International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-05, Issue-02, May 2019

Homogeneous COM-Poisson Pascal Process

G. Shyamala, Department of Mathematics, Seethalakshmi Ramaswami College, Trichy, Tamil Nadu, ratnakrish2003@gmail.com

J. Priyadharshini, Department of Mathematics, Nehru Memorial College, Trichy, Tamil Nadu, priyajaya895@gmail.com

V. Saavithri, Department of Mathematics, Nehru Memorial College, Trichy, Tamil Nadu, saavithriramani@gmail.com

Abstract: COM-Poisson process is a generalization of Poisson process. In this paper, COM-Poisson Pascal process, which is a compound COM-Poisson process with negative binomial distribution as a compounding distribution, is defined and its properties are studied. The traffic accidents and fatalities data are analyzed to find the total number of fatalities during various time interval.

Keywords — Poisson process, Pascal (Negative Binomial) distribution, COM-Poisson process, COM-Poisson Polya-Aeppli process, COM- Poisson Pascal process.

I. INTRODUCTION

Poisson process plays a major role in analyzing counting data. As its mean and variance are equal, it is most suitable for equi-dispersed data. When the equality of mean and variance fails, the data becomes either under dispersed or over dispersed. If the data is over dispersed, the interpretation and decision using Poisson distribution may lead to wrong conclusion.

Lenden & Mantyniemi [6] used the negative binomial distribution to model over dispersion in ecological count data. Minkova[8] described heterogeneous insurance data type using Polya-Aeppli distribution. This is a generalization of count model by adding a new parameter to number of counts into the negative binomial distribution. It is a compound distribution.

Let N be a counting variable. Let $\{X_n, n = 1,2,\ldots\}$ be a sequence of independent and identically distributed positive random variables independent of N with common distribution. The distribution of the random variable $S \equiv X_1 + X_2 + \ldots + X_N$ with $S = 0$ if $N = 0$ is called a compound distribution. The distribution of N is called primary distribution and the distribution of X_n is called secondary distribution or compounding distribution. If N follows Poisson distribution, the distribution is called compound Poisson distribution.

The probability mass function of Polya-Aeppli distribution[3] is

$$P(N = 0) = e^{-\lambda}$$

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

$$n = 1,2,\ldots$$

In 2007, Meintanis [7] fitted certain bivariate distributions to traffic accidents data.

In 1952, the Poisson - Pascal distribution was introduced in the context of the spatial distribution of plants by Skellam[15], who called it a generalized Polya-Aeppli distribution.

The probability mass function of Poisson - Pascal distribution[3] is

$$P(N = 0) = e^{-\lambda}(Q^{-k} - 1)$$

$$P(N = n) = \frac{e^{-\lambda}}{n!} \left[\frac{P}{Q}\right]^n \sum_{j=1}^{\infty} \frac{(kj + n - 1)!}{(kj - 1)!\, j!} (\lambda\, Q^{-k})^j$$

$$n = 1,2,\ldots$$

In 2004, Minkova[8] defines the compound Poisson process with geometric compounding distribution and applied this process in risk models and ruin problems. The corresponding counting process is called a Polya - Aeppli process.

The probability mass function of homogeneous Polya-Aeppli process[8] is

$$P(N_t = 0) = e^{-\lambda t}$$

$$P(N_t = n) = e^{-\lambda t} \sum_{j=1}^{n} \frac{1}{j!} \binom{n-1}{j-1} \left(\lambda\, t\, (1-\rho)\right)^j \rho^{n-j}$$

$$n = 1,2,\ldots$$

The probability mass function of homogeneous Poisson - Pascal process is

 DOI : 10.35291/2454-9150.2019.0052

$$P(N_t = 0) = e^{-\lambda t}(Q^{-k} - 1)$$

$$P(N_t = n) = \frac{e^{-\lambda t}}{n!}\left[\frac{P}{Q}\right]^n \sum_{j=1}^{\infty} \frac{(kj+n-1)!}{(kj-1)!\,j!}(\lambda t\, Q^{-k})^j$$

$$n = 1,2,\ldots$$

COM-Poisson Pascal Distribution [9] which is a combination of COM-Poisson and negative binomial distributions.

