Innovation in Statistics and Probability
1(2), 45–59
DOI: https://doi.org/10.64389/isp.2025.01241
https://sphinxsp.org/journal/index.php/isp/

Research article

On the Univariate and Multivariate Applications of the Poisson Two Parameter Chris-Jerry Distribution to Count Data

Kingsley Kuwubasamni Ajongba[1], Abdul Ghaniyyu Abubakari[2], Suleman Nasiru[3,*]

[1] Department of Statistics and Actuarial Science, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Navrongo, Ghana. kajongba.stu@cktutas.edu.gh

[2] Department of Applied Statistics, Faculty of Social Science and Arts, Simon Diedoung Dombo University of Business and Integrated Development Studies, Wa, Ghana. agabubakari@ubids.edu.gh

[3] Department of Statistics and Actuarial Science, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Navrongo, Ghana. sulemanstat@gmail.com

* **Correspondence:** agabubakari@ubids.edu.gh

ARTICLE INFO

Keywords:
Chris-Jerry distribution
Poisson distribution
Over-dispersion
Count data
Regression model

Mathematics Subject Classification:
62E10, 62H05

Important Dates:
Received: 10 October 2025
Revised: 6 December 2025
Accepted: 16 December 2025
Online: 18 December 2025

ABSTRACT

Statistical analysis and data modeling are necessary in explaining and extracting important information from real-world scenarios. In this study, we illustrated the univariate and multivariate applications of the Poisson two parameter Chris-Jerry (PTPCJ) distribution proposed by [9]. The features of the PTPCJ distribution indicate that the distribution is flexible for modeling positively skewed and approximately symmetric datasets. The PTPCJ distribution has a dispersion index greater than one. That makes it suitable for modeling count data which exhibit over-dispersed characteristic. Thus, we illustrated the univariate applicability of the PTPCJ distribution utilizing three datasets, and in all cases, it outperforms the competing models. Furthermore, multivariate application is demonstrated using the PTPCJ regression model when the response variable conforms to the PTPCJ distribution. The applicability of the regression model shows its superiority in modeling count data.

1. Introduction

The Poisson distribution is the traditional probability distribution for describing discrete data, with the characteristic of the variance of the data being equal to its mean [17]. This characteristic is described as equi-dispersion. However, real-world circumstances from most fields of study such as health, insurance, transportation, sports, engineering, astronomy, agriculture and finance produce discrete data that are over-dispersed, that is with variance to mean ratio of the data being greater than one. This calls for modified forms of the Poisson distribution that are capable of explaining real-world circumstances where data has greater variance as compared to its mean. In this regard, several techniques have been introduced to extend the Poisson distribution in order to adequately explain dispersed count data. Notable among these techniques is the mixed-Poisson (MP) distribution [1].

A well-known and widely used MP distribution is the negative binomial (NB) distribution which is a mixture of the Poisson and gamma distributions. Other MP distributions introduced in literature include: quasi Poisson-Garima [39], new quasi Poisson-Sujatha [38], discrete extended odd Weibull exponential [30], ZI Bell–Touchard [13], Poisson moment exponential [4], size-biased Poisson-Pranav [40], Poisson new XLindley [26], Poisson–NXLindley [34], Poisson Ramos-Louzada [10], Poisson transmuted moment exponential [15], Poisson two sum Lindley [18], Poisson Epanechnikov-exponential [27], Poisson Chris-Jerry [3], Poisson-modified Mishra [33], Poisson area-biased Ailamujia [5], discrete Marshall–Olkin inverse Toppe–Leone distribution [11], discrete Poisson quasi-XLindley [7], Poisson-modified Lindley [19] and Poisson-Mirra [29] distributions among others. The MP distributions have proven to always have greater variance than mean which is a desirable property for modeling over-dispersed observations.

Although several MP distributions have been introduced by researchers to handle over-dispersed count observations, there is the necessity for additional MP models that can be applied to various circumstances. Also, a few of the MP models have been introduced with regression models for modeling data with over-dispersed response variable. Some of these models include: Poisson-Mirra regression model [29], Poisson transmuted record type exponential regression model [23], Poisson-modification of quasi Lindley regression model [41] and Poisson quasi XLindley regression model [8] among others. Though there exist these regression models, investigations are continuously being carried out to discover befitting new regression models under diverse conditions.

The major motivation of this study is to demonstrate the univariate and multivariate applications of the Poisson two parameter Chris-Jerry (PTPCJ) distribution proposed by [9].

The structure of the remaining article is as follows: Section 2 gives a rendering of the formulation of the PTPCJ distribution and Section 3 presents the use of three different data sets to show the utility of the PTPCJ distribution. The PTPCJ regression model is introduced in Section 4 and Section 5 outlines the applicability of the PTPCJ regression model. Finally, Section 6 outlines the conclusion of this study.

2. Poisson Two Parameter Chris-Jerry Distribution

In this section, the PTPCJ distribution developed by [9] is introduced by using the continuous mixtures approach to amalgamate the Poisson and the two Poisson Chris-Jerry (TPCJ) distributions proposed by [20]. The Poisson distribution is presented in equation (2.1) via its probability mass function (pmf) as

$$\mathbb{P}_{Poisson}(z) = \frac{\lambda^z \exp(-\lambda)}{z!}, z = 0, 1, 2, ..., \lambda > 0. \tag{2.1}$$

Also, the probability density function (pdf) is presented for the TPCJ distribution as

$$g(z) = \frac{\varphi^2}{\varsigma\varphi + 2}(\varsigma + \varphi z^2) \exp(-\varphi x), x > 0, \varsigma > 0, \varphi > 0. \tag{2.2}$$

Letting $x|\lambda \sim Poisson(\lambda)$ and $\lambda|\varsigma, \varphi \sim TPCJ(\varsigma, \varphi)$, that is the Poisson distribution parameter in equation (2.1) is random variable which follows the TPCJ distribution defined in equation (2.2).

Proposition 1. *The pmf of the PTPCJ distribution is obtained as*

$$\mathbb{P}_{PTPCJ}(X = x) = \frac{\varphi^2(\varsigma(\varphi + 1)^2 + \varphi\pi(x))}{(\varsigma\varphi + 2)(\varphi + 1)^{x+3}}, x = 0, 1, 2, \dots, \varsigma > 0, \varphi > 0, \tag{2.3}$$

where $\pi(x) = (x + 1)(x + 2)$.

