International Journal of Research in Advent Technology, Vol.7, No.1, January 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

Com-Poisson Thomas Distribution

J.Priyadharshini[1], V.Saavithri[2]
Research Scholar, Department of Mathematics, Nehru Memorial College, Trichy[1].
Assistant Professor, Department of Mathematics, Nehru Memorial College, Trichy[2].
Email: priyajaya895@gmail.com[1], saavithriramani@gmail.com[2]

Abstract- COM-Poisson distribution is a generalization of Poisson, Bernoulli and geometric distributions. Thomas distribution is a compound Poisson distribution with shifted Poisson compounding distribution. In this paper COM-Poisson Thomas distribution, which is a compound COM-Poisson distribution with shifted Poisson compounding distribution, is introduced. This distribution is used to analyze the traffic accident data.

Subject Classification: Primary 60G99 ; Secondary 60K05,60K10.

Keywords: Poisson distribution, Shifted Poisson, COM-Poisson distribution, Thomas distribution, COM-Poisson Thomas distribution.

1. INTRODUCTION

COM-Poisson distribution is a two parameter extension of Possion distribution. This distribution is also a generalization of some well known distributions namely Bernoulli and geometric. This distribution is introduced by Conway & Maxwell [1] in queuing system in 1962. In 2005, Shmueli at el [7] revived this distribution. This distribution is used for both over and under dispersed data.

Neyman [5] 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 ϕ [3][4].

In 1949, Thomas [8] proposed the compound Poisson distribution with compounding shifted Poisson distribution in constructing a model for the distribution of plants of a given species in randomly placed quadrate. Thomas called this distribution a ``double Poisson" distribution, though Douglas [2] has pointed out that the term applies more appropriately to a Neyman type A distribution than to a Thomas distribution.

In 1972, Ord [6] derived the moments for this distribution.

In this paper, COM-Poisson Thomas distribution, which is a compound COM-Poisson distribution with shifted Poisson compounding distribution, is introduced. This distribution is used to analyze the traffic accident data.

This paper is organized as follows: Section 2 describes the study of COM-Poisson distribution and shifted Poisson distribution. In section 3, Thomas distribution is studied and some of its properties are discussed. The COM-Poisson Thomas distribution is defined and some of its properties are derived in section 4. In Section 5, the maximum likelihood estimator of COM-Poisson Thomas distribution is derived. Traffic accidents and fatalities data is analyzed in section 6. Section 7 concludes this paper.

2. COM-POISSON DISTRIBUTION

The COM-Poisson has an extra parameter, denoted by ν, which governs the rate of decay of successive ratio of probabilities such that

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

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

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

where

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

The probability generating function of COM-Poisson distribution is

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

SHIFTED POISSON DISTRIBUTION

Consider the random variable $U \sim Poisson(\phi)$

$$P(U = u) = \frac{e^{-\phi}\phi^u}{(u)!} \quad u = 0,1,2,3\ldots.$$

Now consider, $X = U + 1$, (ie), X is just shifted 1 to the right

∴ The probability mass function of the corresponding shifted Poisson distribution is

$$P(X=x)=\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \quad x = 1,2,3.....$$

The probability generating function is

$$G(s) = se^{\lambda(s-1)}$$

3. THOMAS DISTRIBUTION

Suppose that the several events can happen simultaneously at an instant, then there is a cluster of occurrence 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) $N = X_1 + X_2 + \ldots + X_Y$

Then, the Thomas distribution [8] is derived by assuming that

1. X represents the number of objects within a cluster and X follows shifted Poisson distribution with parameter ϕ

   (ie) $X \sim Shifted\ Poisson\ (\phi)$

2. Y represents the number of clusters and Y follows Poisson distribution with parameter λ.
   (ie) $Y \sim Poisson(\lambda)$

This random variable, N formed by compounding these two random variables X and Y gives the Thomas distribution with parameters λ and ϕ.

