STM JOURNALS

Research & Reviews: Journal of
Statistics

https://journals.stmjournals.com/rrjost/

ISSN: 2278-2273 (Online)
ISSN: 2348-7909 (Print)
Volume 13, Issue 1, 2024
January–April
DOI (Journal): 10.37591/RRJoST

Research RRJoST

The Poisson–Uma Distribution with Properties and Applications to Model Thunderstorm Events

Rama Shanker[1,]*, Kamlesh Kumar Shukla[2]

Abstract

The discrete data available in any field of knowledge is influenced by several known and unknown factors and the factors which affect the discrete data are stochastic. The stochastic nature of discrete data is a challenge for statisticians to model and analyze with the existing discrete distributions. In the present paper, Poisson-Uma distribution, the Poisson compound of Uma distribution, has been proposed to model over-dispersed data of thunderstorm events. The descriptive statistical constants based on moments of the proposed distribution have been studied. Over-dispersion, increasing hazard rate, and unimodality of the distribution have been discussed. One of the important characteristics of the Poisson-Uma distribution is that although it is also a mixture of geometric and negative binomial distributions its nature does not exhibit multiple modes. The method of moment and maximum likelihood estimation for estimating parameters has been studied. The applications of the distribution to model thunderstorm events for June, July, August, and summer have been discussed. The proposed distribution shows a much better fit than the Poisson-Lindley distribution and the Poisson-Sujatha distribution. Therefore, Poisson-Uma distribution can be considered an important over-dispersed distribution for modeling over-dispersed data of thunderstorm events.

Keywords: Uma distribution, compounding, statistical properties, maximum likelihood estimation, goodness of fit

INTRODUCTION

The search for a suitable distribution to model thunderstorm events is very challenging because thunderstorm events are affected by high winds, lighting, extreme turbulence, and combinations of environmental conditions, including unstable air with high moisture content and some type of lifting action during summer months. Initially, Falls et al. (1971) [1] and Carter et al. (2001) [2] discussed in detail the modeling of thunderstorm events and proposed a compound negative binomial-positive binomial distribution. However, because of the complexity of the probability mass function (pmf) of the compound negative binomial-positive binomial distribution and difficulty in the estimation of parameters, the distribution has not attracted the attention of researchers. The search for a suitable and compact distribution that can provide a good fit for thunderstorm events is a challenging task. Because the data relating to thunderstorm events are discrete and over-dispersed, an over-dispersed discrete distribution is required.

We know that the Poisson distribution is an appropriate distribution for equi-dispersed (mean equal to variance) data. However, in real-life situations, stochastic datasets are either over-dispersed (variance greater than mean) or under-dispersed (variance less than mean). Recently, researchers have attempted to derive an over-dispersed one-parameter discrete distribution by

***Author for Correspondence**
Rama Shanker
E-mail: shankerrama2009@gmail.com

[1]Associate Professor, Department of Statistics, Assam University, Silchar, Assam, India
[2]Assistant Professor, Department of Mathematics, Jaypee Institute of Information Technology, Noida, India

Received Date: July 05, 2024
Accepted Date: July 24, 2024
Published Date: September 5, 2024

Citation: Rama Shanker, Kamlesh Kumar Shukla. The Poisson-Uma Distribution with Properties and Applications to Model Thunderstorm Events. Research & Reviews: Journal of Statistics. 2024; 13(1): 20–30p.

compounding the Poisson distribution with one-parameter continuous lifetime distributions. A popular one-parameter over-dispersed discrete distribution is the Poisson-Lindley distribution (PLD) suggested by Sankaran (1970) [3]. The PLD is a Poisson compound of the Lindley distribution (LD) proposed by Lindley (1958) [4]. LD is defined by its probability density function (pdf) and cumulative distribution function (cdf):

$$f_1(x;\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}; x > 0, \theta > 0 \tag{1}$$

$$F_1(x,\theta) = 1 - \left(1 + \frac{\theta x}{\theta+1}\right)e^{-\theta x}; x > 0, \theta > 0 \tag{2}$$

Ghitany et al. (2008) [5] conducted a detailed study on the characterization of the Lindley distribution and showed that the Lindley distribution provides a much better fit than the exponential distribution for the data relating to the waiting time of a customer at a bank because the hazard rate of the Lindley distribution is not constant, such as the hazard rate of the exponential distribution. The PLD is defined by its pmf:

$$P(x,\theta) = \frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}}; x = 0,1,2,\ldots, \theta > 0. \tag{3}$$

Various descriptive properties and estimation methods of PLD have been discussed by Sankaran (1970) and Ghitany and Al-Mutairi (2009) [6], and it has been shown that PLD has an over-dispersed distribution. While testing the goodness of fit of PLD for thunderstorm events, it was observed that although PLD has an over-dispersed discrete distribution, it does not provide a good fit. Shanker and Hagos (2015) [7] extensively studied the applications of PLD for overdispersed data. In recent decades, several lifetime distributions have been proposed, and the Poisson compound of these lifetime distributions provides overdispersed discrete distributions. The Sujatha distribution (SD) is the one-lifetime distribution proposed by Shanker (2016) [8], which provides a better fit than the exponential and Lindley distributions. The SD is defined by its pdf and cdf.

