A NEW POISSON-JUCHEZ DISTRIBUTION WITH ITS PROPERTIES AND APPLICATION

[1]Koyejo S. O., [2*]Akomolafe A. A., [3]OSENI B. M. and [4]Ojo O. O.

Department of Statistics, Federal University of Technology, Akure
[1]koyejosamuel@yahoo.co.uk , [2]akomolafeayotade@gmail.com and [3]bmoseni@futa.edu.ng
[4]ooojo@futa.edu.ng

Abstract: A novel distribution called the Juchez Poisson distribution was proposed and studied. The distribution was developed by combining the Poisson and Juchez distributions, thereby making it more adaptable and better suited for analyzing complex datasets. Various statistical properties of the Poisson-Juchez distribution were derived. The model parameter was estimated using the method of maximum likelihood, where the asymptotic and consistent properties are numerically studied as well. The flexibility of this distribution was shown using simulation study; also, this newly developed distribution was demonstrated through its application to three real-life dataset on the number of Maternity lengths of stay in a clinic. The developed distribution was also compared with some existing Poisson-mixed distributions and the results showed that it performs better than other eleven Poisson-mixed distributions considered, thereby offering a more precise description of the data. This distribution shows a high efficiency when compared with other one parameter distributions. The findings thus show that Poisson-Juchez Distribution is a flexible and effective alternative for modeling count data, especially in situations involving overdispersion and other types of data variability in several fields.

[Koyejo S. O., Akomolafe A. A., OSENI B. M. and Ojo O. O. **A NEW POISSON-JUCHEZ DISTRIBUTION WITH ITS PROPERTIES AND APPLICATION.** *Researcher* 2026;18(4):10-21]. ISSN 1553-9865 (print); ISSN 2163-8950 (online). http://www.sciencepub.net/researcher. 02. doi:10.7537/marsrsj180426.02

Key words: Count Data; Poisson-Juchez; Overdispersion; Index of Dispersion; Compounding

1. INTRODUCTION

The Poisson distribution is commonly used to model the number of events happening in a specific time or space period, assuming these events are independent and occur at a constant average rate (Alomair & Ahsan-ul Haq, 2023). However, in many real-life situations, this assumption of a constant rate may be violated. For example, the rate at which events happen can change due to different people or factors that affect the rate. When this occurs, the data may show some variation than expected which is known as overdispersion. The standard Poisson model is not very good at handling this kind of variation (Hilbe, 2014). To address this, the mixed Poisson distribution, also known as the compound Poisson distribution, was formed. This model adds a random factor to the Poisson setup, allowing the rate parameter to vary depending on a specified distribution, like gamma or log-normal (Moller, et al., 1998). This flexibility helps the mixed Poisson distribution better reflect variation in event rates found in various situations. It has applications in fields like insurance, telecommunications and epidemiology, where event rates can be influenced by unseen factors (Woolson & Clarke, 2011). Many researchers have developed over-dispersed and under-dispersed discrete distributions by compounding the Poisson distribution with continuous distributions to form mixed Poisson distributions. Examples include the discrete Lindley distribution (Shafiq, et al., 2022), the Poisson-Ram Awadh distribution (Shukla, Shanker, & Tiwari, 2022). the Poisson-Chris-Jerry distribution (Ahmad & Wani, 2023), the discrete Poisson Bilal distribution (Ahsan, et al., 2023), and the semi-Poisson distribution (Almuhayfith et al., 2023) Recently, the two-parameter Poisson Rani distribution (Bamigbala et al., 2025), Poisson XRani Distribution (Borbye, Nasiru & Ajongba, 2024), Poisson-Pranav Distribution (Hassan, Wani, & Shafi, 2020), and Size-biased Poisson-Pranav Distribution (Shukla et al., 2023) have also been introduced. Ahmed, Afaq, and Rajnee (2020) proposed the Sauleh distribution, formulated mixture model combining an Exponential distribution (with a scale parameter) and a Gamma distribution (with shape parameter and a scale parameter), weighted by a mixing proportion. In this study, we introduce a new statistical model called the Poisson-Juchez distribution, designed to handle over-dispersed and heavy-tailed count data. This model combines the Poisson distribution with the Juchez distribution. The aim is to offer a better alternative to traditional models like Poisson distribution, Negative Binomial Distribution (NBD) and other similar models. The study explores the model's derivation, statistical properties, parameter estimation, and application to real datasets.

Poisson-Juchez distribution is an alternative Poisson-mixed distribution that is capable of modeling count data exhibiting over-dispersion. By incorporating this mixed distribution, it becomes possible to model the variability and patterns inherent in the data more accurately, leading to better predictions and deeper insights in count data modeling and application. Researchers have explored various applications of Poisson, Negative Binomial and several mixed distributions such as Poisson-negative binomial, Poisson-Lindly negative binomial, Poisson-Gold, Poisson-Bilal, Poisson-Adya, Poisson-Mirra, Poisson-Odoma, e.t.c. to model over-dispersed count data.

The analysis and modelling of count data are very common in many practical lifetime studies, such as the number of days before death in certain diseases or the number of cycles (runs) until a machine stops working and so on. Hence, a number of statistical distributions have been applied to model cases of a count random variable, which takes non-negative integer values. A good review of these distributions can be found in Johnson et al. (2005). The Poisson distribution is commonly used to model count data under the generalized linear model (GLM) framework (Nelder and Wedderburn, 1972, McCullagh and Nelder, 1989, Agresti, 2002). In Count Data analysis, traditional models such as the Poisson and the negative binomial distributions are used to describe the data if the variance and mean are equal. However, real-world count data may not always follow a simple Poisson distribution due to overdispersion (when the variance exceeds the mean) or underdispersion (when the variance is less than the mean). In these cases, more flexible models are required. One such model is the proposed Poisson-Juchez mixed distribution.