Probability generating function (PGF) is

$$G_N(s) = exp[\lambda\,(se^{\phi\,(s-1)}-1)]$$

The probability mass function of N is

$$P(N=n) = \begin{cases} e^{-\lambda} & for\ n=0 \\ \dfrac{e^{-\lambda}}{n!}\displaystyle\sum_{j=1}^{n}\binom{n}{j}(\lambda e^{-\phi})^j (j\phi)^{n-j}, & for\ n=1,2,\ldots \end{cases}$$

4. COM-POISSON THOMAS DISTRIBUTION

Suppose that the several events can happen simultaneously at an instant, then there is a cluster of occurrence 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) $N = X_1 + X_2 + \ldots + X_Y$

COM-Poisson Thomas distribution is derived by assuming that

1. X denotes the number of objects within a cluster and X follows shifted Poisson distribution with parameter ϕ.
   (ie) $X \sim Shifted\ Poisson(\phi)$
2. Y denotes the number of clusters and Y follows COM-Poisson distribution with parameters λ and ν.

   (ie) $Y \sim COM-Poisson(\lambda,\nu)$

This random variable, N formed by compounding these two random variables X and Y gives the COM-Poisson Thomas distribution with parameters λ,ν and ϕ.

The probability generating function of X is,

$$G_X(s) = se^{\phi\,(s-1)}$$

The probability generating function of Y is

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

The probability generating function of the random variable N can be derived as follows

$$G_N\,(s) = E(s^N) = E(s^{X_1+X_2+\ldots+X_Y})$$
$$= \sum_{y=0}^{\infty} E(s^{X_1+X_2+\ldots+X_Y}/Y=y)P(Y=$$
$$= \sum_{y=0}^{\infty}[E(s^x)]^y\,)\,P(Y=y$$
$$= G_Y\left(G_X\,(s)\right)$$
$$= \frac{Z(\lambda G_X\,(s),\nu)}{Z(\lambda\,,\nu)}$$
$$= \frac{1}{Z(\lambda,\nu)}\}\sum_{j=1}^{\infty}\frac{[\lambda\,se^{\phi\,(s-1)}]^j}{j!^{\nu}}$$

Collecting the coefficient of $\boldsymbol{s^n}$ in $\boldsymbol{G_N}\,(\boldsymbol{s})$ we get

$$P(N=n)=\frac{1}{Z(\lambda,\nu)}\sum_{j=1}^{n}\frac{\left(\lambda e^{-\phi}\right)^j (j\phi)^{n-j}}{j!^{\nu}\,(n-j)!}$$

∴ The probability mass function of $\boldsymbol{N}$ is

$$P(N=n) = \begin{cases} \dfrac{1}{Z(\lambda,\nu)} & for\ n=0 \\ \dfrac{1}{Z(\lambda,\nu)}\displaystyle\sum_{j=1}^{n}\frac{\left(\lambda e^{-\phi}\right)^j (j\phi)^{n-j}}{j!^{\nu}\,(n-j)!}, & for\ n=1,2,\ldots \end{cases}$$

where $\lambda > 0, \nu \geq 0$ and $\phi > 0$.

PROPERTIES OF COM-POISSON THOMAS DISTRIBUTION

The mean and variance are

$$Mean(N) = \frac{\lambda(1+\phi)Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)}$$

$$Var(N) = \frac{1}{Z(\lambda,\nu)}\left[\lambda^2(1+\phi)\left(Z_{\lambda\lambda}(\lambda,\nu) - \frac{[Z_\lambda(\lambda,\nu)]^2}{Z(\lambda,\nu)}\right)\right] + \frac{1}{Z(\lambda,\nu)}[\lambda^2(\phi^2+3\phi+1)Z_\lambda(\lambda,\nu)]$$

The expression for ratio between variance and mean is

$$\frac{Var(N)}{Mean(N)} = \lambda\left[\frac{\lambda Z_{\lambda\lambda}(\lambda,\nu)}{Z_\lambda(\lambda,\nu)} - \frac{\lambda Z_\lambda(\lambda,\nu)}{Z(\lambda,\nu)}\right] + \frac{(\phi^2+3\phi+1)}{(1+\phi)}$$