The probability mass function of N is

$$P(N=0) = \frac{Z(\lambda Q^{-k},\ \nu)}{Z(\lambda,\nu)}$$

$$P(N=n) = \frac{1}{Z(\lambda,\nu)}\left(\frac{P}{Q}\right)^n \sum_{j=1}^{\infty} \frac{(kj+n-1)!}{(kj-1)!n!}\frac{(\lambda Q^{-k})^j}{(j!)^\nu}$$

The probability density function of COM-Poisson process [10] is

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

In this paper, COM-Poisson Pascal process is defined and its properties are studied. The traffic accidents and fatalities data are analyzed using COM-Poisson Pascal process.

This paper is organized as follows: In section 2, COM-Poisson Pascal process is defined and some of its properties are studied. In section 3, traffic accidents and fatalities data are analyzed. Section 4 concludes this paper.

II. COM-POISSON PASCAL PROCESS

2.1 Probability density function

Assume that there are Y_t independent random variables in the interval $[0,t]$of the form X, and N_t denotes the sum of these random variables.

(ie) $N_t = X_1 + X_2 + \ldots + X_{Y_t}$

COM-Poisson Pascal process is derived by assuming that

(i) X denotes the number of objects within a cluster and X follows Negative binomial (Pascal) distribution with parameters k and P.

(ie) $X \sim NB(k,P)$

(ii) Y_t denotes the number of clusters in the interval $[0,t]$ and Y_t follows COM-Poisson process with parameters λt and ν

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

This random variable, N_t formed by compounding these two random variables X and Y_t gives the COM-Poisson Pascal process with parameters $\lambda t, \nu, k$ and P.

1. The probability generating function of X is,
$$G_X(s) = (Q-Ps)^{-k} \quad \ldots\ldots (2.1.1)$$
where $Q = 1 + P$.
2. The probability generating function of COM-Poisson process is
$$G_{Y_t}(s) = \frac{Z(\lambda ts,\nu)}{Z(\lambda t,\nu)} \quad \ldots\ldots (2.1.2)$$

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

$$G_{N_t}(s) = E(s^{N_t}) = E\left(s^{X_1+X_2+\ldots+X_{Y_t}}\right)$$
$$= \sum_{y=0}^{\infty} E\left(s^{X_1+X_2+\ldots+X_{Y_t}} / Y(t)=y\right) P(Y(t)=y)$$
$$= \sum_{y=0}^{\infty} [E(s^x)]^y P(Y(t)=y)$$
$$= G_Y(G_X(s))$$
$$= \frac{Z(\lambda t G_X(s),\nu)}{Z(\lambda t,\nu)}$$
$$= \frac{1}{Z(\lambda t,\nu)}\sum_{j=0}^{\infty}\frac{1}{(j!)^\nu}(\lambda t(Q-Ps)^{-k})^j$$
$$= \frac{1}{Z(\lambda t,\nu)}\sum_{j=0}^{\infty}\frac{(\lambda t)^j}{(j!)^\nu}(Q-Ps)^{-kj}$$

Collecting the coefficient of s^n in the above series we get

$$P(N_t=n) = \frac{1}{Z(\lambda t,\nu)}\left(\frac{P}{Q}\right)^n \sum_{j=1}^{\infty}\frac{(kj+n-1)!}{(kj-1)!n!}\frac{(\lambda t Q^{-k})^j}{(j!)^\nu}$$

The probability mass function of N is

$$P(N_t=0) = \frac{Z(\lambda t Q^{-k},\ \nu)}{Z(\lambda t,\nu)}$$

$$P(N_t=n) = \frac{1}{Z(\lambda t,\nu)}\left(\frac{P}{Q}\right)^n \sum_{j=1}^{\infty}\frac{(kj+n-1)!}{(kj-1)!n!}\frac{(\lambda t Q^{-k})^j}{(j!)^\nu} \quad \ldots. (2.1.3)$$

where $\lambda > 0, \nu \geq 0$, $k > 1$ and $P \geq 0$.