Proof. The pmf of the PTPCJ distribution is obtain using

$$\mathbb{P}_{PTPCJ}(X = x) = \int_0^\infty \mathbb{P}_{Poisson}(x)g(\lambda)d\lambda. \tag{2.4}$$

Equations (2.1) and (2.2) are substituted into equation (2.4) and simplified as follows

$$\begin{aligned}\mathbb{P}_{PTPCJ}(X = x) =& \frac{\varphi^2}{(\varsigma\varphi + 2)x!} \int_0^\infty \lambda^x(\varsigma + \varphi\lambda^2)e^{-\lambda(\varphi+1)}d\lambda \\ =& \frac{\varphi^2}{(\varsigma\varphi + 2)x!} \left[\varsigma \int_0^\infty \lambda^x e^{-\lambda(\varphi+1)}d\lambda + \varphi \int_0^\infty \lambda^{x+2}e^{-\lambda(\varphi+1)}d\lambda\right] \\ =& \frac{\varphi^2}{(\varsigma\varphi + 2)x!} \left[\frac{\varsigma}{(\varphi + 1)^{x+1}}\Gamma(x + 1) + \frac{\varphi}{(\varphi + 1)^{x+3}}\Gamma(x + 3)\right]. \end{aligned} \tag{2.5}$$

Further simplifying equation (2.5) completes the proof. □

The corresponding cumulative distribution function (cdf) of the PTPCJ distribution is defined as

$$F_{PTPCJ}(x) = \sum_{y=0}^{x} \frac{\varphi^2(\varsigma(\varphi + 1)^2 + \varphi\pi(y))}{(\varsigma\varphi + 2)(\varphi + 1)^{y+3}}, \tag{2.6}$$

where $\pi(y) = (y + 1)(y + 2)$. Expanding equation (2.6) gives

$$F_{PTPCJ}(x) = 1 - \frac{\varsigma\varphi(\varphi + 1)^2 + x^2\varphi^2 + 5x\varphi^2 + 6\varphi^2 + 2x\varphi + 6\varphi + 2}{(\varsigma\varphi + 2)(\varphi + 1)^{x+3}}.$$

For some values of the parameters of the PTPCJ distribution, graphical illustration of the shapes of the pmf are given in Figure 1. It can be seen that the PTPCJ distribution exhibits right-skewness and approximately bell-shaped shapes.

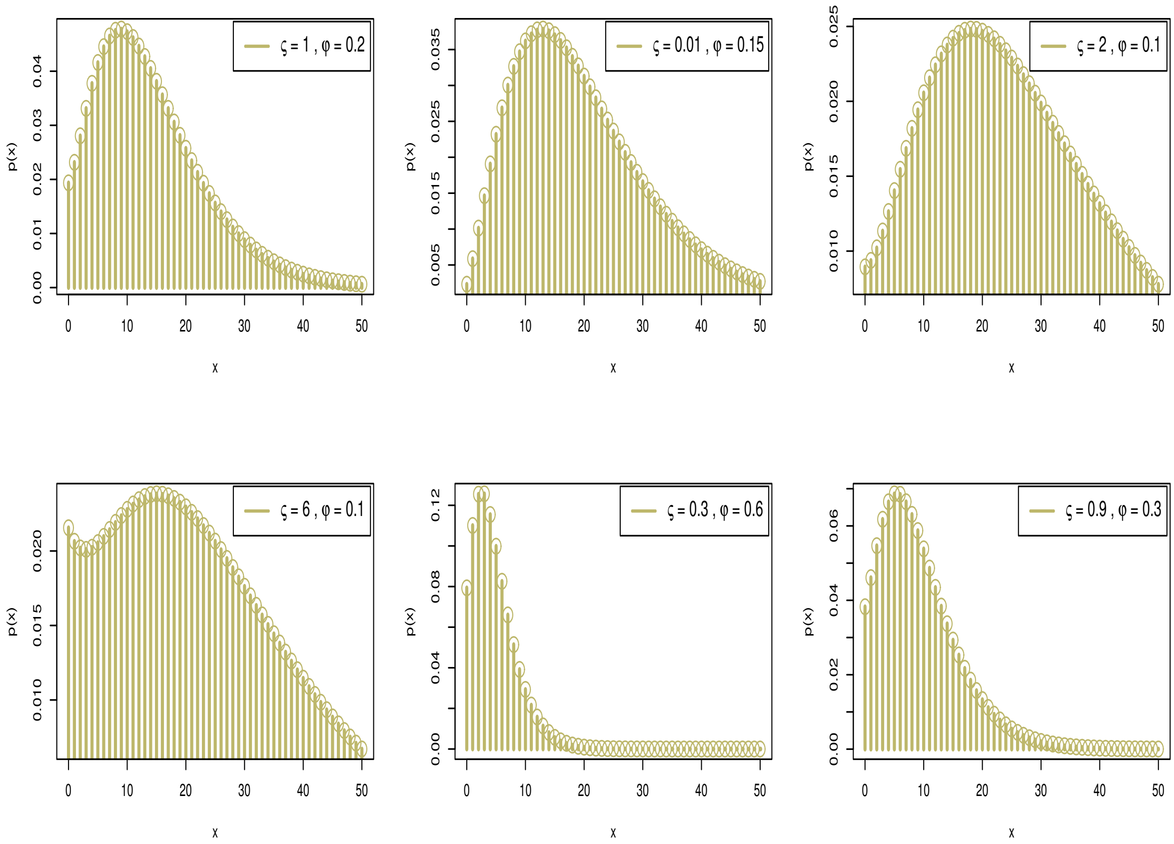

Figure 1. pmf shapes of the PTPCJ distribution

The survival function of the PTPCJ distribution is obtained as

$$
\begin{aligned}
S_{PTPCJ}(x) =& 1 - F_{PTPCJ}(x) \\
=& \frac{\varsigma\varphi(\varphi+1)^2 + x^2\varphi^2 + 5x\varphi^2 + 6\varphi^2 + 2x\varphi + 6\varphi + 2}{(\varsigma\varphi+2)(\varphi+1)^{x+3}}.
\end{aligned}
$$

The failure rate function of the PTPCJ distribution is also obtained as

$$
\begin{aligned}
h_{PTPCJ}(x) =& \frac{\mathbb{P}_{PTPCJ}(X=x)}{S_{PTPCJ}(x)} \\
=& \frac{\varphi^2(\varsigma(\varphi+1)^2 + \varphi\pi(x))}{\varsigma\varphi(\varphi+1)^2 + x^2\varphi^2 + 5x\varphi^2 + 6\varphi^2 + 2x\varphi + 6\varphi + 2}.
\end{aligned}
$$

Figure 2 illustrates some shapes of the failure rate of the PTPCJ distribution for selected parameter values. Increasing, decreasing and bathtub failure rate shapes are seen in Figure 2.