The most recent Poisson-mixed distributions proposed in the last five years include: Poisson- Bilal by Altun (2020), Poisson-Gold by Ahmad and Amjad (2021), Poisson-Mira by Radhakumari et al. (2022), Recently, the Poisson-Adya distribution was introduced by Shanker and Shukla (2022). The existing mixing Poisson distributions (some of which were afore mentioned) can be used for improving count data modeling. In this Research, a new mixed Poisson distribution was developed by combining the Poisson distribution and the Juchez distribution. The Juchez distribution, which is composite of both exponential and gamma distributions, is a continuous distribution proposed by Echebiri and Mbegbu (2022). In a bid to develop a new distribution for modelling count data thereby capturing over-dispersion, this paper proposes a mixed Poisson distribution that will be more flexible and applied in various fields like epidemiology, insurance, social sciences, etc., where count data often exhibit overdispersion. Poisson-Juchez distribution is a mixture of Poisson distribution with the composite of both exponential and gamma distributions.

2. METHODOLOGY

Suppose X follows a Poisson distribution (i.e., $X \sim P\ (\lambda)$), then its Probability Mass Function (PMF) and Cumulative Distribution Function (CDF) are respectively given as

$$f(x,\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0,1,2,\dots ; \lambda > 0 \qquad (1)$$

$$F(x,\lambda) = \sum_{k=0}^{x} \frac{e^{-\lambda}\lambda^k}{k!}, x = 0,1,2,\dots ; \lambda > 0 \qquad (2)$$

Where:

λ is the average of events per interval, x is the number of occurrences

e is the number 2.71828… (Euler's number) the base of the natural logarithms)

Now, suppose the Parameter λ in equation (1) follows Juchez distribution with parameter , then the pdf and cdf of Juchez distribution, proposed by Echebiri and Mbegbu (2022) are respectively given as

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

$$F(x,\theta) = 1 - \left(1 + \frac{\theta x[\theta^2+\theta^2 x^2+3\theta x+6]}{\theta^3+\theta^2+6}\right) e^{-\theta x} \qquad (4)$$

Where:

θ is the scale parameter and,

x is the number of occurrences

2.1. Derivation of Probability Distribution Function and Cummulative Distribution Function for Poisson-Juchez Distribution

Let x follow Poisson distribution with parameter λ> 0 which has the following pdf:

$$f(x,\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0,1,2,\dots ; \lambda > 0$$

Suppose the Parameter λ of the Poisson distribution follows Juchez distribution with parameter θ which has the following pdf:

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

If the parameter λ of the Poisson distribution follows a Juchez distribution, then

$$f(\lambda / \theta) = \frac{\theta^4}{\theta^3+\theta^2+6}(1+\lambda+\lambda^3)e^{-\theta\lambda}$$

According to the Compound-Poisson Theory as demonstrated by Sankaran (1970), the general formula for Compounding is:

$$f(x;\lambda,\theta) = \int_0^\infty f(X = x/\lambda)\, \mathrm{f}(\lambda/\theta) d\lambda$$

This gives:

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

Simplifying further

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

Finally, after solving the integral, the pdf for the Poisson-Juchez distribution is given by equation (5), and the corresponding cdf is given by equation (6)

$$f(x,\theta) = \frac{\theta^4}{(\theta^3+\theta^2+6)}\left[\frac{(1+\theta)^3+x(1+\theta)^2+(1+\theta)^2+x^3+6x^2+11x+6}{(1+\theta)^{x+4}}\right. \qquad (5)$$

The Probability Density Function (PDF) for Poisson-Juchez (PJ) distribution is given by equation (5) and the corresponding Cumulative Distribution Function (CDF) for Poisson-Juchez (PJ) distribution is given by equation (6) below

$$F(x) = \sum_{t=0}^{x} P_t(t)$$

$$= \sum_{t=0}^{x}\frac{\theta^4}{(\theta^3+\theta^2+6)}\left[\frac{(1+\theta)^3+t(1+\theta)^2+(1+\theta)^2+t^3+6t^2+11t+6}{(1+\theta)^{t+4}}\right]$$

$$= \frac{\theta^4}{(\theta^3+\theta^2+6)(1+\theta)^4}\sum_{t=0}^{x}\left[\frac{(1+\theta)^3+t(1+\theta)^2+(1+\theta)^2+t^3+6t^2+11t+6}{(1+\theta)^{t}}\right]$$

$$= \frac{\theta^4\begin{pmatrix}\theta^7+5\theta^6+10\theta^5+16\theta^4\\ +29\theta^3+37\theta^2+24\theta+6\end{pmatrix}}{\theta^4(\theta^3+\theta^2+6)(1+\theta)^4} - \frac{\theta^4(\theta+1)^{-x}\begin{pmatrix}\theta^6+\theta^5x+\theta^3x^3+5\theta^5+2\theta^4x+9\theta^3x^3+8\theta^4\\ +27\theta^3x+3\theta^2x^2+29\theta^3+21\theta^2x+37\theta^2+6\end{pmatrix}}{\theta^4(\theta^3+\theta^2+6)(\theta+1)^4}$$

$$= 1 - \frac{\left(\theta^6+\theta^5x+\theta^3x^3+5\theta^5+2\theta^4x+9\theta^3x^3+8\theta^4+27\theta^3x+3\theta^2x^2+29\theta^3+21\theta^2x+37\theta^2+6\right)}{(\theta^3+\theta^2+6)(\theta+1)^{x+4}} \qquad 6$$

Thus, the Probability density function (pDf) and cumulative distribution function (cdf) of the Poisson-Juchez mixed distribution are derived, using equations (5) and (6), respectively.