2.2 Special cases

1. Replacing k by 1 , P by $\frac{p}{1-p}$, λ by $\frac{\lambda_1}{p}$ in equation (2.1.3), the PGF of COM-Poisson Polya-Aeppli process is obtained as
$$\frac{z\left(\frac{\lambda ts(1-p)}{(1-ps)},\nu\right)}{Z(\lambda t,\nu)}$$
2. When $\nu = 1$ from equation (2.1.3), we get Generalized Polya-Aeppli (Poisson-Geometric) process with PGF
$$exp\{\lambda t[(Q-Ps)^{-k} - 1]\}$$
3. Replacing ν by 1 k by 1 , P by $\frac{p}{1-p}$, λ by $\frac{\lambda_1}{p}$ in equation (2.1.3), the PGF of Polya-Aeppli process

 DOI : 10.35291/2454-9150.2019.0052

$$\exp\left\{\frac{\lambda ts(1-p)}{(1-ps)}-1\right\}$$

2.3 Limiting cases of the COM-Poisson Pascal Process

1. When $k \to \infty, P \to 0, Pk = \mu$, from equation (2.1.3), we get COM-Poisson Neyman Type A process with PGF
$$\frac{Z\big(\lambda t(\exp(\mu(s-1))),v}{Z(\lambda t,v)}$$
2. When $v = 1, k \to \infty, P \to 0, Pk = \mu$, from equation (2.1.3), we get Neyman Type A process with PGF
$$exp[\lambda t(exp(\mu(s-1))-1)]$$

2.4 Mean and variance of COM-Poisson Pascal Process

The mean and variance of the distribution are given by

$$Mean(N_t) = \frac{\lambda\, kP\, Z_\lambda(\lambda t,v)}{Z(\lambda t,v)}$$

$$Var(N_t) = \frac{1}{Z(\lambda t\,,v)}\left[\lambda^2k^2P^2\left(Z_{\lambda\lambda}(\lambda t,v) - \frac{[Z_\lambda(\lambda t,v)]^2}{Z(\lambda t,v)}\right)\right] + \frac{1}{Z(\lambda t,v)}[\lambda tkp(Q+kP)\lambda Z_\lambda(\lambda t,v)]$$

III. DATA ANALYSIS

In this section, two sets of traffic accidents and fatalities data are analyzed

3.1 Data set 1

The following table gives the total Sunday accidents (left entry) and the corresponding number of fatalities (right entry) recorded in the Groningen region for each month during the years 1997-2004 [6]

Month	1997		1998		1999		2000		2001		2001		2003		2004	
January	6	0	6	0	13	1	11	0	8	0	8	0	11	4	2	0
February	10	0	10	1	7	0	4	0	8	1	8	0	9	0	2	0
March	7	0	13	4	8	0	10	0	6	0	12	0	9	0	3	0
April	11	0	5	0	14	1	15	1	9	0	10	1	7	1	1	1
May	12	0	17	2	13	0	18	0	13	2	11	0	12	1	5	0
June	21	1	19	0	14	0	21	1	12	3	12	1	13	0	7	2
July	15	0	10	0	14	0	11	1	10	2	4	0	8	0	1	0
August	11	1	11	1	10	0	8	0	9	0	14	1	6	0	5	0
September	7	0	11	0	7	0	9	0	22	1	16	1	7	0	8	1
October	11	2	13	1	16	1	14	0	15	1	8	1	6	1	2	0
November	15	1	17	1	13	0	13	0	6	0	9	1	11	1	1	0
December	5	0	7	0	10	1	11	0	10	0	8	0	5	0	2	0

Let Y be the number of Sunday's that accidents occur in Groningen between the years 1997-2004.