$$f_2(x;\theta) = \frac{\theta^3}{\theta^2+\theta+2}(1+x+x^2)e^{-\theta x}; x > 0, \theta > 0 \tag{4}$$

$$F_2(x;\theta) = 1 - \left[1 + \frac{\theta x(\theta x+\theta+2)}{\theta^2+\theta+2}\right]e^{-\theta x}; x > 0, \theta > 0. \tag{5}$$

Shanker (2016) [9] also derived the Poisson compound of the Sujatha distribution and named it the Poisson-Sujatha distribution (PSD), which is also an over-dispersed discrete distribution. The PSD is defined by its pmf:

$$P(x;\theta) = \frac{\theta^3}{\theta^2+\theta+2} \cdot \frac{x^2+(\theta+4)x+\left(\theta^2+3\theta+4\right)}{(\theta+1)^{x+3}}; x = 0,1,2,\ldots, \theta > 0 \tag{6}$$

Similar to PLD, PSD is also an over-dispersed discrete distribution. Shanker and Hagos (2016) [10] conducted extensive studies on the applications of PSD for modeling over-dispersed discrete data from biological sciences and found that PSD provides a much better fit than PLD. Furthermore, it has been observed that PSD does not provide a satisfactory fit for the data relating to thunderstorm events, and this is an indication that the search for an over-dispersed discrete distribution should continue, which should provide a better fit for thunderstorm events than PLD.

Shanker (2022) [11] proposed a one-parameter lifetime distribution called the Uma distribution (UD), defined by its pdf and cdf.

$$f_3(x;\theta) = \frac{\theta^4}{\theta^3+\theta^2+6}(1+x+x^3)e^{-\theta x}; x > 0, \theta > 0 \tag{7}$$

$$F_3(x;\theta) = 1 - \left[1 + \frac{\theta x(\theta^2 x^2 + 3\theta x + \theta^2 + 6)}{\theta^3+\theta^2+6}\right] e^{-\theta x}; x > 0, \theta > 0 \tag{8}$$

Detailed studies of Uma distribution are available in Shanker (2022). Shanker et al. (2023, 2023) [12, 13] also proposed the weighted Uma distribution and power Uma distribution and discussed their statistical properties and applications in different fields of knowledge.

The main purpose of this study was to derive an over-dispersed discrete distribution, which is a compound of Poisson and Uma distributions. The derivation of the compound of the Poisson distribution with the Uma distribution lies in the fact that the Lindley distribution provides a better fit than the exponential distribution, and the Sujatha distribution provides a better fit than the Lindley distribution, and it is expected that their corresponding Poisson compound will provide a better fit. Descriptive statistical constants based on moments were studied. Overdispersion, unimodality, and increasing hazard rate of the derived distribution have been discussed. The estimation of the parameters of the proposed distribution is discussed using the method of moment and maximum likelihood. Applications and a good fit of the proposed distribution were also presented.

POISSON-UMA DISTRIBUTION

Let Xfollows Poisson distribution with parameter $\lambda > 0$ having pmf

$$P(X|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}; x = 0,1,2,\ldots$$

Now suppose the parameter λfollows Uma distribution with parameter θ having pdf

$$f(\lambda|\theta) = \frac{\theta^4}{\theta^3+\theta^2+6}(1+\lambda+\lambda^3)e^{-\theta\lambda}; \lambda > 0, \theta > 0$$

Thus, the marginal pmf of X can be obtained as

$$P(X=x) = \int_0^\infty P(X|\lambda) f(\lambda|\theta) d\lambda = \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!} \frac{\theta^4}{\theta^3+\theta^2+6}(1+\lambda+\lambda^3)e^{-\theta\lambda} d\lambda \tag{9}$$
$$= \frac{\theta^4}{(\theta^3+\theta^2+6)x!}\int_0^\infty e^{-(\theta+1)\lambda}\lambda^x (1+\lambda+\lambda^3)\, d\lambda$$

$$= \frac{\theta^4}{\theta^3+\theta^2+6} \frac{x^3+6x^2+(\theta^2+2\theta+12)x+(\theta^3+4\theta^2+5\theta+8)}{(\theta+1)^{x+4}}; x = 0,1,2,\ldots, \theta > 0 \tag{10}$$

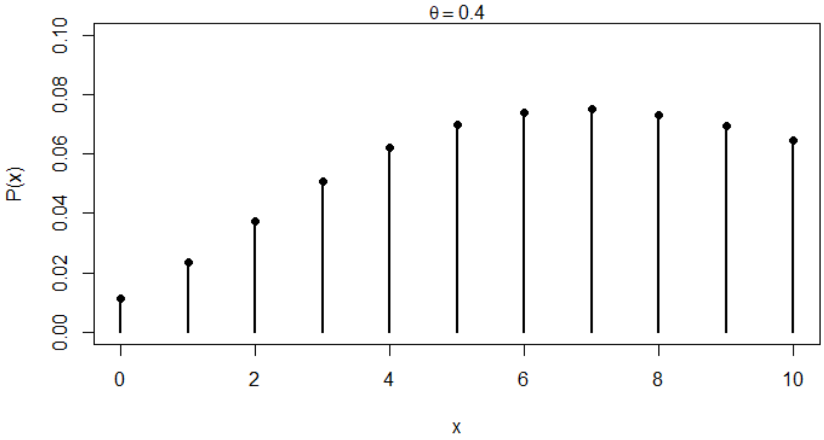
We would call this distribution as Poisson-Uma distribution (PUD). Furthermore, we would show that the pmf of PUD is over-dispersed, unimodal, and has an increasing hazard rate. The pmf of PUD for varying values of parameter has been shown in Figure 1, and it reveals that as the value of parameter increases, the distribution becomes positively skewed and over-dispersed.