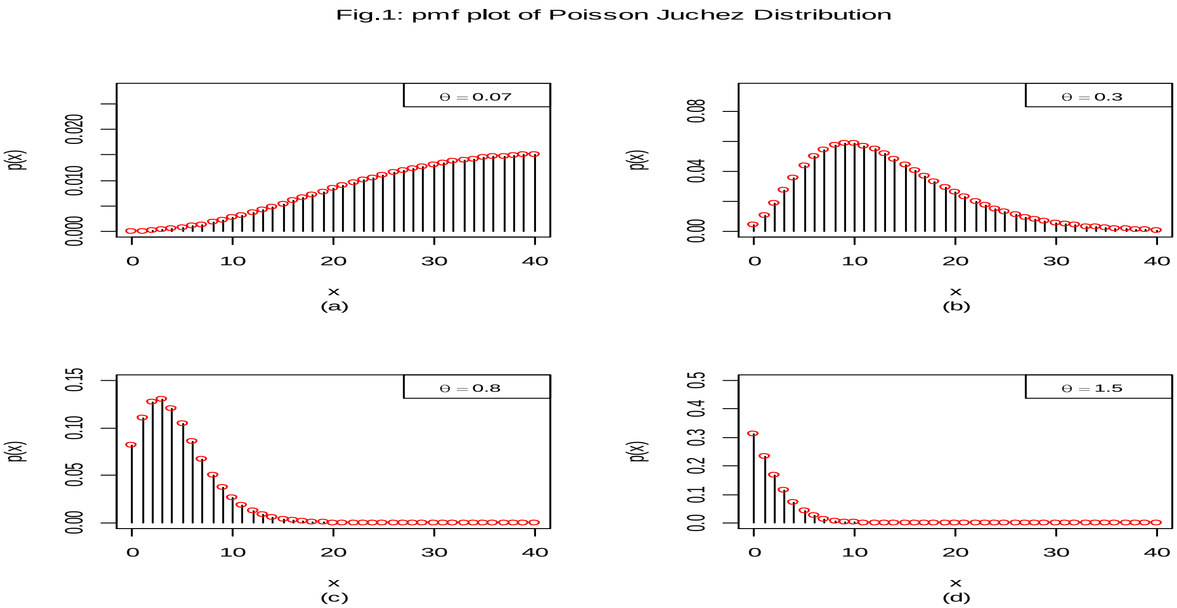

Figure 1: PMF Plot of Poisson-Juchez Distribution with parameter values 0.07, 0.3, 0.8 and 1.5

Figure 1 depicts the probability mass function (pmf) plots for Poisson-Juchez distribution using selected values of θ. It illustrates that the distribution is positively skewed, has one peak (i.e. unimodal and approximately symmetric.

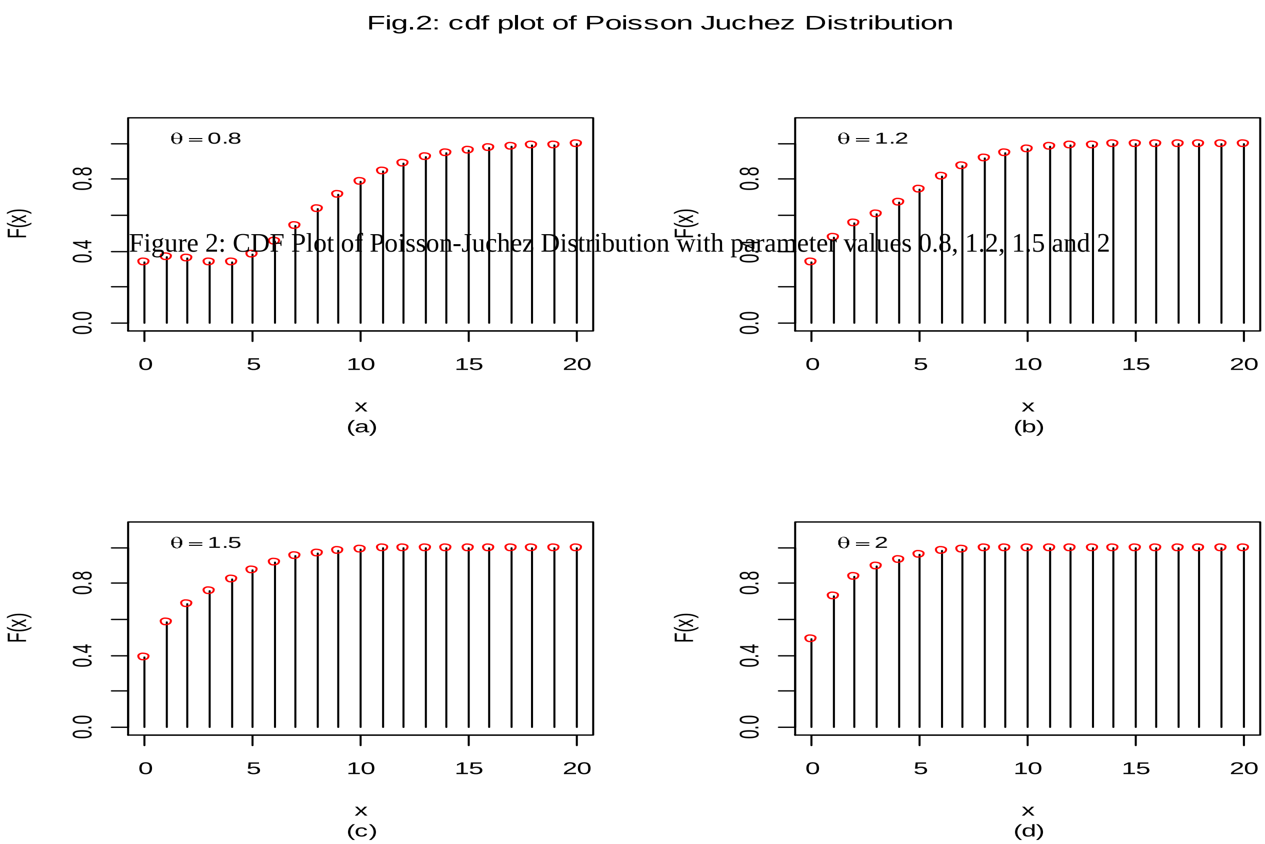

Figure 2: CDF Plot of Poisson-Juchez Distribution with parameter values 0.8, 1.2, 1.5 and 2

Figure 2 depicts the cumulative distribution function plots for Poisson-Juchez distribution, with parameter values 0.8, 1.2, 1.5 and 2. These plots show the distribution is an increasing function, meaning the probability accumulates smoothly over time, eventually reaching $\mathcal{F}(x) = 1$ as expected. For smaller parameter values such as 0.8, the CDF increases slowly, and it takes larger x values to reach 1. As the parameter value increases, the CDF becomes steeper and shifts to the left, showing that most of the probability mass is concentrated at smaller x values. At parameter 2, the CDF quickly approaches 1. The plots confirm the validity of the Poisson-Juchez distribution. Each CDF starts near 0, increasing continuously without decreasing, and approaches 1 as x becomes large, which are key characteristics of a valid cumulative distribution function. Across all parameter values (from 0.8 to 2), the curves continuously show this behavior, proving that the distribution properly accumulates probability. The smooth, non-decreasing nature of the CDFs and their convergence to 1 thus confirm that the Poisson-Juchez distribution meets the basic requirements of a valid probability distribution.