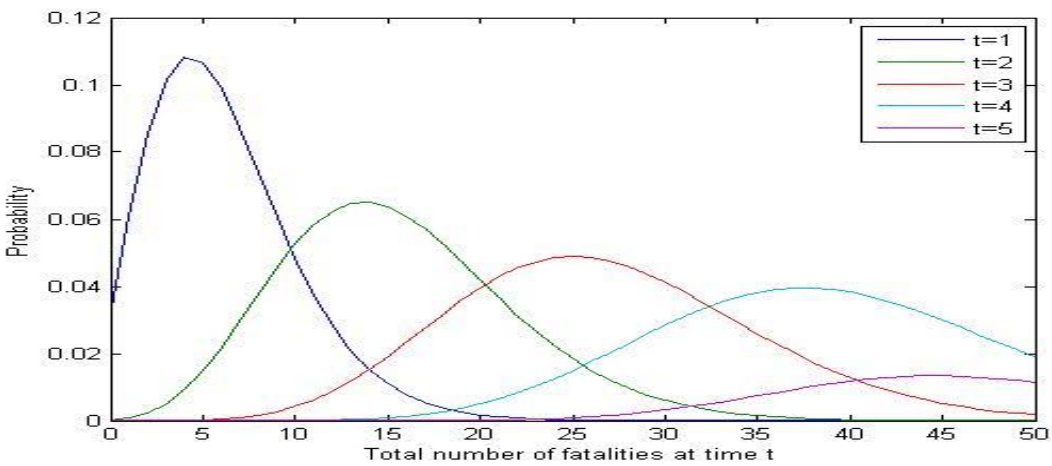
$X_i, i = 1,2,\ldots$, be the number of fatalities of ith accident and N be the total number of fatalities from the year 1997 to 2004.

Assuming that the number of accidents follows COM-Poisson distribution, the estimated values of λ and v are 2.6027 & 0.4320 respectively.

Fitting the Negative binomial distribution to the number of fatalities, the parameters are obtained as $k = 2, P = 0.2300$

Therefore total number of fatalities at time t follows COM-Poisson Pascal process follows with parameters $\lambda = 2.6027\,, v = 0.4320, k = 2$ and $P = 0.23$.

The following figure shows the probabilities of total number of fatalities at time t for various t.



 DOI : 10.35291/2454-9150.2019.0052

3.2 Data set 2

Here the data is taken from fatal crashes and fatalities calendar 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 years 2016.