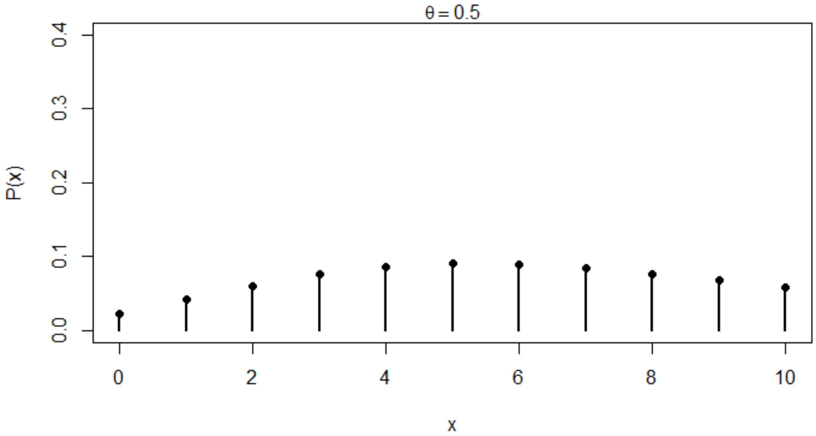
The pmf of the PUD for varying values of the parameter is shown in Figure 1.

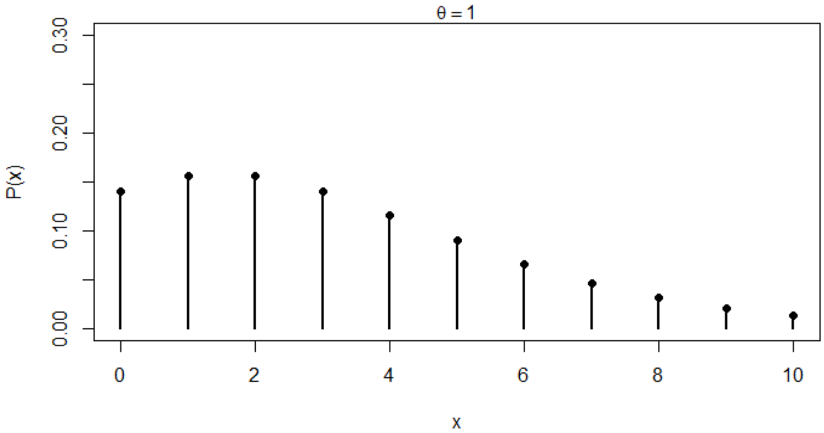
In general, a PUD is a three-component mixture distribution that can be expressed as

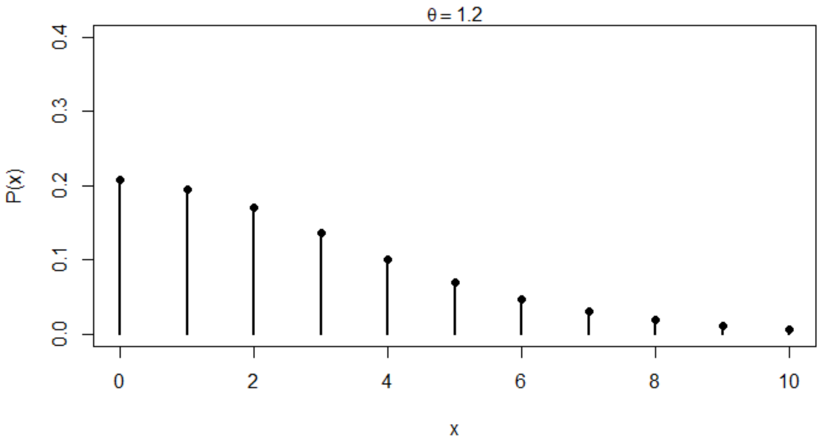
$$P(x;\theta) = p_1 P_1(x;\theta) + p_2 P_2(x;\theta) + p_3 P_3(x;\theta),$$

Where, $P_i(x;\theta)$is the pmf of the negative binomial distribution (NBD) with the number of successes i and with certain proportions. When $i = 1$, $P_i(x;\theta)$is the pmf of the geometric distribution (GD), which is a special case of NBD. The formulae for p_1, p_2, p_3 and $P_i(x;\theta)$for $i = 1,2,3$ are given by


θ = 0.4
P(x)
0.00
0.02
0.04
0.06
0.08
0.10
0
2
4
6
8
10
x


θ = 0.5
P(x)
0.0
0.1
0.2
0.3
0.4
0
2
4
6
8
10
x


θ = 1
P(x)
0.00
0.10
0.20
0.30
0
2
4
6
8
10
x


θ = 1.2
P(x)
0.0
0.1
0.2
0.3
0.4
0
2
4
6
8
10
x

Research & Reviews: Journal of Statistics
Volume 13, Issue 1
ISSN: 2278-2273 (Online), ISSN: 2348-7909 (Print)