2.2. PROPERTIES

In this section, we discuss some of the Statistical properties of the Poisson-Juchez Distribution.

[i] The $\boldsymbol{r^{th}}$ moment is defined as:

$$\mu_r = \boldsymbol{E}(\boldsymbol{x^r}) = \sum_{\boldsymbol{x=0}}^{\infty} \boldsymbol{x^r f(x)}$$

$$\mu_r = \sum_{x=0}^{\infty} x^r . \frac{\theta^4}{(\theta^3+\theta^2+6)}\left[\frac{(1+\theta)^3 + x(1+\theta)^2 + (1+\theta)^2 + x^3 + 6x^2 + 11x + 6}{(1+\theta)^{x+4}}\right]$$

$$\mu_r = \frac{\theta^4}{(\theta^3+\theta^2+6)(1+\theta)^4}\sum_{x=0}^{\infty} x^r \left[\frac{(1+\theta)^3 + x(1+\theta)^2 + (1+\theta)^2 + x^3 + 6x^2 + 11x + 6}{(1+\theta)^{x}}\right]$$

Let $A = \frac{\theta^4}{(\theta^3+\theta^2+6)(1+\theta)^4}$

Let $\boldsymbol{r} = 1,2,3,4$ to derive the first, second, third and fourth Moments, respectively.

First Moment $r = 1$

$$\mu_1 = A\left[(1+\theta)^3 \sum_{x=0}^{\infty} x(1+\theta)^{-x} + (1+\theta)^2 \sum x^2(1+\theta)^{-x} + (1+\theta)^2 \sum x(1+\theta)^{-x} \right.$$
$$\left. + \sum_{x=0}^{\infty} x^4(1+\theta)^{-x} + 6\sum x^3(1+\theta)^{-x} + 11\sum x^2(1+\theta)^{-x} + 6\sum x(1+\theta)^{-x}\right]$$

$$\mu_1 = \frac{(\theta^3+2\theta^2+24)(\theta+1)^3}{\theta(\theta^3+\theta^2+6)(\theta+1)^3}$$

$$\mu_1 = \frac{\theta^3+2\theta^2+24}{\theta(\theta^3+\theta^2+6)} \tag{8}$$

Second Moment $r = 2$

$$\mu_2 = \frac{(\theta^4+4\theta^3+6\theta^2+24\theta+120)}{\theta^2(\theta^3+\theta^2+6)} \tag{9}$$

Third Moment $r = 3$

$$\mu_3 = \frac{(\theta^5+8\theta^4+24\theta^3+48\theta^2+360\theta+720)}{\theta^3(\theta^3+\theta^2+6)} \tag{10}$$

Fourth Moment $r = 4$

$$\mu_4 = \frac{(\theta^6+16\theta^5+78\theta^4+192\theta^3+960\theta^2+4320\theta+5040)}{\theta^4(\theta^3+\theta^2+6)} \tag{11}$$

[ii] Variance $\boldsymbol{\sigma^2} = \mu_2 - {\mu_1}^2$

$$\sigma^2 = \frac{(\theta^4+4\theta^3+6\theta^2+24\theta+120)}{\theta^2(\theta^3+\theta^2+6)} - \left(\frac{\theta^3+2\theta^2+24}{\theta(\theta^3+\theta^2+6)}\right)^2 \tag{12}$$

$$\sigma^2 = \frac{(\theta^7+4\theta^6+6\theta^5)+32\theta^4+120\theta^3+60\theta^2+144\theta+144}{\theta^2(\theta^3+\theta^2+6)^2} \tag{13}$$

[iii]Coefficient of Variation (CV)

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

$$CV = \frac{\sigma}{\mu_1} = \frac{\sqrt{(\theta^7+4\theta^6+6\theta^5)+32\theta^4+120\theta^3+60\theta^2+144\theta+144}}{\theta^3+2\theta^2+24} \quad (14)$$

[iv] Survival Function $S(x,\theta)$

$$S(x,\theta) = 1 - F(x)$$

$$S_{x=1-}\frac{(\theta^6+\theta^5 x+5\theta^5+2\theta^4 x+8\theta^4+\theta^3 x^3+9\theta^3 x^2+27\theta^3 x+29\theta^3+3\theta^2 x^2+21\theta^2 x+37\theta^2+6\theta x+24\theta+6)}{(\theta^3+\theta^2+6)(1+\theta)^{x+4}} \quad (15)$$

[v] Hazard Function $H(x,\theta)$

$$H(x,\theta) = \frac{f(x,\theta)}{S(x,\theta)}$$

$$H(x,\theta) = \frac{\theta^4(1+\theta)^3+(x+1)(1+\theta)^2+(x+3)(x+2)(x+1)}{\theta^6+\theta^5 x+\theta^3 x^3+5\theta^5+2\theta^4 x+8\theta^4+27\theta^3 x+9\theta^3 x^2+29\theta^3+3\theta^2 x^2+21\theta^2 x+37\theta^2+6\theta x+24\theta+6} \quad (16)$$

Table 1: Some Descriptive Statistics of Poisson Juchez Distribution Across Theta Values

Theta	Mean	Variance	Std Dev	CV	Index Dispersion	Skewness	Kurtosis
0.5	7.7255	24.4344	4.9431	0.6398	3.1628	1.0061	4.4881
1.0	3.3750	7.9844	2.8257	0.8372	2.3657	1.1453	4.7311
1.5	1.8280	3.8485	1.9618	1.0732	2.1054	1.4497	5.6766
2.0	1.1111	2.0988	1.4487	1.3038	1.8889	1.7794	7.1250
2.5	0.7480	1.2719	1.1278	1.5078	1.7005	2.0484	8.5897
3.0	0.5476	0.8509	0.9224	1.6845	1.5538	2.2464	9.7859
3.5	0.4271	0.6176	0.7859	1.8400	1.4461	2.3933	10.6853
4.0	0.3488	0.4771	0.6908	1.9802	1.3678	2.5103	11.3722
4.5	0.2946	0.3861	0.6214	2.1089	1.3104	2.6113	11.9364
5.0	0.2551	0.3234	0.5687	2.2289	1.2675	2.7037	12.4389
5.5	0.2251	0.2779	0.5272	2.3418	1.2346	2.7913	12.9144
6.0	0.2016	0.2436	0.4936	2.4489	1.2087	2.8759	13.3807
6.5	0.1826	0.2169	0.4657	2.5510	1.1880	2.9582	13.8461
7.0	0.1669	0.1955	0.4421	2.6488	1.1710	3.0386	14.3140
7.5	0.1538	0.1779	0.4218	2.7428	1.1569	3.1173	14.7858
8.0	0.1426	0.1633	0.4041	2.8336	1.1450	3.1945	15.2613
8.5	0.1330	0.1509	0.3885	2.9213	1.1349	3.2700	15.7404
9.0	0.1246	0.1403	0.3746	3.0063	1.1260	3.3441	16.2226
9.5	0.1172	0.1311	0.3621	3.0889	1.1183	3.4168	16.7074
10.0	0.1107	0.1230	0.3507	3.1692	1.1116	3.4881	17.1944