The factorial moments are,

$$\mu'_{(1)} = \frac{1}{Z(\lambda,\nu)}[\lambda(1+\phi)Z_\lambda(\lambda,\nu)]$$

$$\mu'_{(2)} = \frac{1}{Z(\lambda,\nu)}[\lambda^2(1+\phi)^2 Z_{\lambda\lambda}(\lambda,\nu) + \lambda\phi(2+\phi)Z_\lambda(\lambda,\nu)]$$

$$\mu'_{(3)} = \frac{1}{Z(\lambda,\nu)}[\lambda^3(1+\phi)^3 Z_{\lambda\lambda\lambda}(\lambda,\nu) + \lambda^2\phi(2+\phi)(1+\phi)Z_{\lambda\lambda}(\lambda,\nu) + \lambda\phi^2(3+\phi)Z_\lambda(\lambda,\nu)]$$

$$\mu'_{(4)} = \frac{1}{Z(\lambda,\nu)}[\lambda^4(1+\phi)^4 Z_{\lambda\lambda\lambda\lambda}(\lambda,\nu) + 6\lambda^3\phi(2+\phi)(1+\phi)Z_{\lambda\lambda\lambda}(\lambda,\nu) + \lambda^2\phi^2(7\phi^2+28\phi+24)Z_{\lambda\lambda}(\lambda,\nu) + \lambda\phi^3(4+\phi)Z_\lambda(\lambda,\nu)]$$

5. MAXIMUM LIKELIHOOD ESTIMATION

Let $N_1, N_2, \ldots, N_n$ be the independent samples follows the COM-Poisson Thomas distribution with parameters $\lambda > 0, \nu \geq 0$ and $\phi > 0$. The likelihood function of $N_1, N_2, \ldots, N_n$ is

$$L = \prod_{i=1}^{n} P(N=N_i)$$

$$= \prod_{i=1}^{n}\frac{1}{Z(\lambda,\nu)}\sum_{j=1}^{N_i}\frac{\left(\lambda e^{-\phi}\right)^j (j\phi)^{N_i-j}}{j!^{\nu}\,(N_i-j)!}$$

$$= \frac{1}{[Z(\lambda,\nu)]^n}\prod_{i=1}^{n}\sum_{j=1}^{N_i}\frac{\left(\lambda e^{-\phi}\right)^j (j\phi)^{N_i-j}}{j!^{\nu}\,(N_i-j)!}$$

The log likelihood function is

$$logL = l = -nlog\left[\sum_{j=0}^{\infty}\frac{\lambda^j}{j!^{\nu}}\right] + \sum_{i=1}^{n} log\left[\sum_{j=1}^{N_i}\frac{\left(\lambda e^{-\phi}\right)^j (j\phi)^{N_i-j}}{j!^{\nu}\,(N_i-j)!}\right]$$

The estimators for $\boldsymbol{\lambda}, \boldsymbol{\nu}$ and $\boldsymbol{\phi}$ are

$$\sum_{i=1}^{n}\frac{I_2(N_i)}{I_1(N_i)} - n\frac{J_2}{J_1} = 0$$

$$n\frac{J_3}{J_1} - \sum_{i=1}^{n}\frac{I_3(N_i)}{I_1(N_i)} = 0$$

$$\sum_{i=1}^{n}\frac{I_4(N_i)}{I_1(N_i)} = 0$$

where

$$I_1(N_i) = \sum_{j=1}^{N_i}\frac{\left(\lambda e^{-\phi}\right)^j (j\phi)^{N_i-j}}{j!^{\nu}\,(N_i-j)!}$$

$$I_2(N_i) = \sum_{j=1}^{N_i}\frac{j\lambda^{j-1}e^{-j\phi}(j\phi)^{N_i-j}}{j!^{\nu}\,(N_i-j)!}$$

$$I_3(N_i) = \sum_{j=1}^{N_i}\frac{\left(\lambda e^{-\phi}\right)^j \log(j!)\,(j\phi)^{N_i-j}}{j!^{\nu}\,(N_i-j)!}$$

$$I_2(N_i) = \sum_{j=1}^{N_i}\frac{j\lambda^j e^{-j\phi}(j\phi)^{N_i-j-1}(N_i-j-j\phi)}{j!^{\nu}\,(N_i-j)!}$$

$$J_1 = \sum_{j=0}^{\infty}\frac{\lambda^j}{j!^{\nu}}$$

$$J_2 = \sum_{j=0}^{\infty} \frac{j\lambda^{j-1}}{j!^{\nu}}$$

$$J_3 = \sum_{j=0}^{\infty} \frac{\lambda^j \log(g!)}{j!^{\nu}}$$

6. DATA ANALYSIS

The data is taken from fatal crashes and fatalities calender 2016 of Texas department of transportation, Austin. The one day accidents (left entry) and the corresponding number of fatalities (right entry) for each month during the year 2016 is given in the table 6.1.