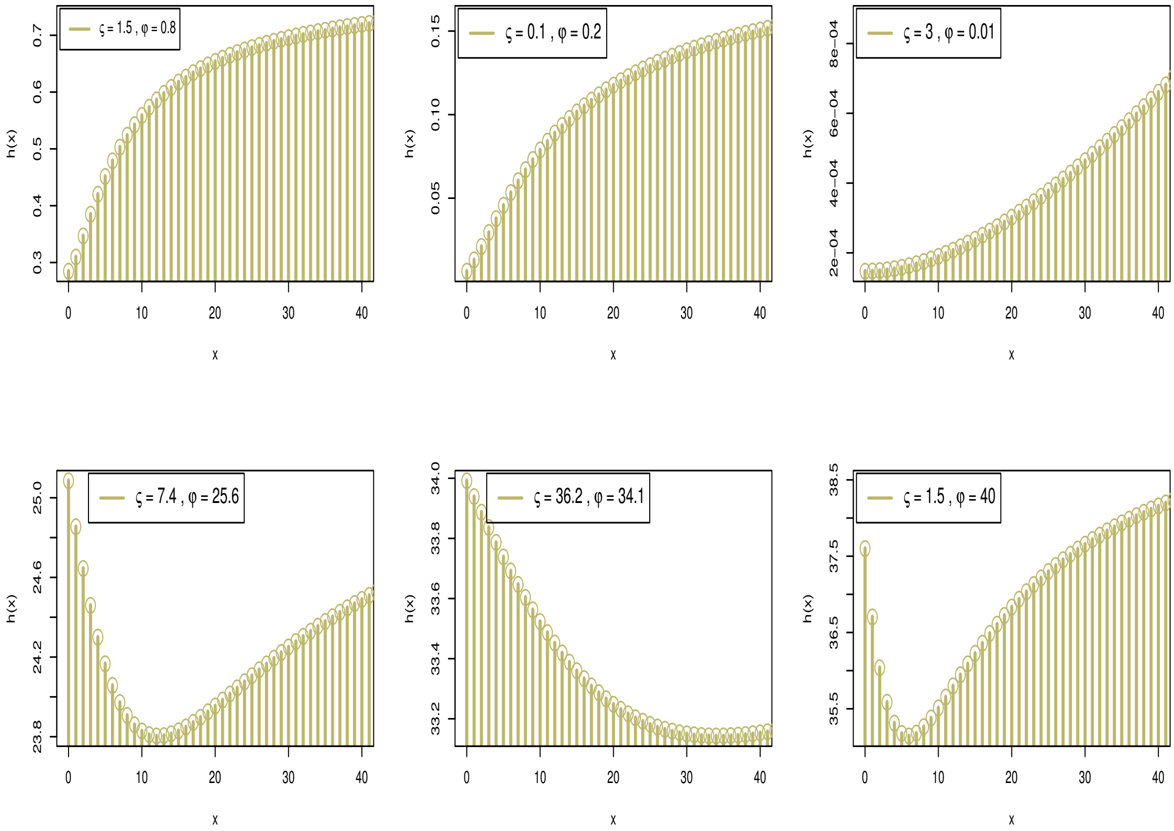

Figure 2. Failure rate shapes of the TPPCJ distribution

3. Univariate Applications of the PTPCJ Distribution

The utility of the PTPCJ distribution is established in this section using three real datasets. The PTPCJ distribution is compared with the Poisson distribution, Poisson entropy based weighed exponential (PEBWE) distribution [14], Poisson Janardan (PJ) distribution [37], Poisson quasi XLindley (PQXL) distribution [8] and discrete Marshal Olkin inverted Topp-Leone (DMOITL) distribution [12]. The comparison is carried out using log-likelihood ($\hat{\ell}$) and some information criteria, including Akaike Information Criterion (AIC), corrected Akaike Information Criterion (AICc), and Bayesian Information Criterion (BIC). Also, chi-square (χ^2) goodness-of-fit test is used for the comparison. The model with the least value of information criteria is considered the best model as well as the model with the highest p-value of the χ^2 goodness-of-fit test. The pmfs of the PEBWE, PJ, PQXL and DMOITL distributions are given as

$$\mathbb{P}_{PEBWE}(X = x) = \frac{\beta((1+\beta)\ln(\beta) - (1+x)\beta)}{(1+\beta)^{2+x}(\ln(\beta) - 1)}, x = 0, 1, 2, \ldots, \beta > 0,$$

$$\mathbb{P}_{PJ}(X = x) = \left(\frac{\theta}{\theta+\alpha}\right)^2 \left(\frac{\alpha}{\theta+\alpha}\right)^x \left[1 + \frac{\alpha(1+\alpha x)}{\theta+\alpha^2}\right], x = 0, 1, 2, \ldots, \alpha > 0, \theta > 0,$$

$$\mathbb{P}_{PQXL}(X = x) = \frac{\eta(\alpha(1+\eta)^2 + \eta(2+x+\eta))}{(1+\alpha)(1+\eta)^{x+3}}, x = 0, 1, 2, \ldots, \alpha > -1, \eta > 0$$

and

$$\mathbb{P}_{DMOITL}(X = x) = \frac{\alpha(1 + 2x)^{\theta}}{(1 + x)^{2\theta} - \bar{\alpha}(1 + 2x)^{\theta}} - \frac{\alpha(3 + 2x)^{\theta}}{(2 + x)^{2\theta} - \bar{\alpha}(3 + 2x)^{\theta}}, x = 0, 1, 2, \ldots.$$

3.1. Dataset I: Mammalian Cytogenetic Dosimetry Lesions (Exposure -60 μg | kg)

The first data used to establish the utility of the PTPCJ distribution is the mammalian cytogenetic dosimetry lesions in rabbit lymphoblast induced by streptonigrin (NSC-45383), Exposure -60 μg | kg. The data is studied by [32], [36] and [35]. The mean, DI, CS and excess kurtosis of the data are respectively given as 0.4742, 1.5600, 2.4098 and 7.6598. Thus the data is over-dispersed, positively skewed and leptokurtic.

The ML estimates, with their corresponding standard errors (SE), ℓ, AIC, AICc, BIC and χ^2 test values of the fitted models are highlighted in Table 1. It is revealed that the proposed distribution outperforms the other competing models since it displays the least AIC, AICc, BIC values and highest ℓ and p-value of the χ^2 test

Table 1. ML estimates, ℓ, AIC, AICc, BIC and χ^2 values for dataset I

		Expected Frequencies					
X	**Observed**	**PTPCJ**	**Poisson**	**PEBWE**	**PJ**	**PQXL**	**DMOITL**
0	413	409.8706	374.0484	148.9031	407.6760	407.6761	410.7554
1	124	128.4790	177.3774	158.4473	131.1374	131.1373	133.7505
2	42	41.8199	42.0570	116.8015	42.1830	42.1830	35.0761
3	15	13.8822	6.6480	75.1514	13.5690	13.5690	11.7207
4	5	4.6384	0.7881	45.0120	4.3647	4.3648	4.7184
5	0	1.5473	0.0747	25.7918	1.4040	1.4040	2.1754
6	2	0.7627	0.0063	30.8931	0.6658	0.6658	2.8035
Total	601	601	601	601	601	601	601
ML Estimate (SE)		ς=5.4586	λ=0.4742	β=1.1085	α=0.0024	α=999.3720	α=1.0291
		(14.1658)	(0.0281)	(0.0385)	(1.42×10^{-4})	(1.03×10^{-7})	(0.2050)
		φ=2.6253			θ=0.0051	η=2.1095	θ=4.0669
		(1.0181)			(6.71×10^{-5})	(0.1517)	(0.3985)
$\hat{\ell}$		-556.3000	-582.6800	-861.9900	-556.5200	-556.5200	-558.7200
AIC		1116.6010	1167.3550	1725.9780	1117.0330	1117.0320	1121.4380
AICc		1116.6210	1167.3620	1725.9850	1117.0530	1117.0530	1121.4580
BIC		1125.3980	1171.7540	1730.3760	1125.8300	1125.8300	1130.2350
χ^2		0.2712	48.022	660.33	0.6594	0.65937	3.7575
df		2	2	5	2	2	2
p-value		0.8732	<0.0001	<0.0001	0.7191	0.7192	0.1528