The index of dispersion values in the table shows some interesting characteristics of the Poisson Juchez distribution. All values are greater than 1.0 which suggests that the distribution is overdispersed compared to standard Poisson distribution for all theta values. The skewness values in the table show how the shape of the distribution changes as theta changes. For lower theta values (0.5 to 1.5), the skewness is moderate and positive (ranging from 1.01 to 1.45), indicating that the distribution is somewhat balanced with a slight right tail. However, as theta increases, the skewness increases sharply, reaching 3.49 when theta is 10.0.

The kurtosis in the table starts at 4.49 when theta is 0.5 and rises significantly to 17.19 when theta is 10.0. Since the kurtosis of a normal distribution is 3.0, all the values in the table suggest that the distribution has heavy tails, making it leptokurtic.

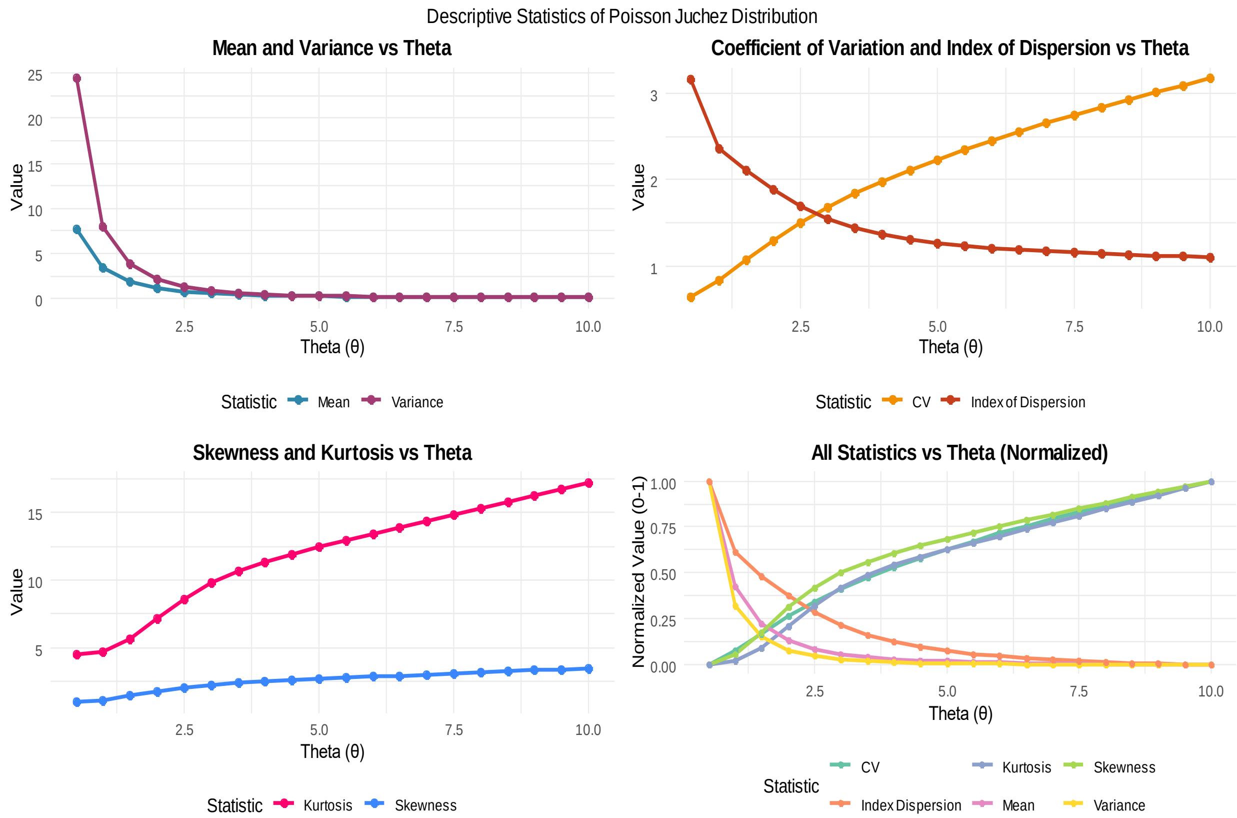

Figure 1: Descriptive Statistics of P[oisson-Juchez Distribution

3. PARAMETER ESTIMATION OF THE PJ DISTRIBUTION

In this study, the Maximum Likelihood Estimation is used to estimate the parameter, θ.
Let $X_1,, X_n$ be a random sample of size n drawn from the Poisson-Juchez distribution with the pmf in equation (5), then the likelihood function is given by:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i, \theta) \tag{17}$$

The logarithm of the likelihood function (log-likelihood function) for the parameter ϴ can be written as the product of individual probabilities for each observation, represented by the equation

$$\ell(\theta) = \sum_{i=1}^{n} \log\, f(x_i, \theta) \tag{18}$$

Recall:

$$f(x,\theta) = \frac{\theta^4[(1+\theta)^3 + x(1+\theta)^2 + (1+\theta)^2 + x^3 + 6x^2 + 11x + 6]}{(\theta^3+\theta^2+6)(1+\theta)^{x+4}}$$

so,

$$\log\, f(x,\theta) = 4\log\,\theta - \log\,(\theta^3+\theta^2+6) - (x+4)\log\,(1+\theta) + \log\,((1+\theta)^3 + x(1+\theta)^2 + (1+\theta)^2 + x^3 + 6x^2 + 11x + 6) \tag{19}$$

$$\ell(\theta) = \sum_{i=1}^{n} \{4\log\,\theta - \log\,(\theta^3+\theta^2+6) - (x+4)\log\,(1+\theta) + \log\,((1+\theta)^3 + x(1+\theta)^2 + (1+\theta)^2 + x^3 + 6x^2 + 11x + 6)\} \tag{20}$$