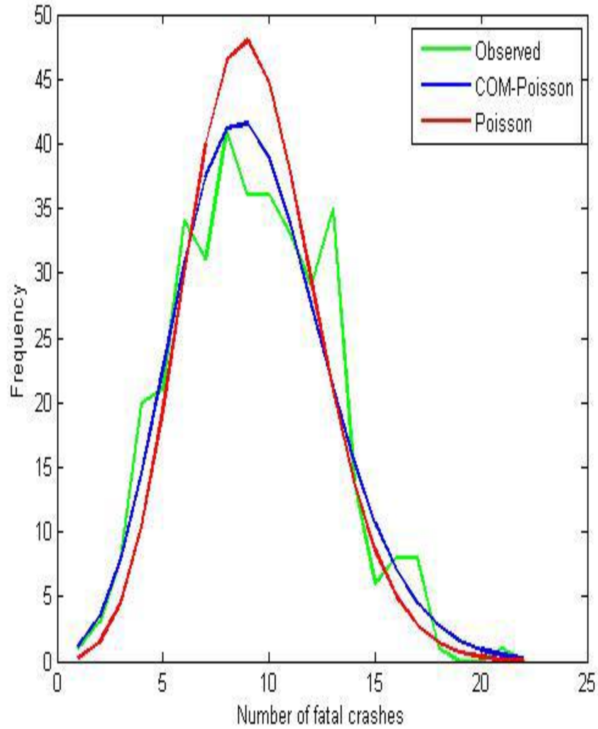
Let Y be the number of day's that accidents occurred at the year 2016.

$X_i, i = 1,2,\ldots,$ be the number of fatalities of ith accident and N be the total number of fatalities from January 2016 to December 2016.

Fitting the Poisson and COM-Poisson distribution to the number of fatal crashes, the parameters are obtained as follows.

Distribution	Parameters
Poisson	$\lambda = 9.3005$
COM-Poisson	$\lambda = 5.1161$ $\nu = 0.7385$

Table 6.2



The above figure gives the curves for the observed frequency and expected frequencies using Poisson & COM-Poisson distributions for the number of fatal crashes.

Date	Jan		Feb		Mar		Apr		May		June		July		Aug		Sep		Oct		Nov		Dec	
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31	1081096711111285127767836511111176755121415	12810106711131485137979837612121176755131415	575991312699101417129891216102165310121086	57511914126131012142112108915201123654101410107	61161210136581061195788121171065139171388109	911715131465912611105989121191165131119171081011	8911138969811610841510671137121376658918	812121511979111271294151088143814159885910118	1354871649131010981313911355141715847121081313134	135411716141813101181421911355142216115712111013154	257898101461113161341313591371778712164673	259911910158111418154131461217718798121947113	811111099413131344411156129471488149108810812	811121011941314164541115712115918981101098101112	111210112161211101012131799271461011532412779104	1112101131712121113131319111027146141153251410891104	6141217686910137488149121391114456496813	6151513888691015868101510121310111445849781 6	10166851110101213410611178121168141697615139161711	141978511111212174106141881311681618101971516122319	101010813141067711812115121211131391311131014101213	111015813171167711812115121411151691411131016101416	615813118687999773813133111261112574114414	617131412961079910783815133121561117574125416

																				13	10	10		

Table 6.1

Model	log-likelihood	AIC	BIC	$\chi.^2-$DF	$\chi.^2-$ Statistic	P-Value	Decision
Poisson	-982.5257	1967.1	1971	13	81.8673	0.0000	Reject
COM-Poisson	-974.1474	1952.3	1960.1	12	15.9479	0.1936	Accept

Table 6.3. Goodness of fit

From the above table, it is clear that COM-Poisson has the maximum likelihood and minimum AIC and BIC values, and at all the levels of significance, COM-Poisson distribution is accepted and Poisson distribution is rejected. This shows that COM-Poisson fits better than Poisson distribution as for as the data given in the table 6.1 is concerned.