The pmf plots of the empirical and fitted distributions for dataset I are highlighted in Figure 3. It is revealed that the pmf of PTPCJ distribution closely mimics the empirical pmf, thus providing an optimal fit to dataset I.

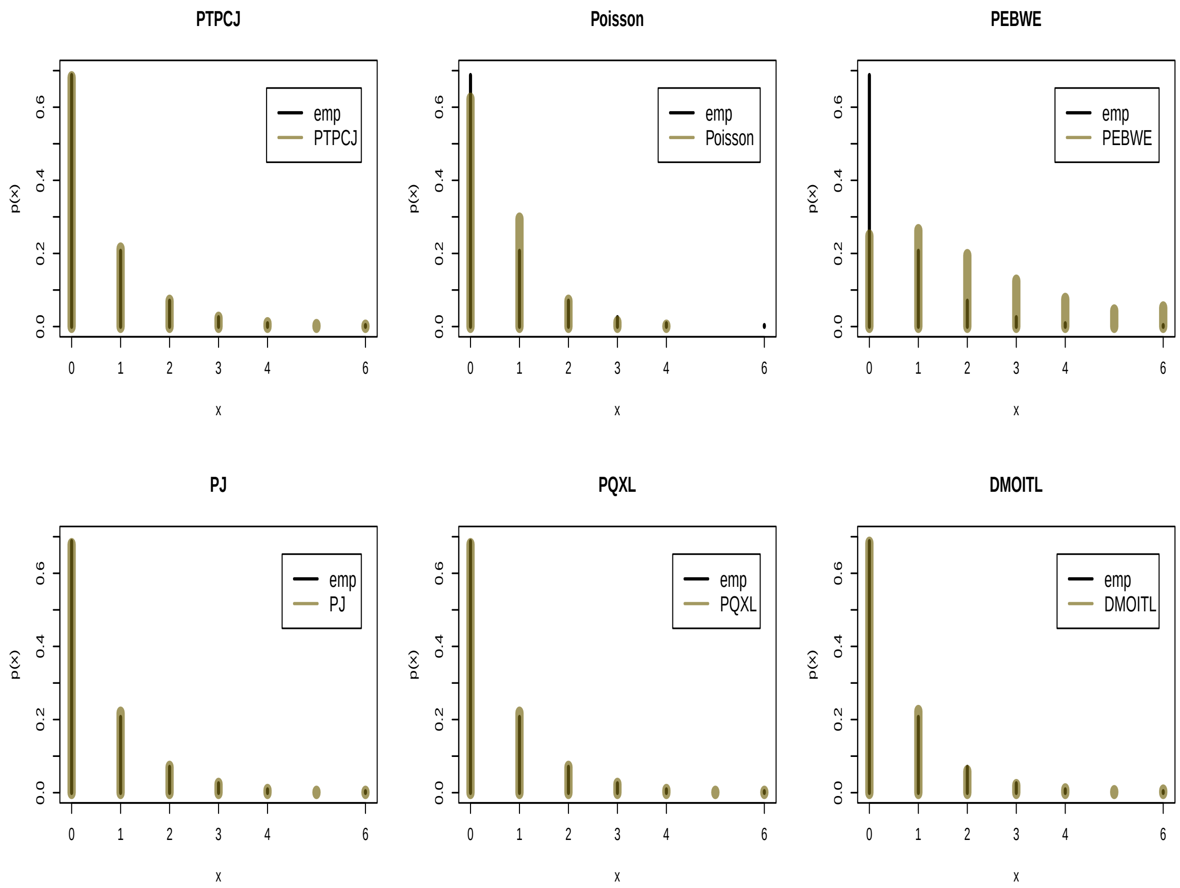

Figure 3. pmf plots of the empirical and fitted models for dataset I

3.2. Dataset II: Mammalian Cytogenetic Dosimetry Lesions (Exposure -70 μg | kg)

The second data used to establish the utility of the PTPCJ distribution is the mammalian cytogenetic dosimetry lesions in rabbit lymphoblast induced by streptonigrin (NSC-45383), Exposure -70 μg | kg. The data is studied by [36], [24] and [2]. The mean, DI, CS and excess kurtosis of the second data are respectively given as 0.5533, 1.7175, 2.3441 and 7.0685. The second data is over-dispersed, positively skewed and leptokurtic.

The ML estimates with their SE, ℓ, AIC, AICc, BIC and χ^2 values of the fitted distributions for dataset II are display in Table 2. It is evident from Table 2 that the PTPCJ distribution performs better than the competing models since it displays the least AIC, AICc and BIC values. It also has the highest ℓ value and p-value of the χ^2 test.

In Figure 4, we present the pmf plots of the empirical and fitted models for dataset II. It is seen from Figure 4 that the pmf of the PTPCJ distribution closely mimics the empirical pmf, thus it provides an optimal fit to the dataset

Table 2. ML estimates, ℓ, AIC, AICc, BIC and χ^2 values for dataset II

X	Observed	Expected Frequencies					
		PTPCJ	**Poisson**	**PEBWE**	**PJ**	**PQXL**	**DMOITL**
0	200	194.4326343	172.5090	74.3327	193.1330	193.1330	197.5178
1	57	67.3113	95.4550	79.0763	68.7985	68.7985	67.8629
2	30	24.2113	26.4092	58.2957	24.5076	24.5076	19.8936
3	7	8.8779	4.8710	37.5130	8.7302	8.7302	7.3784
4	4	3.2742	0.6738	22.4720	3.1099	3.1099	3.2496
5	0	1.2046	0.0746	12.8786	1.1078	1.1078	1.6189
6	2	0.6880	0.0074	15.4317	0.6130	0.6130	2.4787
Total	300	300	300	300	300	300	300
ML Estimate (SE)		ς=5.5244	λ=0.5533	β=1.1081	α=0.0011	α=999.9606	α=0.9242
		(11.9560)	(0.0429)	(0.0555)	(7.92×10^{-5})	(1.12×10^{-7})	(0.2485)
		φ=2.2990			θ=0.0019	η=1.8079	θ=3.5556
		(0.8020)			(4.37×10^{-5})	(0.1748)	(0.4888)
$\hat{\ell}$		-303.1900	-323.4500	-436.5900	-303.4700	-303.4700	-305.4000
AIC		610.3866	648.8907	875.1714	610.9316	610.9316	614.8056
AICc		610.4270	648.9042	875.1848	610.9720	610.9720	614.8460
BIC		617.7942	652.5945	878.8752	618.3391	618.3391	622.2132
χ^2		3.6546	30.0230	296.9200	3.5216	3.5216	7.1708
df		2	2	5	1	1	2
p-value		0.1608	<0.0001	<0.0001	0.0606	0.0606	0.0277