Taking the derivative with respect to the parameter θ yields

$$\frac{\partial\ell}{\partial\theta} = \frac{4}{\theta} - \frac{3\theta^2+2\theta}{\theta^3+\theta^2+6} - \frac{x+4}{1+\theta} + \frac{3(1+\theta)^2+2x(1+\theta)+2(1+\theta)}{(1+\theta)^3+x(1+\theta)^2+(1+\theta)^2+x^3+6x^2+11x+6} \tag{21}$$

The Maximum Likelihood Estimate of the parameter θ is obtained by equating (21) to zero and solving the resulting equation. The solution of (21) is best solved iteratively since it does not have a closed-form solution.

4. Simulation

A simulation study was conducted to evaluate the parameter estimation performance of the of the Poissonj-Juchez distribution. The simulation set up is as follows:
1. The sample sizes used were n = 10, 30, 50, 100, 250, 300, 500, 750 and 1000.
2. The data were generated from the Poisson-Juchez with different true parameter values as specified in Table 1.
3. Each scenario was repeated 500 times.

The simulation results are summarized in Tables 2a through 2b.
As the sample size increases, the average bias and root mean squared error (RMSE) of the MLE decrease for all cases.

Table 2a: Simulation Results for Poisson Juchez Distribution (θ = 0.25) Across Sample Sizes

Sample Size (n)	**Mean Estimate**	**Bias**	**Variance**	**MSE**	**RMSE**
10	0.257513	0.007513	0.002043	0.002099	0.045819
30	0.250833	0.000833	0.000693	0.000694	0.026337
50	0.251172	0.001172	0.000366	0.000367	0.019170
100	0.251250	0.001250	0.000208	0.000210	0.014482
250	0.250328	0.000328	0.000078	0.000078	0.008819
300	0.249866	-0.000134	0.000065	0.000065	0.008070
500	0.250192	0.000192	0.000036	0.000036	0.006002
750	0.250641	0.000641	0.000025	0.000026	0.005053
1000	0.250386	0.000386	0.000021	0.000021	0.004564

Table 2a presents the simulation results for the Poisson Juchez Distribution when the parameter θ ¸ = 0.25 across various sample sizes.

The mean estimates range from 0.257513 at n = 10 to 0.250386 at n = 1000, indicating consistent accuracy in estimation. The bias values are very small, with the highest absolute bias of 0.007513 occurring at the smallest sample size (n = 10), and decreasing significantly as the sample size increases. The negative bias observed at n = 300 (-0.000134) suggests minimal systematic error in estimation.

The variance and mean squared error (MSE) values also decrease with increasing sample size, from 0.002043 and 0.002099 at n = 10 to 0.000021 for both measures at n = 1000.
The root mean squared error (RMSE) follows an expected decline, decreasing from 0.045819 to 0.004564, showcasing improved estimation precision with larger samples. Overall, the results indicate that Î¸ = 0.25 is accurately estimated even at moderate sample sizes.

Table 2b: Simulation Results for Poisson Juchez Distribution (θ = 0.50) Across Sample Sizes

Sample Size	**Mean_Est**	**Bias**	**Variance**	**MSE**	**RMSE**
n = 10	0.521245	0.021245	0.009394	0.009846	0.099225
n = 30	0.506712	0.006712	0.003190	0.003235	0.056879
n = 50	0.500616	0.000616	0.001993	0.001994	0.044650
n = 100	0.501994	0.001994	0.000795	0.000799	0.028263
n = 250	0.500745	0.000745	0.000296	0.000296	0.017212
n = 300	0.500548	0.000548	0.000287	0.000288	0.016961
n = 500	0.499389	-0.000611	0.000163	0.000163	0.012779
n = 750	0.500411	0.000411	0.000122	0.000123	0.011074
n = 1000	0.500011	0.000011	0.000092	0.000092	0.009566

Table 2b shows robust estimation performance for θ = 0.50, though with slightly higher variability compared to the lower parameter value in Table 2a. The mean estimates in Table 2b range from 0.521245 at n = 10 to 0.500011 at n = 1000, demonstrating good convergence to the true parameter value. Table 2b reveals that bias decreases consistently from 0.021245 at small sample sizes to near-zero values (0.000011) at n = 1000, with one instance of slight negative bias (-0.000611) at n = 500.

The variance measures presented in Table 2b decrease systematically from 0.009394 at n = 10 to 0.000092 at n = 1000, indicating improved precision with larger samples. Table 2b shows that MSE values follow a similar declining pattern, from 0.009846 to 0.000092, while RMSE decreases from 0.099225 to 0.009566. The estimation performance documented in Table 2b suggests that θ = 0.50 requires moderate sample sizes (n ≥ 100) for highly accurate estimation, with acceptable performance even at smaller sample sizes.

Table 2c: Simulation Results for Poisson Juchez Distribution (θ = 1.00) across Sample Sizes

Sample Size	Mean_Est	Bias	Variance	MSE	RMSE
n = 10	1.032130	0.032130	0.038001	0.039033	0.197568
n = 30	1.003933	0.003933	0.012238	0.012254	0.110696
n = 50	1.008826	0.008826	0.007649	0.007727	0.087906
n = 100	1.004635	0.004635	0.003398	0.003420	0.058480
n = 250	0.999293	-0.000707	0.001475	0.001476	0.038418
n = 300	0.999639	-0.000361	0.001415	0.001415	0.037619
n = 500	1.000424	0.000424	0.000745	0.000745	0.027301
n = 750	0.999960	-0.000040	0.000505	0.000505	0.022464
n = 1000	1.000513	0.000513	0.000391	0.000391	0.019778

Table 2c demonstrated adequate estimation performance for θ = 1.00, with more pronounced variability at smaller sample sizes compared to lower theta values. The mean estimates shown in Table 2c range from 1.032130 at n = 10 to 1.000513 at n = 1000, indicating reasonable accuracy across sample sizes. Table 2c reveals larger bias values at small sample sizes, with the maximum bias of 0.032130 at n = 10, though bias approaches zero as sample size increases, with values near ±0.001 for larger samples.