Date	Jan		Feb		Mar		Apr		May		June		July		Aug		Sep		Oct		Nov		Dec	
1	10	12	5	5	6	9	8	8	13	13	2	2	8	8	11	11	6	6	10	14	10	11	6	6
2	8	8	7	7	11	11	9	12	5	5	5	5	11	11	12	12	14	15	16	19	10	10	15	17
3	10	10	5	5	6	7	11	12	4	4	7	9	11	12	10	10	12	15	6	7	10	15	8	13
4	9	10	9	11	12	15	13	15	8	11	8	9	10	10	1	1	1	13	8	8	8	8	13	14
5	6	6	9	9	10	13	8	11	7	7	9	11	9	11	12	13	7	8	5	5	13	13	11	12
6	7	7	13	14	13	14	9	9	16	16	8	9	9	9	16	17	6	8	11	11	14	17	8	9
7	11	11	12	12	6	6	6	7	9	14	10	10	4	4	12	12	8	8	10	11	10	11	6	6
8	11	13	6	6	5	5	9	9	13	18	14	15	13	13	11	12	6	6	10	12	6	6	8	10
9	12	14	9	13	8	9	8	11	10	13	6	8	13	14	10	11	9	9	12	12	7	7	7	7
10	8	8	9	10	10	12	11	12	10	10	11	11	13	16	10	13	10	10	13	17	7	7	9	9
11	5	5	10	12	6	6	6	7	9	11	13	14	4	4	12	13	13	15	4	4	11	11	9	9
12	12	13	14	14	11	11	10	12	8	8	16	18	4	5	13	13	7	8	10	10	8	8	9	10
13	7	7	17	21	9	10	8	9	13	14	13	15	4	4	17	19	4	6	6	6	12	12	7	7
14	7	9	12	12	5	5	4	4	13	21	4	4	11	11	9	11	8	8	11	14	11	11	7	8
15	6	7	9	10	7	9	15	15	9	9	13	13	15	15	9	10	8	10	17	18	5	5	3	3
16	7	9	8	8	8	8	10	10	11	11	13	14	6	7	2	2	14	15	8	8	12	12	8	8
17	8	8	9	9	8	9	6	8	3	3	5	6	12	12	7	7	9	10	12	13	12	14	13	15
18	3	3	12	15	12	12	7	8	5	5	9	12	9	11	14	14	12	12	11	11	11	11	13	13
19	6	7	16	20	11	11	11	14	5	5	13	17	4	5	6	6	13	13	6	6	13	15	3	3
20	5	6	10	11	7	9	3	3	14	14	7	7	7	9	10	14	9	10	8	8	13	16	11	12
21	11	12	21	23	10	11	7	8	17	22	17	18	14	18	11	11	11	11	14	16	9	9	12	15
22	11	12	6	6	6	6	12	14	15	16	7	7	8	9	5	5	14	14	16	18	13	14	6	6
23	11	11	5	5	5	5	13	15	8	11	8	9	8	8	3	3	4	4	9	10	11	11	11	11
24	7	7	3	4	13	13	7	9	4	5	7	8	14	1	2	2	5	5	17	19	13	13	12	17
25	6	6	10	10	9	11	6	8	7	7	12	12	9	10	4	5	6	8	6	7	10	10	5	5
26	7	7	12	14	17	19	6	8	12	12	16	19	10	10	12	14	4	4	15	15	14	16	7	7
27	5	5	10	10	13	17	5	5	10	11	4	4	8	9	7	10	9	9	13	16	10	10	4	4
28	5	5	8	10	8	10	8	9	8	10	6	7	8	8	7	8	6	7	9	12	12	14	11	12
29	12	13	6	7	8	8	9	10	13	13	7	11	10	10	9	9	8	8	16	23	13	16	4	5
30	14	14			10	10	18	18	13	15	3	3	8	11	10	10	13	16	17	19	10	10	4	4
31	15	15			9	11			4	4			12	12	4	4			11	13			14	16

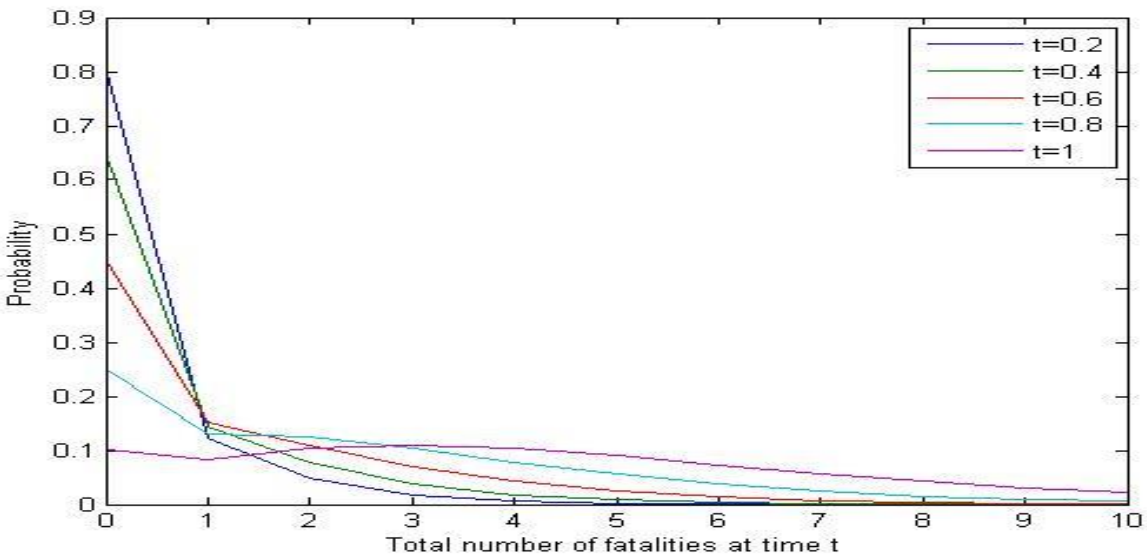
Let Y be the number of day's that accidents occurred at the year 2016. X_i, $i = 1,2,\ldots,$ be the number of fatalities of ith accident and N be the total number of fatalities from January 2016 to December 2016.