3.3. Dataset III: Accidents of 647 Women

The last data used to establish the utility of the proposed model is the accidents data of 647 women working on high explosive shells in 5 weeks. The data is studied by [31], [22] and [6]. The mean, DI, CS and excess kurtosis of the third dataset are respectively given by 0.4652, 1.4872, 2.1163 and 4.8964. This implies that the data is over-dispersed, positively skewed and leptokurtic.

Table 3 highlights the ML estimates, SE, AIC, AICc, BIC and χ^2 test values of the fitted models for dataset III. It is seen that the PTPCJ distribution produce the least AIC, AICc and BIC values as well as the highest ℓ and p-values of the χ^2 test. Thus the PTPCJ distribution provides the best fit to dataset III compared to the other competing models.

In Figure 5, we present the pmf plots of the fitted models for dataset III. It is reveal that the pmf of the proposed distribution closely mimics the empirical pmf, thus it provides an optimal fit to dataset III.

4. PTPCJ Regression model for Multivariate Applications

In this section, we introduce the PTPCJ count regression model. It can be used to study the relationship between a response variable and a set of independent variables. The Poisson regression model is the classical regression model for analysing discrete data. However, when the response variable is not equi-dispersed, the Poisson regression fails to be the best regression model for modeling such data. Suppose that Y follows

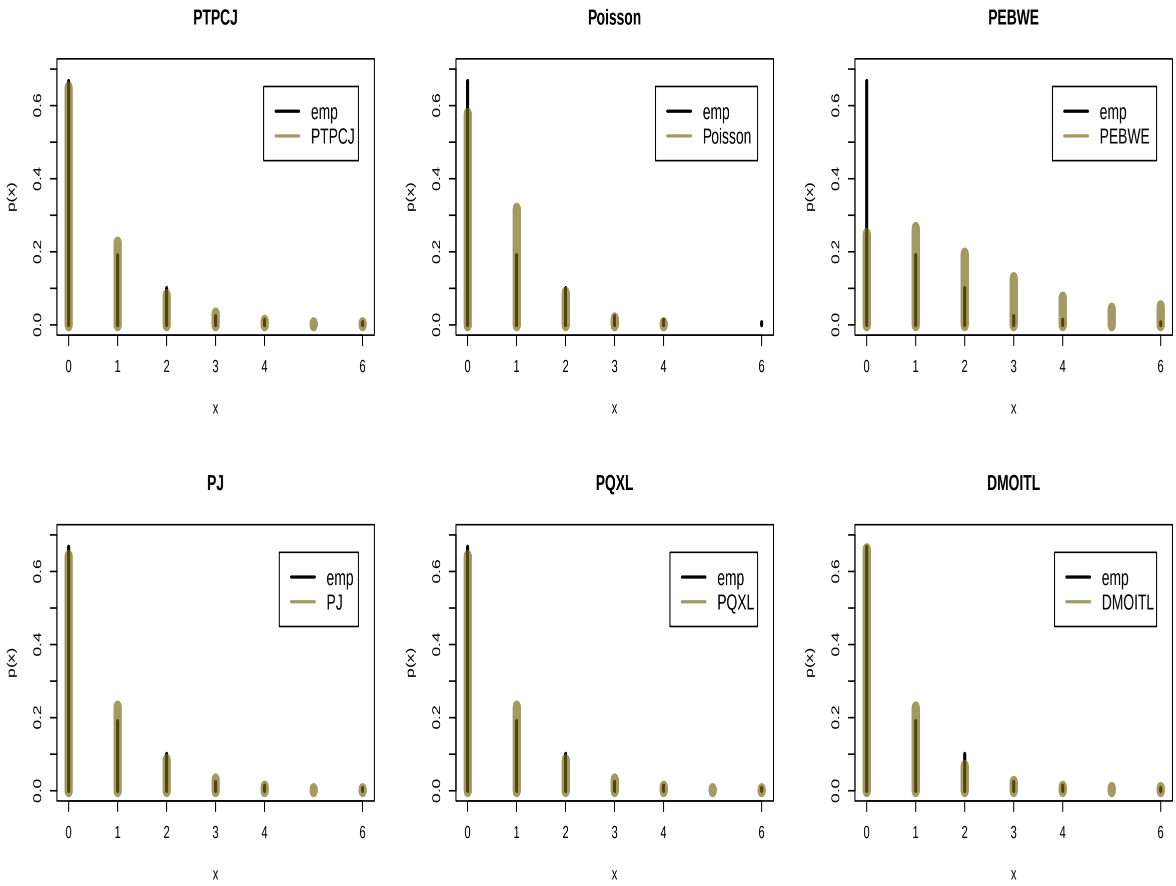

Figure 4. pmf plots of the empirical and fitted models for dataset II

the PTPCJ distribution with mean

$$\mu = \frac{\varsigma\varphi + 6}{(\varsigma\varphi + 2)\varphi}. \tag{4.1}$$

Making ς the subject in equation (4.1), we obtain $\varsigma = \frac{2\mu\varphi - 6}{\varphi(1-\mu\varphi)}$. Substituting $\varsigma = \frac{2\mu\varphi - 6}{\varphi(1-\mu\varphi)}$ into the pmf of the PTPCJ distribution produce a re-parameterized pmf given as

$$\mathbb{P}(Y = y) = \frac{\varphi^2 \left(\frac{(2\mu\varphi-6)(\varphi+1)^2}{\varphi(1-\mu\varphi)} + \varphi\pi(y)\right)}{\left(\frac{2\mu\varphi-6}{(1-\mu\varphi)} + 2\right)(\varphi + 1)^{y+3}}, y = 0, 1, 2, \ldots, \mu > 0, \varphi > 0.$$

The logarithm link function $\log(\mu_i) = \mathbf{X}_i^T\mathbf{\Omega}$ is used in this study to link the independent variables to the response variable, where $\mathbf{X}_i^T = (x_{1i}, x_{2i}, x_{3i}, \ldots x_{ki})$ is the vector of independent variables and $\mathbf{\Omega} = (\omega_0, \omega_1, \omega_2, \ldots, \omega_k)$ is the vector of regression coefficients. Thus $\mu_i = e^{\mathbf{X}_i^T\mathbf{\Omega}}$. The log-likelihood function of the PTPCJ regression model is given as

