i=1}^{m} \frac{\mu^i e^{-i\lambda}(i\lambda)^m}{(i!)^\nu m!}\right]$$

$$P(N=x) = \frac{\lambda^x}{Z(\mu,\nu)x!}\sum_{j=0}^{x} \frac{(\mu e^{-\lambda})^j (j)^x}{(j!)^\nu}$$

This is the probability mass function of the Com-Pisson Neyman type A distribution and denote it by $N \sim CPNTA(\mu,\lambda,\nu)$

Special Case

When $\nu = 1$ we have Neyman type A distribution.

5 Properties

The mean and variance can be calculated from the first and second derivatives of the probability generating function by setting $s = 1$

$$G'_N(s) = \frac{1}{Z(\mu,\nu)}\sum_{j=0}^{\infty} \frac{\mu^j e^{-\lambda j}}{(j!)^\nu}(\lambda j)e^{\lambda j s}$$

$$\text{Mean} = G'_X(1) = \frac{\lambda \sum_{j=0}^{\infty} \frac{j\mu^j}{(j!)^\nu}}{Z(\mu,\nu)}$$

$$G''_N(s) = \frac{1}{Z(\mu,\nu)}\sum_{j=0}^{\infty} \frac{\mu^j e^{-\lambda j}}{(j!)^\nu}(\lambda j)^2 e^{\lambda j s}$$

$$G''_N(1) = \frac{\lambda^2 \sum_{j=0}^{\infty} \frac{j^2\mu^j}{(j!)^\nu}}{Z(\mu,\nu)}$$

$$Var(N) = G''_X(1) + G'_X(1) - [G'_X(1)]^2$$

$$Var(N) = \frac{\lambda^2 \sum_{j=0}^{\infty} \frac{j^2\mu^j}{(j!)^\nu}}{Z(\mu,\nu)} + \frac{\lambda \sum_{j=0}^{\infty} \frac{j\mu^j}{(j!)^\nu}}{Z(\mu,\nu)} - \left[\frac{\lambda \sum_{j=0}^{\infty} \frac{j\mu^j}{(j!)^\nu}}{Z(\mu,\nu)}\right]^2$$

From these,we find the ratio between mean and variance to be

$$\frac{Var(N)}{Mean} = \frac{\lambda \sum_{j=0}^{\infty} \frac{j^2 \mu^j}{(j!)^\nu}}{\sum_{j=0}^{\infty} \frac{j \mu^j}{(j!)^\nu}} - \frac{\lambda \sum_{j=0}^{\infty} \frac{j \mu^j}{(j!)^\nu}}{Z(\mu, \nu)} + 1$$

6 Maximum Likelihood Estimation

Let $x_1, x_2, .., x_n$ be the samples follows the COM-Poisson Neyman Type A distribution with parameter $\lambda, \mu > 0$, and $\nu \geq 0$.

$$\begin{aligned} L &= \prod_{i=1}^{n} P(N = x_i) \\ &= \prod_{i=1}^{n} \left[\frac{\lambda^{x_i}}{Z(\mu, \nu)(x_i!)} \sum_{j=0}^{\infty} \frac{(\mu e^{-\lambda})^j (j)^{x_i}}{(j!)^\nu} \right] \\ &= \prod_{i=1}^{n} \left[\frac{\lambda^{x_i}}{Z(\mu, \nu)(x_i!)} \right] \prod_{i=1}^{n} \left[\sum_{j=0}^{\infty} \frac{(\mu e^{-\lambda})^j (j)^{x_i}}{(j!)^\nu} \right] \end{aligned}$$

The log likelihood function is

$$l = \log L = \log \sum_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} - n \log \left(\sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^\nu} \right) + \sum_{i=1}^{n} \log \left(\sum_{j=0}^{n} \frac{\mu^j}{(j!)^\nu} (e^{-\lambda})^j (j)^{x_i} \right)$$

Differentiating l partially with respect to λ, μ and ν and equating to zero, we get.

$$\frac{\sum_{i=1}^{n} \frac{x_i \lambda^{x_i - 1}}{x_i!}}{\sum_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!}} - \sum_{i=1}^{n} \left[\frac{\sum_{j=0}^{\infty} \frac{\mu^j}{(j!)^\nu} j (e^{-\lambda})^{j-1} (e^{-\lambda}) (j)^{x_i}}{\sum_{j=0}^{\infty} \frac{\mu^j}{(j!)^\nu} (e^{-\lambda})^j (j)^{x_i}} \right] = 0$$

$$-n \left[\frac{\sum_{j=0}^{\infty} \frac{j\mu^{j-1}}{(j!)^\nu}}{\sum_{j=0}^{\infty} \frac{\mu^j}{(j!)^\nu}} \right] + \sum_{i=1}^{n} \left[\frac{\sum_{j=0}^{\infty} \frac{j\mu^{j-1}}{(j!)^\nu} (e^{-\lambda})^j (j)^{x_i}}{\sum_{j=0}^{\infty} \frac{\mu^j}{(j!)^\nu} (e^{-\lambda})^j (j)^{x_i}} \right] = 0$$

$$-n \left[\frac{\sum_{j=0}^{\infty} \frac{\mu^j \log(j!)}{(j!)^\nu}}{\sum_{j=0}^{\infty} \frac{\mu^j}{(j!)^\nu}} \right] + \sum_{i=1}^{n} \left[\frac{\sum_{j=0}^{\infty} \frac{\mu^j \log(j!)}{(j!)^\nu} (e^{-\lambda})^j (j)^{x_i}}{\sum_{j=0}^{\infty} \frac{\mu^j}{(j!)^\nu} (e^{-\lambda})^j (j)^{x_i}} \right] = 0$$

Solving the above three equations using numerical techniques, estimators of λ, μ and ν can be obtained.

References

[1] Conway, R.W. and W.L.Maxwell(1962). *A queuing model with state dependent service rates.* Journal of Industrial Engineering, 12, pp. 132-136.

[2] Shmeli G, Minka T.P, Kadane J.B, Borle S and Boatwright P (2005): *A useful distribution for fitting discrete data:Revival of the COM-Poisson ditribution*, J.R.Stat.Soc.Ser.C(Appl. Stat), 54, 127-142.

[3] Davit,F.N, and Moore,P.G.,(1954), *Notes on contagious distribution in plant populations*, annals of Botany, New series, 18, 47-53,[9.6.1-9.6.5].

[4] Cernuschi,F, and Cstangetto,L.,(1946) *Chains of rare events*, Annals of Mathematical statistics, 17, 53-61,[9.6.1].

[5] Shenton,L.R., and Bowman,K.O.,(1967), *Remark on large sample estimators for some discrete distributions*, Teechnoneetrics, 9, 581-598, [9.6.1,9.6.4.]

[6] Barton,D.E,(1957), *The Modality of Neyman's Contagious distribution of Type A*, Trabajors de Estadistica, 8, 13-22, [9.6.1].

[7] Priyadharshini J., and Saavithri,V., *COM-Poisson Polya-Aeppli Process*, Paper communicated to the journal of Statistics and probability letters.