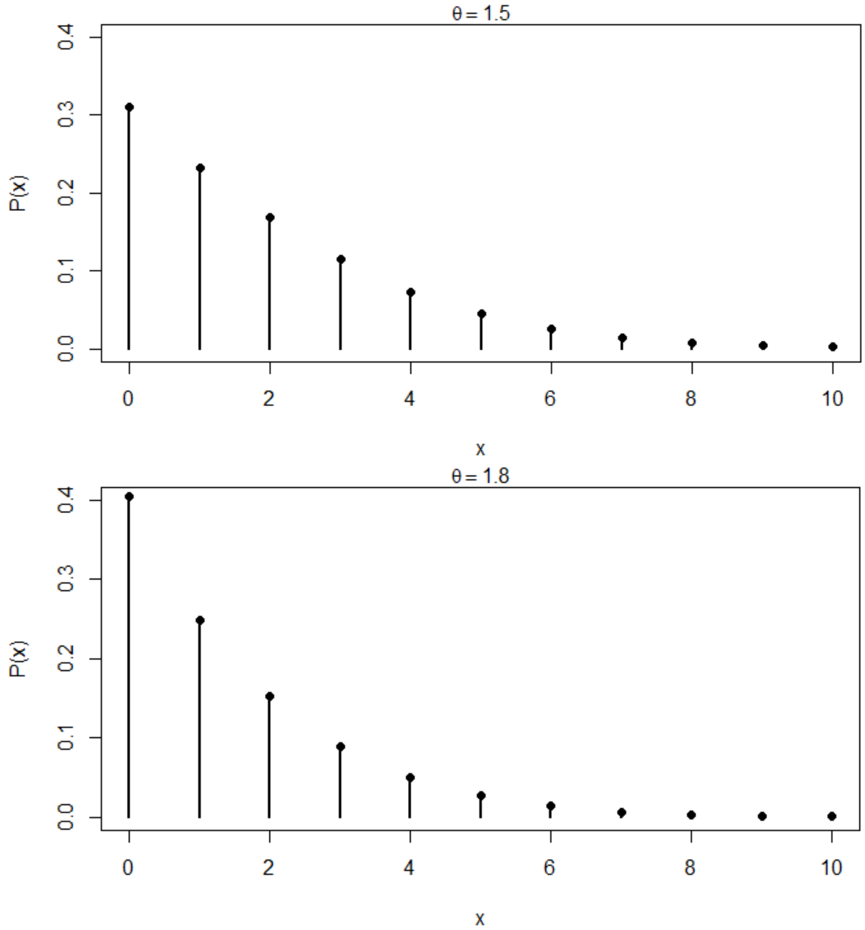


Figure1. pmf of PUD.

$$p_1 = \frac{\theta^3}{\theta^3+\theta^2+6}, p_2 = \frac{\theta^2}{\theta^3+\theta^2+6}, p_3 = \frac{6}{\theta^3+\theta^2+6}$$

$$P_1(x;\theta) = \frac{\theta}{(\theta+1)^{x+1}}, P_2(x;\theta) = \frac{(x+1)\theta^2}{(\theta+1)^{x+2}}, P_3(x;\theta) = \frac{(x+1)(x+2)(x+3)\theta^4}{6(\theta+1)^{x+4}}.$$

Even though the PUD is a three-component mixture of NBD, the presence of three modes is not noticeable in any of the plots of pmf of PUD in Figure 1 for the selected values of the parameter θ. This indicates that the three modes from the three subpopulations must be located close to each other. Tajuddin et al. (2022) [14] observed that if the modes of the sub-populations are located very close to each other, the population will have a single mode. This means that the distribution in which the existence of the modes of the sub-populations, each with very close mode values, is certain can be a model for over-dispersed data.

DESCRIPTIVE STATISTICAL CONSTANTS

Using Equation 9, the r th factorial moment about origin, $\mu_{(r)}{}'$, of the PUD can be obtained as

$$\mu_{(r)}{}' = E\left[E\left(X^{(r)}|\lambda\right)\right] = \frac{\theta^4}{\theta^3+\theta^2+6}\int_0^\infty \left[\sum_{x=0}^{\infty} x^{(r)} \frac{e^{-\lambda}\lambda^x}{x!}\right](1+\lambda+\lambda^3)e^{-\theta\lambda}d\lambda$$

$$= \frac{\theta^4}{\theta^3+\theta^2+6}\int_0^\infty \lambda^r \left[\sum_{x=r}^{\infty} \frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!}\right](1+\lambda+\lambda^3)e^{-\theta\lambda}d\lambda$$

$$= \frac{\theta^4}{\theta^3+\theta^2+6}\int_0^\infty \lambda^r\,(1+\lambda+\lambda^3)e^{-\theta\lambda}d\lambda$$

$$= \frac{r!\{\theta^3+(r+1)\theta^2+(r+1)(r+2)(r+3)\}}{\theta^r(\theta^3+\theta^2+6)}; r = 1,2,3,.. \quad (11)$$

Substituting $r = 1,2,3$ and 4 in Equation 11, the first four factorial moments about the origin of the PUD can be obtained as

$$\mu_{(1)}{}' = \frac{\theta^3+2\theta^2+24}{\theta(\theta^3+\theta^2+6)}, \mu_{(2)}{}' = \frac{2(\theta^3+3\theta^2+60)}{\theta^2(\theta^3+\theta^2+6)}$$

$$\mu_{(3)}{}' = \frac{6(\theta^3+4\theta^2+120)}{\theta^3(\theta^3+\theta^2+6)}, \mu_{(4)}{}' = \frac{24(\theta^3+5\theta^2+210)}{\theta^4(\theta^3+\theta^2+6)}.$$

