Compound COM-Poisson Distribution with Binomial Compounding Distribution

V.Saavithri
Department of Mathematics
Nehru Memorial College
Trichy.

J.Priyadharshini
Department of Mathematics
Nehru Memorial College
Trichy.

Z. Parvin Banu
Department of Mathematics
Nehru Memorial College
Trichy.

February 26, 2018

Abstract

Conway-Maxwell Poisson distribution is two parameter Poisson distribution and also a generalization of Bernoulli, Geometric distributions. In this paper, the compound COM-Poisson distribution with binomial compounding distribution is proposed. Its properties are also derived. The parameters of newly introduced distribution are estimated by the method of profile likelihood estimation.

1 Introduction

In 1952, Poisson - Binomial distribution was discussed by Skellam and fitted by him to quadrat data on the sedge Carex flacca. In McGuire et al. this distribution was used to represent variation in the numbers of corn-borer larvae in randomly chosen areas of a field.

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. They use the acronym COM-Poisson for this distribution.

The COM-Poisson belongs to the exponential family as well as to the two-parameter power series distributions family. Its even stronger, it is easy to use flexible for fitting over and under-dispersed data.

In many practical applications, the equidispersion property of the Poisson distribution is not observed in the count data at hand, it motivates the search for more flexible models for this type of data.

In this paper we consider the COM-Poisson distribution and study the mean and variance. In this paper we define the Compound COM-Poisson distribution with binomial compounding distribution. Its properties are also derived.

This paper is laid out as follows: Section 2 describes the study of COM-Poisson distribution. In section 3, the Compound COM-Poisson distribution with binomial compounding distribution is defined and discussed some of its properties. Section 4 deals with the profile likelihood estimation of the newly introduced distribution. Section 5 concludes this paper.

2 COM-Poisson Distribution

The probability density function of COM-Poisson distribution [7] is

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

where $Z(\lambda,\nu)=\sum\limits_{j=0}^{\infty}\dfrac{\lambda^j}{(j!)^\nu}$ for $\lambda>0$ and $\nu\geq 0$.
Here the parameter ν governs the rate of decay of successive ratios of probabilities such that

$$\frac{P(X=x-1)}{P(X=x)}=\frac{x^\nu}{\lambda} \tag{2}$$

The probability generating function of COM-Poisson distribution is

$$G_X(s)=\frac{Z(\lambda s,\nu)}{Z(\lambda,\nu)} \tag{3}$$

The mean and variance are

$$Mean(X)=G_X^{'}(1)=\frac{\lambda Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)}$$

$$Var(X)=\frac{\lambda^2 Z_{\lambda\lambda}(\lambda,\nu)}{Z(\lambda,\nu)}+\frac{\lambda Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)}-\left[\frac{\lambda Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)}\right]^2$$

$$\text{where}\quad Z_\lambda(\lambda,\nu)\equiv\frac{d}{d\lambda}[Z(\lambda,\nu)],$$

$$Z_{\lambda\lambda}(\lambda,\nu)\equiv\frac{d^2}{d\lambda^2}[Z(\lambda,\nu)]$$

3 Compound COM-Poisson Distribution with Binomial Compounding Distribution

Consider the several events that can happen simultaneously at an instant, we have a cluster (of occurences) at a point.
Assume that there are Y independent random variables of the form X, and N denotes the sum of these random variables.
(ie)

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

Then, the compound COM-Poisson binomial model is derived by the following assumptions

(i) X denotes the number of objects within a cluster and it follows binomial distribution with parameters (n, p)
(ie)

$$X \sim Binomial(n, p)$$

(ii) Y denotes the number of clusters and it follows COM-Poisson distribution with parameters λ and ν.
(ie)

$$Y \sim CMP(\lambda, \nu)$$

This random variable, N formed by compounding these two random variables X and Y gives the Compound COM-Poisson Distribution with Binomial Compounding Distribution.
Its probability generating function (PGF) can be derived as follows.
The probability mass function (PMF) of X is

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, \qquad x = 0, 1, 2, ...$$

where $p > 0, q > 0, p + q = 1$
Its probability generating function is

$$\begin{aligned} G_X(s) = E(s^X) &= \sum_{x=1}^{\infty} s^x P(X = x) \\ &= \sum_{x=1}^{\infty} s^x \binom{n}{x} p^x q^{n-x} \\ &= \sum_{x=1}^{\infty} \binom{n}{x} (ps)^x q^{n-x} \\ &= (q + ps)^n \end{aligned}$$

$$G_X(s) = (q + ps)^n \tag{4}$$

Also, the random variable Y having probability mass function in (1) and the probability generating function is

$$G_Y(s) = \frac{Z(\lambda s, \nu)}{Z(\lambda, \nu)} \tag{5}$$

Since $X_i's$ are iid and independent of Y, probability generating function of the random variable N is given by

$$\begin{aligned} G_N(s) = E(s^N) &= E(s^{X_1+X_2+\ldots+X_Y}) \\ &= \sum_{y=0}^{\infty} E(s^{X_1+X_2+\ldots+X_Y}/Y=y)P(Y=y) \\ &= \sum_{y=0}^{\infty} [E(s^x)]^y P(Y=y) \\ &= G_Y(G_X(s)) \\ &= \frac{Z(\lambda G_X(s), \nu)}{Z(\lambda, \nu)} \\ &= \frac{1}{Z(\lambda,\nu)} \sum_{j=0}^{\infty} \frac{[\lambda(q+ps)^n]^j}{(j!)^\nu} \\ &= \frac{1}{Z(\lambda,\nu)} \sum_{j=0}^{\infty} \frac{\lambda^j \, [q+ps]^{nj}}{(j!)^\nu} \end{aligned} \tag{6}$$