The variance patterns in Table 2c show substantial improvement with sample size, decreasing from 0.038001 at n = 10 to 0.000391 at n = 1000. Table 2c indicates that MSE values follow a similar trajectory, from 0.039033 to 0.000391, while RMSE decreases dramatically from 0.197568 to 0.019778. The results in Table 2c suggest that accurate estimation of θ = 1.00 requires larger sample sizes compared to smaller parameter values, with n ≥ 250 recommended for high precision

5. APPLICATIONS TO REAL DATA SETS

The real data set considered to fit the proposed distribution was the number of Maternity lengths of stay, obtained from Bamigbala et al. (2025). These data set was fitted to the proposed distribution (PJD) and eleven others from the existing Poisson-mixed distributions. Descriptive statistics of this data are shown in Table 1. The index of dispersion is greater than unity for the data set, indicating that the data set is over-dispersed. In this study, the Raphson-Newton method based on the R programming language (R Core Team, 2014) is used to compute the maximum likelihood estimates (Nash, 2014). being a global optimization method, Tables 2, 3, 4 and 5 present the results of fitting the different distributions to these real data sets. We use the estimated log-likelihood (LogLik), AIC, BIC, Chi-square and P-value model criteria selection to compare the observed and expected values of each data set.

Table 3: Descriptive Statistics for the real-life data set

Data sets	Descriptions of data	Mean	Standard Deviation	Variance	Skewness
I	number of maternity lengths of stay	0.7833	1.1213	1.2573	1.341

Table 3 presents the fundamental characteristics of the real-world dataset used in this comparative study. The maternity length of stay dataset exhibits a mean of 0.7833with a variance of 12573, resulting in a variance-to-mean ratio of approximately 1.60, indicating clear overdispersion and thus demonstrates the inadequacy of the standard Poisson distribution, which assumes equal mean and variance, necessitating more flexible modeling approaches.

Table 4: Observed and Expected Frequency for Data Fitted on Number of Maternity Length of Stay

Count	Observed	Poisson	Poisson Janardan	Poisson Lindley	Poisson Odoma	Poisson Pranav	Poisson Quasi Lindley	Poisson Shanker	Negative Binomial	Poisson Ishita	Poisson Sushila	Poisson Mirra	Poisson Juchez
0	176	120.15	168.66	170.34	178.25	178.35	168.66	199.18	168.88	174.28	168.67	169.22	171.21
1	86	139.29	100.41	99.21	89.14	89.89	100.41	80.49	101.27	94.63	100.41	98.36	96.56
2	52	80.74	55.40	54.58	51.14	50.83	55.40	32.52	54.80	52.83	55.40	55.88	54.62
3	33	31.20	29.18	28.94	29.88	29.36	29.18	13.14	28.59	29.14	29.18	30.09	29.93
4	25	9.04	14.90	14.95	16.83	16.52	14.90	5.31	14.64	15.67	14.90	15.39	15.71
5	11	2.10	7.44	7.57	9.01	8.94	7.43	2.15	7.41	8.20	7.43	7.54	7.90

Table 4 analyzes the maternity length of stay dataset, The observed data shows 176 zero counts and decreasing frequencies for longer stays. The Poisson Juchez model provides reasonable fit with expected frequency (171.21) close to observed zero counts, though not as precise as in previous datasets.

The Poisson Odoma model demonstrates competitive performance, while the standard Poisson model again underestimates zero counts (120.15 vs. observed 176). The Poisson Shanker model significantly overestimates zero counts (199.18), indicating poor model specification. Most mixed Poisson distributions show improved performance over the standard Poisson, though the differences are less pronounced than in the previous datasets.

Table 5: Model Selection Criteria for Number of Maternity Length of Stay Across Different Count Distributions

Distribution	LogLik	AIC	BIC	ChiSq	PValue	Parameters
Poisson Juchez	568.1375	1138.275	1142.223	8.437356	7.680897e-02	theta = 1.9475
Poisson Odoma	568.2358	1138.472	1142.420	4.890816	2.986851e-01	theta = 2.119
Poisson Pranav	568.8586	1139.717	1143.665	5.502863	2.394780e-01	theta = 1.7854
Poisson Ishita	568.9947	1139.989	1143.937	7.828296	9.807433e-02	theta = 1.494
Poisson Mirra	568.2725	1140.545	1148.441	9.967584	1.884370e-02	theta = 1.7466, alpha = 3.2061
Poisson Lindley	569.5634	1141.127	1145.075	10.949824	2.713374e-02	theta = 1.2458
Poisson Quasi Lindley	569.5125	1143.025	1150.921	11.658496	8.649440e-03	theta = 1.3028, alpha = 0.9596
Poisson Sushila	569.5125	1143.025	1150.921	11.657158	8.654803e-03	theta = 0.96, alpha = 0.7369
Poisson Janardan	569.5125	1143.025	1150.921	11.657142	8.654869e-03	theta = 1.7689, alpha = 1.3578
Negative Binomial	570.2619	1144.524	1152.420	12.499999	5.852665e-03	r = 1.2421, p = 0.4827
Poisson	605.6669	1213.334	1217.282	122.656548	1.445953e-25	theta = 1.1593
Poisson Shanker	652.7254	1307.451	1311.399	154.260725	2.486160e-32	theta = 1.4747

Table 5 identifies the Poisson Juchez model as optimal for maternity length of stay data, with the lowest AIC (1138.275) and reasonable goodness-of-fit statistics (Chi-square: 8.437356, p-value: 0.07680897). The Poisson Odoma model ranks second with AIC of 1138.472, showing competitive performance. The Poisson Pranav model also demonstrates strong performance with AIC of 1139.717.

The standard Poisson distribution exhibits severe inadequacy with substantially higher AIC (1213.334) and extremely low p-value (1.445953e-25). The Poisson Shanker model shows the worst performance with AIC of 1307.451 and p-value of 2.486160e-32, confirming its consistent inability to adequately model overdispersed count data across all three datasets.