The relationship between moments about the origin and factorial moments about the origin gives the following four moments about the origin:

$$\mu_1{}' = \frac{\theta^3+2\theta^2+24}{\theta(\theta^3+\theta^2+6)}$$

$$\mu_2{}' = \frac{\theta^4+4\theta^3+6\theta^2+24\theta+120}{\theta^2(\theta^3+\theta^2+6)}$$

$$\mu_3{}' = \frac{\theta^5+8\theta^4+24\theta^3+48\theta^2+360\theta+720}{\theta^3(\theta^3+\theta^2+6)}$$

$$\mu_4{}' = \frac{\theta^6+16\theta^5+78\theta^4+192\theta^3+960\theta^2+4320\theta+5040}{\theta^4(\theta^3+\theta^2+6)}.$$

Using the relationship between moments about the mean and the moments about the origin, moments about the mean are obtained as:

$$\mu_2 = \frac{\theta^7+4\theta^6+6\theta^5+32\theta^4+120\theta^3+60\theta^2+144\theta+144}{\theta^2(\theta^3+\theta^2+6)^2}$$

$$\mu_3 = \frac{\begin{pmatrix}\theta^{11}+7\theta^{10}+22\theta^9+68\theta^8+372\theta^7+952\theta^6+1080\theta^5+2520\theta^4+2160\theta^3\\+1728\theta^2+2592\theta+1728\end{pmatrix}}{\theta^3(\theta^3+\theta^2+6)^3}$$

$$\mu_4 = \frac{\begin{pmatrix}\theta^{15}+15\theta^{14}+87\theta^{13}+300\theta^{12}+1402\theta^{11}+6404\theta^{10}+15348\theta^9+29376\theta^8\\+68472\theta^7+100944\theta^6+131328\theta^5+207792\theta^4+222912\theta^3+171072\theta^2\\+186624\theta+93312\end{pmatrix}}{\theta^4(\theta^3+\theta^2+6)^4}.$$

The descriptive measures of PUD, including the coefficient of variation (C.V), skewness, kurtosis, and index of dispersion, are obtained as:

$$C.V = \frac{\sigma}{\mu_1{}'} = \frac{\sqrt{\theta^7+4\theta^6+6\theta^5+32\theta^4+120\theta^3+60\theta^2+144\theta+144}}{\theta^3+2\theta^2+24}$$

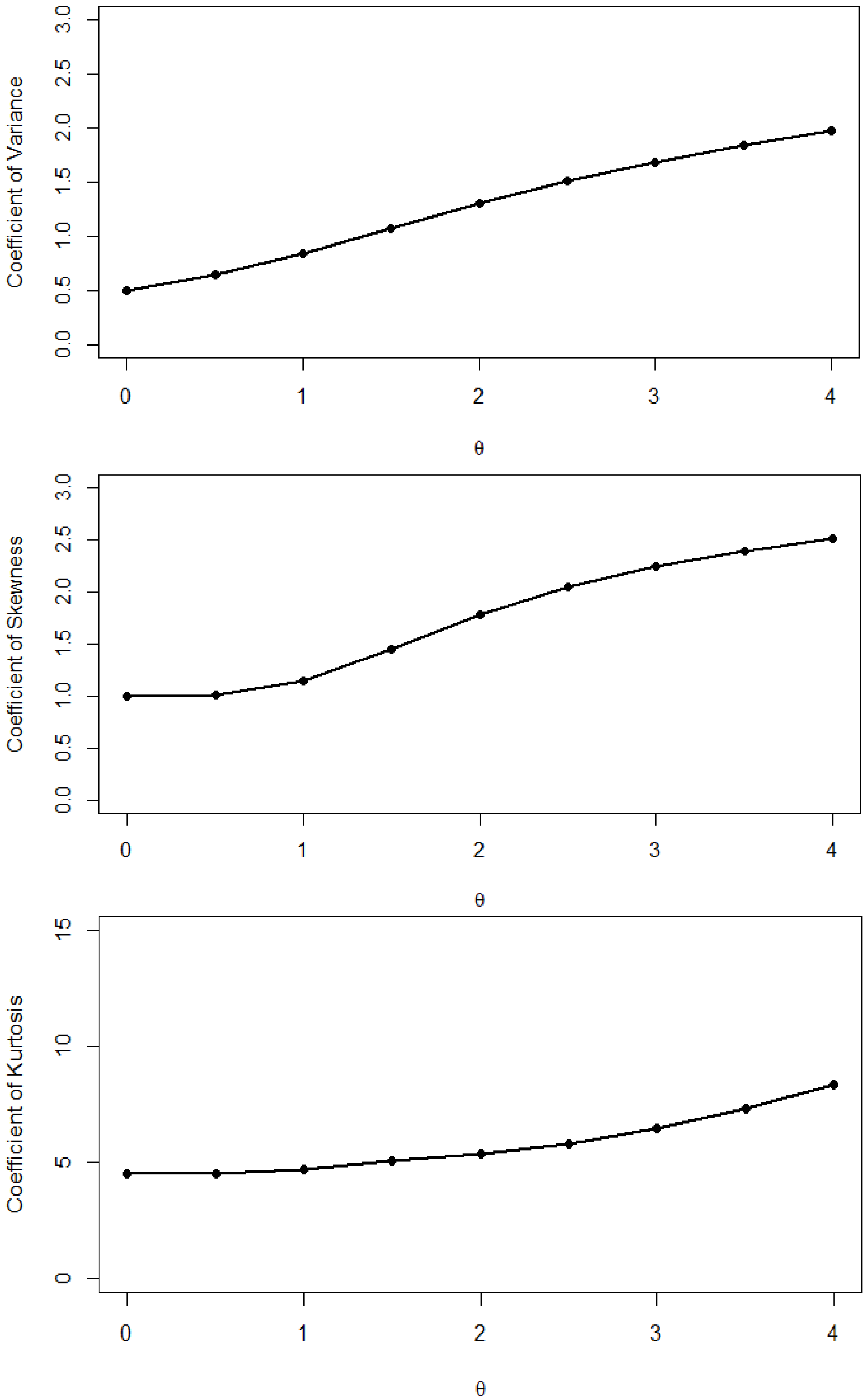
$$\sqrt{\beta_1} = \frac{\mu_3}{(\mu_2)^{\frac{3}{2}}} = \frac{\begin{pmatrix}\theta^{11}+7\theta^{10}+22\theta^9+68\theta^8+372\theta^7+952\theta^6+1080\theta^5+2520\theta^4+2160\theta^3\\+1728\theta^2+2592\theta+1728\end{pmatrix}}{(\theta^7+4\theta^6+6\theta^5+32\theta^4+120\theta^3+60\theta^2+144\theta+144)^{\frac{3}{2}}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2{}^2} = \frac{\begin{pmatrix}\theta^{15}+15\theta^{14}+87\theta^{13}+300\theta^{12}+1402\theta^{11}+6404\theta^{10}+15348\theta^9+29376\theta^8\\+68472\theta^7+100944\theta^6+131328\theta^5+207792\theta^4+222912\theta^3+171072\theta^2\\+186624\theta+93312\end{pmatrix}}{(\theta^7+4\theta^6+6\theta^5+32\theta^4+120\theta^3+60\theta^2+144\theta+144)^2}$$