Expanding the summation and collecting the coefficient of s^m in the above equation we get

$$P(N=m) = \frac{1}{Z(\lambda,\nu)} \sum_{j \geq m/n} \frac{\lambda^j}{(j!)^\nu} \binom{nj}{m} (p)^m q^{nj-m} \qquad \text{for} \quad m = 1, 2, \ldots$$

The probability mass function of N is

$$P(N=m) = \begin{cases} \dfrac{1}{Z(\lambda,\nu)} & \text{for} \quad m = 0 \\ \dfrac{1}{Z(\lambda,\nu)} \sum_{j \geq m/n} \dfrac{\lambda^j}{(j!)^\nu} \binom{nj}{m} (p)^m q^{nj-m} & \text{for} \quad m = 1, 2, \ldots \end{cases} \tag{7}$$

The mean and variance are derived as follows

$$\begin{aligned} G_N^{'}(s) &= \frac{1}{Z(\lambda,\nu)} \sum_{j=0}^{\infty} \frac{j\,[\lambda(q+ps)^n]^{j-1}\,\lambda np(q+ps)^{n-1}}{(j!)^\nu} \\ &= \frac{\lambda np(q+ps)^{n-1}}{Z(\lambda,\nu)} \sum_{j=0}^{\infty} \frac{j\,[\lambda(q+ps)^n]^{j-1}}{(j!)^\nu} \\ G_N^{'}(1) &= \frac{\lambda np}{Z(\lambda,\nu)} \sum_{j=0}^{\infty} \frac{j(\lambda)^{j-1}}{(j!)^\nu} \end{aligned}$$

$$Mean(N) = G_N^{'}(1) = \frac{\lambda np Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)} \tag{8}$$

$$\begin{aligned}
G_N^{''}(s) &= \frac{\partial}{\partial s}\left[\frac{1}{Z(\lambda,\nu)}\sum_{j=0}^{\infty}\frac{j\left[\lambda(q+ps)^n\right]^{j-1}\lambda np(q+ps)^{n-1}}{(j!)^\nu}\right] \\
&= \frac{\lambda np}{Z(\lambda,\nu)}\left[(q+ps)^{n-1}\sum_{j=0}^{\infty}\frac{j(j-i)\left[\lambda(q+ps)^n\right]^{j-2}\lambda n(q+ps)^{n-1}p}{(j!)^\nu}\right] \\
&\quad + \frac{\lambda np}{Z(\lambda,\nu)}\left[\sum_{j=0}^{\infty}\frac{j\left[\lambda(q+ps)^n\right]^{j-1}(n-1)p(q+ps)^{n-2}}{(j!)^\nu}\right] \\
&= \frac{\lambda np}{Z(\lambda,\nu)}\left[\lambda np\sum_{j=0}^{\infty}\frac{j(j-1)\lambda^{j-2}}{(j!)^\nu} + (n-1)p\sum_{j=0}^{\infty}\frac{j\lambda^{j-1}}{(j!)^\nu}\right] \\
G_N^{''}(1) &= \frac{\lambda np}{Z(\lambda,\nu)}\left[\lambda np Z_{\lambda\lambda}(\lambda,\nu) + (n-1)pZ_\lambda(\lambda,\nu)\right] \\
Var(N) &= G_N^{''}(1) + G_N^{'}(1) - [G_N^{'}(1)]^2 \\
&= \frac{\lambda np}{Z(\lambda,\nu)}\left[\lambda np Z_{\lambda\lambda}(\lambda,\nu) + (n-1)pZ_\lambda(\lambda,\nu)\right] + \frac{\lambda np Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)} - \left[\frac{\lambda np Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)}\right]^2 \\
&= \frac{\lambda np}{Z(\lambda,\nu)}\left[\lambda np Z_{\lambda\lambda}(\lambda,\nu) + (np-p+1)Z_\lambda(\lambda,\nu) - \frac{\lambda np\left[Z_\lambda(\lambda,\nu)\right]^2}{Z(\lambda,\nu)}\right]
\end{aligned}$$

$$Var(N) = \frac{\lambda np}{Z(\lambda,\nu)}\left[\lambda np Z_{\lambda\lambda}(\lambda,\nu) + (np+q)Z_\lambda(\lambda,\nu) - \frac{\lambda np\left[Z_\lambda(\lambda,\nu)\right]^2}{Z(\lambda,\nu)}\right] \tag{9}$$