Table 3. ML estimates, ℓ, AIC, AICc, BIC and χ^2 values for dataset III

		Expected Frequencies					
X	**Observed**	**PTPCJ**	**Poisson**	**PEBWE**	**PJ**	**PQXL**	**DMOITL**
0	447	443.0252	406.3123	160.307	441.5706	441.5707	444.3621
1	132	138.0255	189.0263	170.5532	140.2035	140.2034	142.9739
2	42	44.6747	43.9698	125.73048	44.5161	44.5160	37.1826
3	21	14.5095	6.8186	80.90324	14.1343	14.1343	12.3519
4	3	4.6528	0.7930	48.4619	4.4878	4.4878	4.9496
5	2	2.1122	0.0799	61.0442	2.0878	2.0878	5.1798
Total	647	647	647	647	647	647	647
ML Estimate (SE)		ς=2.4276	λ=0.4652	β=1.1082	α=0.0036	α=997.6623	α=1.0236
		(3.6830)	(0.0268)	(0.0370)	(2.05×10^{-4})	(1.03×10^{-7})	(0.1972)
		φ=3.0612			θ=0.0077	η=2.1502	θ=4.0913
		(0.9032)			(9.50×10^{-5})	(0.1500)	(0.3870)
$\hat{\ell}$		-592.2100	-617.1800	-926.2100	-592.4800	-592.4800	-596.2400
AIC		1188.4260	1236.3690	1854.4230	1188.9600	1188.9600	1196.4860
AICc		1188.4450	1236.3750	1854.4300	1188.9780	1188.9780	1196.5040
BIC		1197.3710	1240.8410	1858.8960	1197.9040	1197.9040	1205.4300
χ^2		3.8227	64.9460	721.3100	4.4015	4.4015	10.1340
df		2	2	4	2	2	2
p-value		0.1479	<0.0001	<0.0001	0.1107	0.1107	0.0063

$$\ell(\varphi,\mathbf{\Omega}) = 2n\log[\varphi] + \sum_{i=1}^{n}\log\left[\frac{(2\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}}-6)(\varphi+1)^2}{\varphi(1-\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}})}+\varphi\pi(y_i)\right] - \sum_{i=1}^{n}\log\left[\frac{2\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}}-6}{(1-\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}})}+2\right] - \sum_{i=1}^{n}(y_i+3)\log[\varphi+1]. \tag{4.2}$$

Differentiating equation (4.2) with respect φ and the vector of the regression coefficients, $\mathbf{\Omega}$, gives the score functions as

$$\frac{\partial\ell(\varphi,\mathbf{\Omega})}{\partial\varphi} = \frac{2n}{\varphi} - \sum_{i=1}^{n}\frac{e^{\mathbf{X}_i^T\mathbf{\Omega}}}{1-\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}}} - \sum_{i=1}^{n}\frac{(y_i+3)}{(\varphi+1)} + \sum_{i=1}^{n}\frac{\left[\begin{array}{l}2\varphi(\varphi+1)(1-\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}})(3\varphi+1)e^{\mathbf{X}_i^T\mathbf{\Omega}}\\ \quad-(\varphi+1)^2(1-2\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}})(2\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}}-6)+\left[\varphi(1-\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}})\right]^2\pi(y_i)\end{array}\right]}{\left[\varphi(1-\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}})\right]\left[(2\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}}-6)(\varphi+1)^2+\varphi(1-\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}})\pi(y_i)\right]}$$

and

$$\frac{\partial\ell(\varphi,\mathbf{\Omega})}{\partial\mathbf{\Omega}} = \sum_{i=1}^{n}\frac{4\varphi(1+\varphi)^2\mathbf{X}_i^T e^{\mathbf{X}_i^T\mathbf{\Omega}}}{2(1+\varphi)^2(\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}}-3)+\varphi^2(1-\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}})\pi(y_i)} - \sum_{i=1}^{n}\frac{\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}}}{1-\varphi e^{\mathbf{X}_i^T\mathbf{\Omega}}}.$$

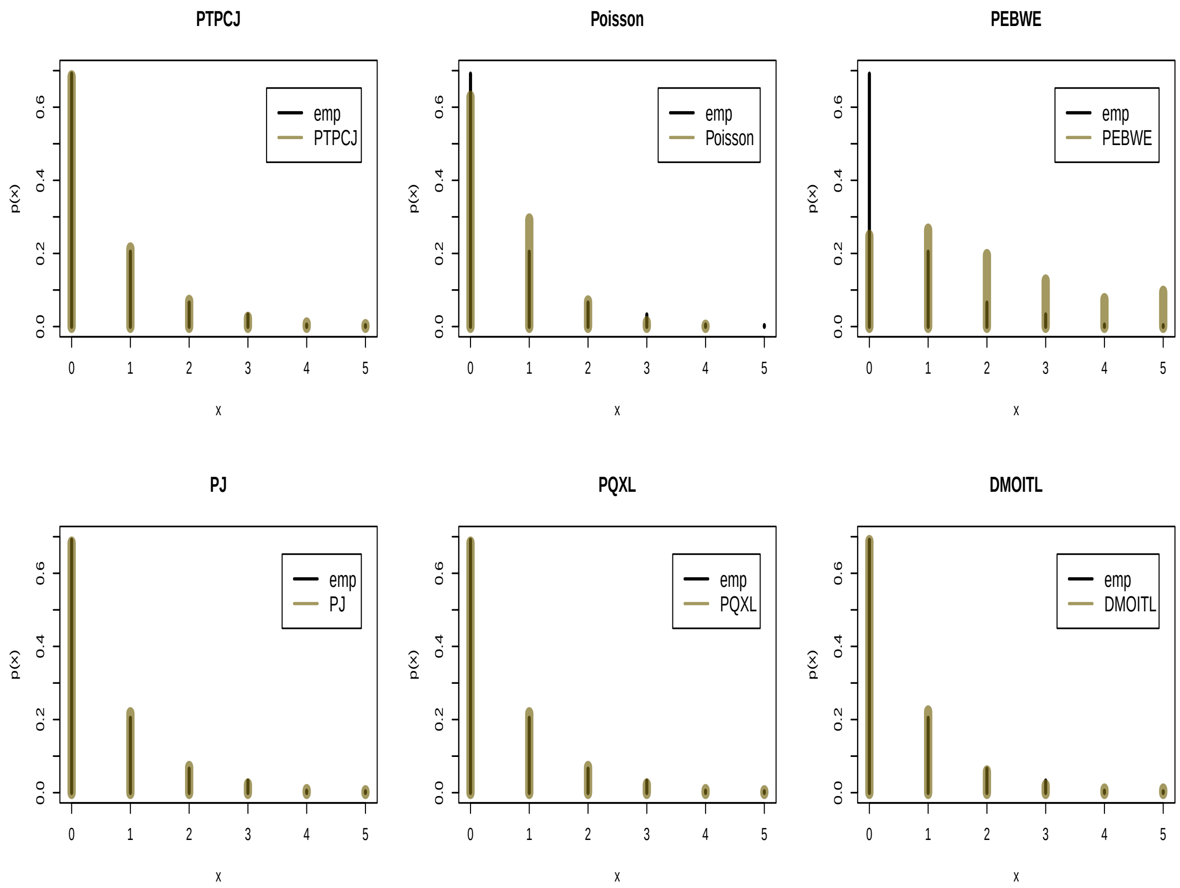

Figure 5. pmf plots of the empirical and fitted models for dataset III

Equating each function to zero and solving them simultaneously will give the estimators of the parameters. However, due to the non-linear nature of the score functions, the log-likelihood function in equation (4.2) is optimized directly using R statistical software [21].