From the model selection Criteria the dataset was subjected to, the Poisson Juchez distribution emerges as the superior model, demonstrating robust performance in handling overdispersed count data with varying characteristics. The standard Poisson distribution consistently fails to provide adequate fit, confirming the necessity of more flexible modeling approaches for real-world count data. Mixed Poisson distributions, particularly Poisson Juchez, Poisson Odoma, and Poisson Ishita, consistently outperform traditional alternatives, supporting their utility in practical applications involving overdispersed count phenomena.

6. CONCLUSION

The newly proposed mixed distribution (Poisson-Juchez distribution) presents a significant advancement in modeling over-dispersed count data. Through property derivations and analysis, we have shown that the PJD effectively captures the complexities of real-world datasets, outperforming traditional models such as the Poisson and Negative Binomial distributions. The application of the PJD to Maternity length of stay dataset had demonstrated its practical flexibility and robustness. The goodness-of-fit tests indicate that the PJD provides a superior fit, as evidenced by lower Chi-Square values and better AIC and BIC scores.

The shapes of the distribution are also investigated, showing that the distribution is positively skewed, unimodal and approximately symmetric. The smooth, non-decreasing nature of the CDFs and their convergence to unity further confirm that the Poisson-Juchez distribution satisfies the fundamental requirements of a valid probability distribution. The result from this study confirms the consistency of the Poisson Juchez distribution as the superior model for the dataset used and in all model selection criteria employed, demonstrating robust performance in handling overdispersed count data with varying characteristics. The standard Poisson distribution consistently fails to provide adequate fit, confirming the necessity of more flexible modeling approaches for real-world count data. Mixed Poisson distributions, particularly the proposed Poisson Juchez, consistently outperformed traditional alternatives, supporting its flexible utility in practical applications involving overdispersed count phenomena.

Future studies may explore further applications of the PJD and its potential extensions to other types of data, projecting its role as a versatile tool in statistical modeling.

REFERENCES

Ahmad, A., Amjad, D., (2021). New Distribution for fitting Discrete Data: The Poisson-Gold Distribution and its Statistical Properties. Austrian Journal of Statistics. July 2021, Volume 50, 19–35. http://www.ajs.or.at/ doi:10.17713/ajsv50i4.1091

Anwar et al. (2020). Poisson-Pranav distribution and its applications. Pak. J. Statist. 2020 Vol. 36(1), 57-72

Akomolafe et al. (). Exponential Pareto Negative Binomial distribution with its properties and applications. Int'l. Journal of Advanced Research and publications. ISSN: 2456-9992

Ayat et al., (2022). Poisson-Odoma Distribution. Journal of Statistics Applications & Probability. 11, No. 1, 155-161

Bamigbala, et al., (2025). Poisson Sauleh Distribution and its Properties. International Journal of Development Mathematics Vol 2 Issue 1 Page No. 246 – 266. ISSN: 3026-8656 (Print) | 3026-8699 (Online). journal homepage: https://ijdm.org.ng/index.php/Journals

Bhati et al., (2017). A New Count Model generated from mixed Poisson-transmuted exponential family.

Bodhisuwan,W., Pudprommarat, C., Bodhisuwan, R. and Saothayanun, L. (2017). Zero truncated Negative Binomial-Erlang Distribution. AIP conference Proceedings, http://www.bjs-uniben.org/ On Zero Truncated... 30 1905, 050011, American Institute of Physics

Cameron, A. C, & Trivedi, P. K. (2013). *Regression Analysis of Count Data*. Cambridge University Press.

Cheng, L., Geedipally, S.R., Lord, D. (2013). The Poisson-Weibull generalized linear model for analyzing motor vehicle crash data.

Déniz., E.G. (2013). A New Discrete Distribution: Properties and Applications in Medical care. Journal of Statistics

Echebiri, U.V., & Mbegbu, J.I., (2022). *Juchez Probability Distribution: Properties and Applications*. Asian Journal of Probability and Statistics. 20(2): 57-71, Article no. AJPAS.92104

Emrah, A., (2020). A New One-Parameter Distribution with associated Regression and Integer-Valued Autoregressive Models. Math. Slovaca . 2; (4):979-994

Famoye and Singh (2006). G eneralized Poisson regression model,

Grine, R., Zeghbud, H. (2017). On Poisson quasi-Lindley distribution and Applications Count Model generated from mixed Poisson-transmuted exponential family.

Hilbe, J.M. (2011). *Negative Binomial Regression*. Cambridge University Press

Imoto et al., (2017). A Modified Conway-Maxwell Poisson type Binomial Distribution and its Properties.

Mahmoudi, E., Zakerzadeh, H. (2010). Generalized Poisson-Lindley Distribution. Communications in Statistics: Theory and Methods, 39, 1785-1798. https://doi.org/10.1080/03610920902898514

Munindra et al. (2017). A New Quasi Poison-Lindley Distribution: Properties and Applications. Journal of Statistical Theory and Applications, Vol. 16, No. 4 (576–588)

Shukla, K.K, (2018). Pranav distribution with properties and its applications. Biom Biostat Int J. 2018;7(3): 244–254. DOI: 10.15406/bbij.2018.07.00215

Radhakumari et al. (2022). On Discrete Poisson-Mirra Distribution: Regression, INAR (1) Process. Axioms, 193

Shanker, R., Hagos, F., (2015). On Poisson quasi-Lindley distribution and its Applications to Biological Sciences. Biometrics & Biostatistics International Journal; 2 (4):103–107.

Shanker, R., Shukla, K. K., Ranjan, A. et al. (2021). Adya Distribution with properties and application. Biometrics & Biostatistics International Journal: 10 (3):81–88.

Willmot, G. E. (1987). The Poisson-Inverse Gaussian Distribution as an Alternative to the Negative Binomial. *Scandinavian Actuarial Journal*.

Winai, B., and Sirinapa, A., (2022). The Poisson-Transmuted Janardan Distribution for Modelling Count Data. Trends in Sciences; 19(5): 2898

Wongrin, W., and Bodhisuwan, W. (2016). The Poisson-generalised Lindley distribution and its applications. Songklanakarin J. Sci. Technol. 38 (6), 645-656, Nov. - Dec. 2016

Zamani, H., Ismail, N., Faroughi, P. (2014). Poisson-weighted exponential univariate version and regression model with applications.

Zhang, Z., et al. (2019). Modeling Hospital Admission Rates using Poisson-Tweedie Distributions." *Statistics in Medicine*.