The ratio between variance and mean is

$$\begin{aligned}
Ratio &= \frac{Var(N)}{Mean(N)} \\
&= \frac{\frac{\lambda np}{Z(\lambda,\nu)}\left[\lambda np Z_{\lambda\lambda}(\lambda,\nu) + (np+q)Z_\lambda(\lambda,\nu) - \frac{\lambda np\left[Z_\lambda(\lambda,\nu)\right]^2}{Z(\lambda,\nu)}\right]}{\frac{\lambda np Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)}} \\
&= \lambda np\left[\frac{Z_{\lambda\lambda}(\lambda,\nu)}{Z_\lambda(\lambda,\nu)} - \frac{Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)}\right] + (np+q)
\end{aligned}$$

4 Profile Likelihood Estimation

Let $N_1, N_2, ..., N_m$ be the samples follows the Compound COM-Poisson distribution with binomial compounding distribution with parameters $\lambda > 0, \nu \geq 0, n > 0$ and $p > 0$.

$$
\begin{aligned}
L &= \prod_{i=1}^{m} P(N = N_i) \\
&= \prod_{i=1}^{m} \frac{1}{Z(\lambda,\nu)} \sum_{j \geq N_i/n} \frac{\lambda^j}{(j!)^\nu} \binom{nj}{N_i} (p)^{N_i} (1-p)^{nj-N_i}
\end{aligned}
$$

The log likelihood function is

$$
l = logL = -mlog\left[\sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^\nu}\right] + \sum_{i=1}^{m} log\left[\sum_{j \geq N_i/n} \frac{\lambda^j}{(j!)^\nu} \binom{nj}{N_i} (p)^{N_i} (1-p)^{nj-N_i}\right] \tag{10}
$$

Differentiating equation (10) partially with respect to λ and equating to zero

$$
\frac{\partial l}{\partial \lambda} = 0
$$

$$
0 = \sum_{i=1}^{m} \left[\frac{\sum\limits_{j \geq N_i/n} \frac{j\lambda^{j-1}}{(j!)^\nu} \binom{nj}{N_i} (p)^{N_i} (1-p)^{nj-N_i}}{\sum\limits_{j \geq N_i/n} \frac{\lambda^j}{(j!)^\nu} \binom{nj}{N_i} (p)^{N_i} (1-p)^{nj-N_i}}\right] - m\left[\frac{\sum\limits_{j=0}^{\infty} \frac{j\lambda^{j-1}}{(j!)^\nu}}{\sum\limits_{j=0}^{\infty} \frac{\lambda^j}{(j!)^\nu}}\right]
$$

Differentiating equation (10) partially with respect to p and equating to zero

$$
\frac{\partial l}{\partial p} = 0
$$

$$
0 = \sum_{i=1}^{m} \left[\frac{\sum\limits_{j \geq N_i/n} \frac{\lambda^j}{(j!)^\nu} \binom{nj}{N_i} (p)^{N_i} (1-p)^{nj-N_i} \left[\frac{N_i}{p} - \frac{nj - N_i}{1-p}\right]}{\sum\limits_{j \geq N_i/n} \frac{\lambda^j}{(j!)^\nu} \binom{nj}{N_i} (p)^{N_i} (1-p)^{nj-N_i}}\right]
$$

Then $\max\limits_{\lambda,\nu,n,p} log(\lambda,\nu,n,p|m) = \max\limits_{\nu}\left[\max\limits_{\lambda,n,p} log(\lambda,n,p|m)\right]$ is calculated.

5 Conclusion

The Compound COM-Poisson distribution with binomial compounding distribution is introduced and its properties are derived. Also the parameters of the newly introduced distribution are estimated. This distribution can be applied to Bacterial count data.

References

[1] CONSUL P.C (1989) : Generalized Poisson Distributions: Properties and Applications, Marcel Dekker Inc., New York/Basel.

[2] CONWAY R.W, AND W.L. MAXWELL (1962): A queuing model with state dependent service rates, Journal of Industrial Engineering, 12, pp.132-136.

[3] JOHNSON N.L, KOTZ S AND KEMP A.W (2005): Univariate Discrete Distributions, 3rd edition, Wiley Series in Probability and Mathematical Sciences.

[4] MACEDA E.C (1948) : On the Compound and Generalized Poisson Distributions, Annals of Mathematical Statistics, Vol.19, No.3. (sept), pp. 414-416.

[5] MCGUIRE J.U, BRINDLEY T.A, and BANCROFT T.A (1957) : The distribution of European corn borer Pyrausta Nubilalis (Hbn.) in field corn, Biometrics, 13, 65-78 [errata and extensions (1958) 14, 432-434]. [9.4, 9.5]

[6] MEDHI J (2002): Stochastic Processes, 2nd edition, New Age International (P) Ltd., Publishers.

[7] 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.

[8] SKELLAM J.G. (1952) : Studies in statistical ecology I: Spatial pattern, Biometrika, 39, 346-362. [9.4, 9.5, 9.8]