5. Applications of the PTPCJ Regression Model

In this section, we present the applications of the PTPCJ regression model using count data with 189 observations called low birth weight data (lbw) obtain from the COUNT package [25] in R software. In this study, we consider number of physician visits as the response variable (y_i) and whether or not mother smokes (history of mother smoking = 1, mother non-smoker = 0) (x_{i1}) as the explanatory variable. The PTPCJ regression model is compared with th Poisson regression model, Poisson extended exponential (PEE) regression model [28] and new Poisson generalized Lindley (PGL) regression model [16] based on their ℓ, AIC, AICc and BIC values. The regression structure used for this application is given as

$$\mu_i = \exp(\omega_0 + \omega_1 x_{i1}), i = 1, \dots, n. \tag{5.1}$$

Substituting equation (5.1) into the following pmfs give the regression models for the PEE and PGL regression models:

$$f(y) = \frac{\alpha\left(1+\alpha+\left[\frac{\alpha-\alpha^2\mu}{\alpha\mu-2}\right]+\left[\frac{\alpha-\alpha^2\mu}{\alpha\mu-2}\right]y\right)}{\left(\alpha+\left[\frac{\alpha-\alpha^2\mu}{\alpha\mu-2}\right]\right)(\alpha+1)^{y+2}}, \alpha>0, \mu>0$$

and

$$f(y) = \frac{\theta^{\frac{\theta(\theta\mu+\mu+1)}{\theta+1}}\Gamma\left(\frac{\theta(\theta\mu+\mu+1)}{\theta+1}+y-1\right)\left(\frac{\theta(\theta\mu+\mu+1)}{\theta+1}(\theta+2)-\theta+y-2\right)}{\Gamma\left(\frac{\theta(\theta\mu+\mu+1)}{\theta+1}\right)\Gamma(y+1)(\theta+1)^{\frac{\theta(\theta\mu+\mu+1)}{\theta+1}+y+1}}, \theta>0, \mu>0.$$

Also, substituting (5.1) into the pmf of Poisson in equation (2.1) for λ gives the Poisson regression model.

Table 4 displays the ML estimates, SE, ℓ, AIC, AICc and BIC values of the fitted regression models. It is observed that the PTPCJ regression model outperforms the competing regression models since the proposed regression model has the highest log-likelihood value and the least AIC, AICc and BIC values.

Table 4. ML estimates, ℓ, AIC, AICc and BIC of the fitted regression models

Regression Model	ML estimates(SE)	ℓ	AIC	AICc	BIC
PTPCJ	ω_0 =-0.1483(0.1132)	-230.4100	466.8103	466.9401	476.5356
	ω_1 =-0.2270(0.1819)				
	φ =3.1309(0.6133)				
Poisson	β_0 =-0.2016(0.1031)	-237.1900	478.3801	478.4446	484.8636
	β_1 =-0.0771(0.1688)				
PEE	ω_0 =-0.1492(0.1141)	-230.6100	467.2153	467.3450	476.9405
	ω_1 =-0.2242(0.1828)				
	α =2.4072(0.5075)				
PGL	β_0 =-0.1594(0.1220)	-230.9400	467.8868	468.0165	477.6120
	β_1 =-0.1934(0.2054)				
	θ =2.3408(0.8291)				

6. Conclusion

In real-world situations, discrete data is mostly over-dispersed. However, the Poisson distribution is only adequate for modeling equi-dispersed data. Therefore, to obtain better fit for most real-world data, several new count distributions are continuously being developed to adequately explain over-dispersed count data. For this reason, the PTPCJ distribution was developed by [9] for analyzing over-dispersed discrete data. We demonstrated the univariate and multivariate applications of the PTPCJ distribution in this study using three datasets and it was revealed that the PTPCJ distribution outperforms some competing models. The study introduced the PTPCJ regression model for studying the relationship between a response variable and explanatory variables. The applications of the regression model was established using a real dataset and it was evident that the proposed regression model performs better than the Poisson, Poisson extended exponential and new Poisson generalized Lindley regression models. The study recommends that in future integer auto-regressive process for modeling count time series data can be explored where the innovation follows the PTPCJ distribution. Finally, the study recommends a bivariate extension of the new distribution.

Conflict of Interest

The authors declare there is no existing conflict of interest.

References

1.Adetunji, A. and Sabri, S. R. M. (2023). A new three-parameter mixed Poisson transmuted weighted exponential distribution with applications to insurance data. *Science and Technology Indonesia*, 8(2):235–244.

2.Agarwal, A. and Pandey, H. (2023). An inflated model based on sized biased Himanshu distribution and its applications. *Journal of Probability and Statistical Science*, 21(1):1–7.

3.Ahmad, P. B. and Wani, M. K. (2023). A new compound distribution and its applications in over-dispersed count data. *Annals of Data Science*, 14:1799–1820.

4.Ahsan-ul Haq, M. (2024). On Poisson moment exponential distribution with applications. *Annals of Data Science*, 11(1):137–158.

5.Aijaz, A., Ain, S., Afaq, A., and Tripathi, R. (2022). Poisson area-biased Ailamujia distribution and its applications in environmental and medical sciences. *Statistics in Transition new series*, 23(3):167–184.

6.Alaa, R. and Eledum, H. (2022). Discrete extended Erlang-truncated exponential distribution and its applications. *Applied Mathematics & Information Sciences*, 16(1):127–138.

7.Alghamdi, F. M., Ahsan-ul Haq, M., Hussain, M. N. S., Hussam, E., Almetwally, E. M., Aljohani, H. M., Mustafa, M. S., Alshawarbeh, E., and Yusuf, M. (2024a). Discrete poisson quasi-xlindley distribution with mathematical properties, regression model, and data analysis. *Journal of Radiation Research and Applied Sciences*, 17(2):100874.