$$\gamma = \frac{\sigma^2}{\mu_1{}'} = \frac{\theta^7+4\theta^6+6\theta^5+32\theta^4+120\theta^3+60\theta^2+144\theta+144}{\theta(\theta^3+\theta^2+6)(\theta^3+2\theta^2+24)}.$$

Research & Reviews: Journal of Statistics
Volume 13, Issue 1
ISSN: 2278-2273 (Online), ISSN: 2348-7909 (Print)

Graphical presentations of these descriptive measures of PUD for varying values of parameters are shown in Figure 2. The coefficient of variation, skewness, and kurtosis increased, and the index of dispersion decreased with increasing parameter values.



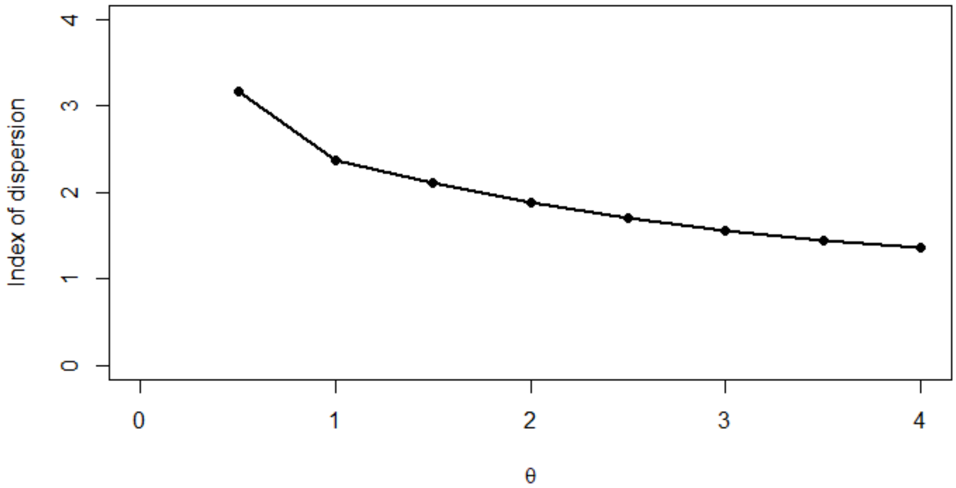


Figure 2. Coefficients of variation, skewness, kurtosis, and index of dispersion.

STATISTICAL PROPERTIES
Overdispersion
We have

$$\mu_2=\frac{\theta^7+4\theta^6+6\theta^5+32\theta^4+120\theta^3+60\theta^2+144\theta+144}{\theta^2\left(\theta^3+\theta^2+6\right)^2}$$

$$=\frac{\theta^3+2\theta^2+24}{\theta\left(\theta^3+\theta^2+6\right)}\left[\frac{\theta^7+4\theta^6+6\theta^5+32\theta^4+120\theta^3+60\theta^2+144\theta+144}{\theta\left(\theta^3+\theta^2+6\right)\left(\theta^3+2\theta^2+24\right)}\right]$$

$$=\frac{\theta^3+2\theta^2+24}{\theta\left(\theta^3+\theta^2+6\right)}\left[1+\frac{\theta^6+4\theta^5+2\theta^4+84\theta^3+60\theta^2+144}{\theta\left(\theta^3+\theta^2+6\right)\left(\theta^3+2\theta^2+24\right)}\right]$$

$$=\mu_1'\left[1+\frac{\theta^6+4\theta^5+2\theta^4+84\theta^3+60\theta^2+144}{\theta\left(\theta^3+\theta^2+6\right)\left(\theta^3+2\theta^2+24\right)}\right]$$

This shows that $\mu_2>\mu_1{}'$ and thus PUD is always over-dispersed distribution. Therefore, PUD can be used for discrete datasets that are overdispersed in nature.