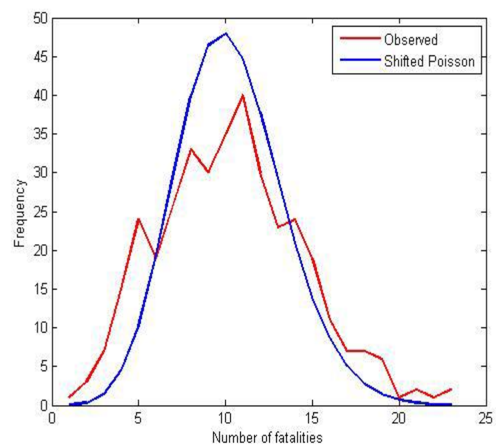


Figure 2

Fitting the shifted Poisson distribution to the number of fatalities, the parameter is obtained as $\phi = 9.3005$. The total number of fatalities follows COM-Poisson Thomas distribution. The three parameters are estimated as

$$\lambda = 5.1161$$

$$\nu = 0.7385$$

$$\phi = 9.3087$$

The mean, variance, coefficient of skewness and coefficient of kurtosis of total number of fatalities are

$$\text{Mean} = 95.8845$$
$$\text{Variance} = 1.3979 \times 10^3$$
$$\beta_1 = 0.1768$$
$$\beta_2 = 3.1855$$

Figure 2 gives the observed frequency curve and expected frequency curve using shifted Poisson distribution for the number of fatalities. Figure 3 gives the plot for probability density functions of COM-Poisson Thomas and Thomas distributions.

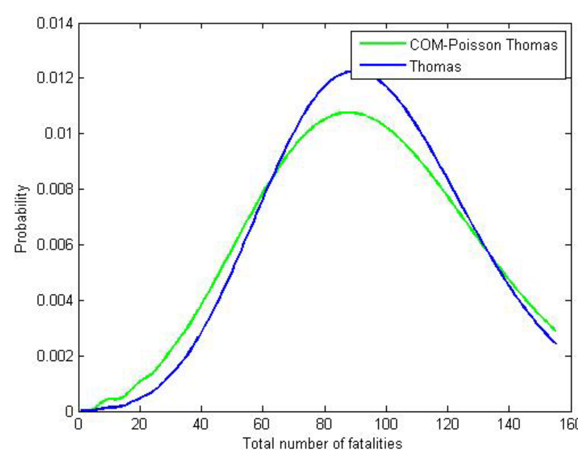


Figure. 3

7. CONCLUSION

In this paper, COM-Poisson Thomas distribution is defined and its properties are derived. Fatal accidents and fatalities data is analyzed and it is proved that COM-Poisson Thomas distribution is a better than Thomas distribution. The probability curves for COM-Poisson Thomas and Thomas distributions are plotted.

REFERENCES

[1] Conway R.W, and W.L. Maxwell. "A queuing model with state dependent service rates", Journal of Industrial Engineering, 12, pp.132-136, 1962.
[2] Douglas J.B, "Analysis with Standard Contagious Distributions", Burtonsville, MD: International Co-operative Publishing House. 1980.
[3] Feller W., "On a general class of "contagious" distribution", Annals of Mathematical Statistics, 14, pp. 389-400, 1943.
[4] Johnson N.L, Kotz S and Kemp A.W "Univariate Discrete Distributions", 3rd edition, Wiley Series in Probability and Mathematical Science, 2005.
[5] Neyman J. " On a new class of "contagious" distributions, applicable in entomology and bacteriology". The Annals of Mathematical Statistics 10, 1, pp. 35-57, 1939.
[6] Ord J.K "Families of Frequency Distributions", London: Griffin, 1972.
[7] Shmeli G, Minka T.P, Kadane J.B, Borle S and Boatwright P, "A useful distribution for fitting discrete data: Revival of the COM-Poisson distribution", J.R.Stat.Soc.Ser.C(Appl. Stat), 54, pp. 127-142, 2005.
[8] Thomas M.,(1949): A Generalization of Poisson's binomial limit for use in ecology, Biometrika, 36, pp. 18-25.