Assuming that the number of accidents follows COM-Poisson distribution, the estimated values of λ and ν are 5.1161 & 0.7385 respectively.

Fitting the Negative binomial distribution to the number of fatalities, the parameters are obtained as $k = 15$, $P = 0.6622$.

Therefore total number of fatalities at time t follows COM-Poisson Pascal process follows with parameters $\lambda = 5.1161$, $\nu = 0.7385$, $k = 15$ and $P = 0.6622$.

The following figure shows the probabilities of total number of fatalities at time t for various t.



As both the data sets are over dispersed, COM-Poisson Pascal process will be more suitable than the Poisson Pascal process.

For the above data sets 1& 2 , when time interval with length t increases the probable range of total number of fatalities increases. Hence more number of preventive steps have to be taken to save the lives of people by identifying the actual causes of accidents.

IV. CONCLUSION

In this paper, COM-Poisson Pascal process is defined and its properties are studied. The traffic accidents and fatalities data are analyzed using this process. It is shown

 DOI : 10.35291/2454-9150.2019.0052

that COM-Poisson Pascal process is better than compound Poisson process as for as these data are concerned.

REFERENCES

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

[2] Feller W, "On a general class of "contagious" distributions", Annals of Mathematical Statistics, 1943). 14, 389-400.

[3] Johnson N.L, Kotz S and Kemp A.W, "Univariate Discrete Distributions", 3rd edition, Wiley Series in Probability and Mathematical Sciences. 2005.

[4] Josemar Rodrigues, Mario De Castro, Vicente G. Cancho and N.Balakrishnan, "COM-Poisson cure rate survival models and an application to a cutaneous melanoma data", Journal of Statistical Planning and Inference, 2009, 139, 3605-3611

[5] Katti S.K., and Gurland J, "The Poisson Pascal distribution", Biometrics, 1961, 17, 527-538.

[6] Linden, A. and Mantyniemi, S. (2011). Using the negative binomial distribution to model overdispersion in ecological count data. Ecology, 92, 1414-1421.

[7] Meintanis S.G, " A new goodness of fit test for certain bivariate distributions applicable to traffic accidents", Statist. Methodol. 2007, 4, pp. 22-34

[8] Minkova L.D, "A Generalization of the Classical Discrete Distributions", Commun. Statist - Theory and Methods, 2002, 31 (6), 871-888.

[9] Polya G, "Sur quelques points de la th'eorie des probabilities", Annales de l'Institut H. Poincre, 1930 , 1, 117-161.

[10] Priyadharshini J and Saavithri V , "COM-Poisson Pascal distribution", International Journal for Research in Engineering Application & Management (IJREAM), Vol-04, Issue-08, Nov 2018, 146-151.

[11] Priyadharshini J and Saavithri V (2018): A study on COM-Poisson Process and Its Applications, Journal of Applied Science and Computations, Vol 5, Issue 10, 2018, 487-497.

[12] 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), 2005, 54, 127-142.

[13] Shumway R., and Gurland J, "A fitting procedure for some generalized Poisson distributions", Skandivanavisk Aktuarietidskrift, 1960a, 43, 87-108.

[14] Shumway R., And Gurland J, "Fitting the Poisson binomial distributions", Biometrics, 1960b, 16, 522-533.

[15] Skellam J.G, "Studies in statistical ecology I: Spatial pattern", Biometrika, 1952, 39, 346-362.

 DOI : 10.35291/2454-9150.2019.0052