Increasing Hazard Rate and Unimodality
It can easily be shown that PUD has an increasing hazard rate (IHR) and is unimodal. Because $\frac{P(x+1,\theta)}{P(x,\theta)}=\frac{1}{\theta+1}\left[1+\frac{3x^2+15x+(\theta^2+2\theta+14)}{x^3+6x^2+(\theta^2+2\theta+12)x+(\theta^3+4\theta^2+5\theta+8)}\right]$ is a decreasing function of x for a given θ, $P(x,\theta)$ is log-concave. This implies that PUD has an increasing hazard rate and that it is unimodal. Grandell (1997) [15] discussed the relationship between log-concavity, IHR, and the Unimodality of discrete distributions in detail.

ESTIMATION OF PARAMETER
Method of Moment Estimate
Let $x_1,x_2,\ldots,x_n$ be a random sample of size n from PUD. Equating the first moment about the origin to the corresponding sample moment, the moment estimate (ME) $\tilde{\theta}$ of θ is the solution to the following fourth-degree polynomial equation:

$\bar{x}\theta^4 + (\bar{x}-1)\theta^3 - 2\theta^2 + 6\theta\bar{x} - 24 = 0$, where $\bar{x}$ is the sample mean.

This equation can be solved using the Newton-Raphson method to obtain an estimate of the parameter.

Maximum Likelihood Estimate

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from PUD, and let f_x be the observed frequency in the sample corresponding to $X = x (x = 1,2,3,\ldots,k)$ such that $\sum_{x=1}^{k} f_x = n$, where k is the largest observed value with non-zero frequency. The likelihood function L of the PUD is given by

$$L = \left(\frac{\theta^4}{\theta^3+\theta^2+6}\right)^n \frac{1}{(\theta+1)^{\sum_{x=1}^{k} f_x(x+4)}} \prod_{x=1}^{k} [x^3 + 6x^2 + (\theta^2+2\theta+12)x + (\theta^3+4\theta^2+5\theta+8)]^{f_x}$$

The log -likelihood function is obtained as

$$log\, L = n\, log\left(\frac{\theta^4}{\theta^3+\theta^2+6}\right) - \sum_{x=1}^{k} f_x(x+4)\, log(\theta+1) + \sum_{x=1}^{k} f_x\, log[x^3 + 6x^2 + (\theta^2+2\theta+12)x + (\theta^3+4\theta^2+5\theta+8)]$$

The first derivative of the log -likelihood function is given by

$$\frac{d\, log\, L}{d\theta} = \frac{4n}{\theta} - \frac{n(3\theta^2+2\theta)}{\theta^3+\theta^2+6} - \frac{n(\bar{x}+4)}{\theta+1} + \sum_{x=1}^{k} \frac{[2(\theta+1)x + (3\theta^2+8\theta+5)]f_x}{x^3+6x^2+(\theta^2+2\theta+12)x+(\theta^3+4\theta^2+5\theta+8)}$$

Where, $\bar{x}$ is the sample mean.

The maximum likelihood estimate (MLE), $\hat{\theta}$ of θ is the solution of the equation $\frac{d\, log\, L}{d\theta} = 0$ and is given by the solution of the nonlinear equation:

$$\frac{4n}{\theta} - \frac{n(3\theta^2+2\theta)}{\theta^3+\theta^2+6} - \frac{n(\bar{x}+4)}{\theta+1} + \sum_{x=1}^{k} \frac{[2(\theta+1)x + (3\theta^2+8\theta+5)]f_x}{x^3+6x^2+(\theta^2+2\theta+12)x+(\theta^3+4\theta^2+5\theta+8)} = 0$$

Because these log-likelihood equations are not expressible in closed form, the MLE of the parameter θ can be computed iteratively by solving the log-likelihood equation using the Newton-Raphson iteration available in the R software. The initial value of the parameter θ may be taken as the value given by the method of moment estimation.

GOODNESS OF FIT

The applications of PUD have been discussed using four-count datasets relating to over-dispersed thunderstorm events. The goodness of fit of the PUD was compared with the one-parameter over-dispersed distributions, including PLD and PSD. The expected values given by the PLD and PSD are listed in the table for comparison. It is very clear based on values of chi-square from the goodness of fit presented in Tables 1, 2, 3, and 4 that PUD provides a better fit than PLD and PSD.

Table 1. Observed and expected number of days experienced X thunderstorm events at Cape Kennedy, Florida for the 11 years of record for June, January 1957 to December 1967, Falls et al. (1971) [1].

No. of thunderstorms	Observed frequency	PLD	PSD	PUD
0	187	185.3	186.4	187.5
1	77	83.5	82.3	80.6
2	40	35.9	35.5	35.4

No. of thunderstorms	Observed frequency	PLD	PSD	PUD
3	17	15.0	15.0	15.4
4	6	6.1	6.3	6.5
5	2	2.5	2.6	2.7
6	1	1.7	1.9	1.9
Total	330	330	330	330
ML estimate		$\hat{\theta} = 1.8042$	$\hat{\theta} = 1.6790$	$\hat{\theta} = 2.4866$
χ^2		1.43	1.48	1.32
d.f.		3	3	3
p-value		0.6985	0.6869	0.7243

Table 2. Observed and expected number of days experienced X thunderstorm events at Cape Kennedy, Florida for the 11 years of record for July, January 1957 to December 1967, Falls et al. (1971) [1].

No. of thunderstorms	Observed frequency	PLD	PSD	PUD
0	177	177.7	178.7	179.4
1	80	88.0	86.9	85.1
2	47	41.5	41.0	41.1
3	26	18.9	18.9	19.5
4	9	8.4	8.6	9.0
5	2	6.5	6.9	6.9
Total	341	341	341	341
ML estimate		$\hat{\theta} = 1.5835$	$\hat{\theta} = 1.4972$	$\hat{\theta} = 2.2750$
χ^2		5.15	5.41	4.86
d.f.		3	3	3
p-value		0.1611	0.1441	0.1824

Table 3. Observed and expected number of days experienced X thunderstorm events at Cape Kennedy, Florida for the 11 years of record for August, January 1957 to December 1967, Falls et al. (1971) [1].

No. of thunderstorms	Observed frequency	PLD	PSD	PUD
0	185	184.8	186.0	187.1
1	89	87.2	86.1	84.3
2	30	39.3	38.8	38.6
3	24	17.1	17.1	17.6
4	10	7.3	7.4	7.8
5	3	5.3	5.6	5.6
Total	341	341	341	341
ML estimate		$\hat{\theta} = 1.6934$	$\hat{\theta} = 1.5867$	$\hat{\theta} = 2.3844$
χ^2		5.03	4.87	4.54
d.f.		3	3	3
p-value		0.1696	0.1816	0.2087

Table 4. Observed and expected number of days experienced X thunderstorm events at Cape Kennedy, Florida for the 11 years of record for the summer, January 1957 to December 1967, Falls et al. (1971) [1].

No. of thunderstorms	Observed frequency	PLD	PSD	PUD
0	549	547.5	550.8	553.7
1	246	259.0	255.7	250.2

No. of thunderstorms	Observed frequency	PLD	PSD	PUD
2	117	116.9	115.5	115.3
3	67	51.2	51.1	52.5
4	25	21.9	22.3	23.3
5	7	9.2	9.6	9.6
6	1	6.3	7.0	7.0
Total	1012	1012	1012	1012
ML estimate		$\hat{\theta} = 1.6889$	$\hat{\theta} = 1.5824$	$\hat{\theta} = 2.3766$
χ^2		9.60	10.09	9.02
d.f.		4	4	4
p-value		0.0477	0.0389	0.0606

CONCLUDING REMARKS

In this paper, a Poisson compound of the Uma distribution called the Poisson-Uma distribution (PUD) has been suggested. The coefficients of variation, skewness, kurtosis, and index of dispersion were obtained, and their behaviors were studied. The obtained distribution is unimodal, has an increasing hazard rate, and is overdispersed. The method of moment and maximum likelihood estimation is discussed for estimating the parameter. The goodness of fit of the proposed distribution and its comparison with the Poisson-Lindley distribution (PLD) and Poisson-Sujatha distribution (PSD) on four datasets relating to thunderstorm events were studied.

REFERENCES

1. Falls LW, Williford WO, Carter MC. Probability distributions for thunderstorm activity at Cape Kennedy, Florida. J Appl Meteorol Climatol. 1971;10(1):97–104.
2. Carter MC. A model for thunderstorm activity: Use of the compound negative binomial-positive binomial distribution. J R Stat Soc Ser C Appl Stat. 1972;21(2):196–201.
3. Sankaran M. Note: The discrete Poisson-Lindley distribution. Biometrics. 1970;26(1):145–9.
4. Lindley DV. Fiducial distributions and Bayes' theorem. J R Stat Soc Ser B Methodol. 1958;20(1):102–7.
5. Ghitany ME, Atieh B, Nadarajah S. Lindley distribution and its application. Math Comput Simul. 2008;78(4):493–506.
6. Ghitany ME, Al-Mutairi DK. Estimation methods for the discrete Poisson-Lindley distribution. J Stat Comput Simul. 2009;79(1):1–9.
7. Shanker R, Hagos F. On Poisson-Lindley distribution and its applications to biological sciences. Biometrics Biostat Int J. 2015;2(4):103–7.
8. Shanker R. Sujatha distribution and its applications. Stat Transit New Ser. 2016;17(3):391–410.
9. Shanker R. The discrete Poisson-Sujatha distribution. Int J Probab Stat. 2016;5(1):1–9.
10. Shanker R, Fesshaye H. On Poisson-Sujatha distribution and its applications to model count data from biological sciences. Biometrics Biostat Int J. 2016;3(4):1–7.
11. Shanker R. Uma distribution with properties and applications. Biometrics Biostat Int J. 2022;11(5):165–9.
12. Shanker R, Zakaria AF, Prodhani HR, Ray M. Weighted Uma distribution with properties and applications. J Xidian Univ. 2023;17(10):830–46.
13. Shanker R, Baishya J, Ray M, Prodhani HR. Power Uma distribution with properties and applications in survival analysis. J Xidian Univ. 2023;17(12):849–60.
14. Tajuddin RR, Ismail N, Ibrahim K. Several two-component mixture distributions for count data. Commun Stat Simul Comput. 2022;51(7):3760–71.
15. Grandell J. Mixed Poisson processes. 2nd ed. Boca Raton: CRC Press; 2020.