8.Alghamdi, F. M., Ahsan-ul Haq, M., Hussain, M. N. S., Hussam, E., Almetwally, E. M., Aljohani, H. M., Mustafa, M. S., Alshawarbeh, E., and Yusuf, M. (2024b). Discrete Poisson quasi-XLindley distribution with mathematical properties, regression model, and data analysis. *Journal of Radiation Research and Applied Sciences*, 17(2):100874.

9.Alharti, M. F. and Alzubaidi, S. (2025). A novel statistical model with applications on medical and health real data. *Alexandria Engineering Journal*, 125:42–54.

10.Alkhairy, I. (2023). Classical and Bayesian inference for the discrete Poisson Ramos-Louzada distribution with application to covid-19 data. *Mathematical Biosciences and Engineering*, 20(8):14061–14080.

11.Almetwally, E. M., Abdo, D. A., Hafez, E., Jawa, T. M., Sayed-Ahmed, N., and Almongy, H. M. (2022a). The new discrete distribution with application to covid-19 data. *Results in Physics*, 32:104987.

12.Almetwally, E. M., Abdo, D. A., Hafez, E., Jawa, T. M., Sayed-Ahmed, N., and Almongy, H. M. (2022b). The new discrete distribution with application to covid-19 data. *Results in Physics*, 32:104987.

13.Almetwally, E. M., Al-Essa, L. A., Imran, M., Jamal, F., Chesneau, C., and Abdelkawy, M. A. (2025). A novel flexible three-parameter zero-inflated discrete distribution with applications in medical-related fields. *Journal of Radiation Research and Applied Sciences*, 18(2):101515.

14.Alomair, A. and Ahsan-ul Haq, M. (2023). A new extension of Poisson distribution for asymmetric count data: theory, classical and Bayesian estimation with application to lifetime data. *PeerJ Computer Science*, 9:e1748.

15.Alrumayh, A. and Khogeer, H. A. (2023). A new two-parameter discrete distribution for overdispersed and asymmetric data: Its properties, estimation, regression model, and applications. *Symmetry*, 15(6):1289.

16.Atikankul, Y. (2023). A new Poisson generalized Lindley regression model. *Austrian Journal of Statistics*, 52(1):39–50.

17.Bodhisuwan, W. and Aryuyuen, S. (2022). The Poisson-transmuted Janardan distribution for modelling count data. *Trends in Sciences*, 19(5):2898–2898.

18.Chesneau, C., D'cruz, V., Maya, R., and Irshad, M. (2023). A novel discrete distribution based on the mixture of Poisson and sum of two lindley random variables. *REVSTAT-Statistical Journal*, 23(3):371–396.

19.Chesneau, C., Tomy, L., and Veena, G. (2022). The Poisson-modified Lindley distribution. *Applied Mathematics E-Notes*, 22:18–31.

20.Chinedu, E. Q., Chukwudum, Q. C., Alsadat, N., Obulezi, O. J., Almetwally, E. M., and Tolba, A. H. (2023). New lifetime distribution with applications to single acceptance sampling plan and scenarios of increasing hazard rates. *Symmetry*, 15(10):1881.

21.Core Team, R. (2015). R: A language and environment for statistical computing. r foundation for statistical computing, vienna, austria.

22.Eledum, H. and El-Alosey, A. (2023). Discrete Kumaraswamy Erlang-truncated exponential distribution with applications to count data. *Journal of Statistics Applications & Probability*, 12(2):725–739.

23.Erbayram, T. and Akdoğan, Y. (2023). A new discrete model generated from mixed poisson transmuted record type exponential distribution. *Ricerche di Matematica*, 74:1225–1247.

24.Hassan, A., Shalbaf, G. A., Bilal, S., and Rashid, A. (2020). A new flexible discrete distribution with applications to count data. *Journal of Statistical Theory and Applications*, 19(1):102–108.

25.Hilbe, J. M. (2016). *COUNT: Functions, Data and Code for Count Data*. R package version 1.3.4.

26.Irshad, M. R., Aswathy, S., Maya, R., and Nadarajah, S. (2023). New one-parameter over-dispersed discrete distribution and its application to the nonnegative integer-valued autoregressive model of order one. *Mathematics*, 12(1):81.

27.Karakaya, K. (2023). A new discrete distribution with applications to radiation, smoking, and health data. *Journal of Radiation Research and Applied Sciences*, 16(4):100735.

28.Maya, R., Chesneau, C., Krishna, A., and Irshad, M. R. (2022a). Poisson extended exponential distribution with associated INAR (1) process and applications. *Stats*, 5(3):755–772.

29.Maya, R., Irshad, M. R., Chesneau, C., Nitin, S. L., and Shibu, D. S. (2022b). On discrete Poisson–Mirra distribution: Regression, inar (1) process and applications. *Axioms*, 11(5):193.

30.Nagy, M., Almetwally, E. M., Gemeay, A. M., Mohammed, H. S., Jawa, T. M., Sayed-Ahmed, N., and Muse, A. H. (2021). The new novel discrete distribution with application on covid-19 mortality numbers in kingdom of saudi arabia and latvia. *Complexity*, 2021(1):7192833.

31.Para, B. A., Jan, T. R., and Bakouch, H. S. (2020). Poisson Xgamma distribution: A discrete model for count data analysis. *Model Assisted Statistics and Applications*, 15(2):139–151.

32.Rashid, A. (2016). A new count data model with application in genetics and ecology. *Electronic Journal of Applied Statistical Analysis*, 9(1):213–226.

33.Sah, B. K. and Sahani, S. K. (2023). Poisson-modified Mishra distribution. *Jilin Daxue Xuebao (Gongxueban) Journal of Jilin University (Engineering and Technology Edition)*, 42(1):261–275.

34.Seghier, F., Ahsan-ul Haq, M., Zeghdoudi, H., and Hashmi, S. (2023). A new generalization of poisson distribution for over-dispersed, count data: Mathematical properties, regression model and applications. *Lobachevskii Journal of Mathematics*, 44(9):3850–3859.

35.Shanker, R. and Fesshaye, H. (2016). On Poisson-Sujatha distribution and its applications to model count data from biological sciences. *Biometrics & Biostatistics International Journal*, 3(4):1–7.

36.Shanker, R., Fesshaye, H., Shanker, R., Leonida, T. A., and Sium, S. (2017). On discrete Poisson-shanker distribution and its applications. *Biometrics & biostatistics International journal*, 5(1):00121.

37.Shanker, R., Sharma, S., Shanker, U., Shanker, R., and Leonida, T. A. (2014). The discrete Poisson–Janardan distribution with applications. *International Journal of Soft Computing and Engineering*, 4(2):31–33.

38.Shanker, R. and Shukla, K. K. (2024). A new quasi Poisson-Sujatha distribution with properties and applications. *AIJR Abstracts*, page 86.

39.Shanker, R. and Tabbasum, R. (2024). A quasi Poisson-Garima distribution with properties and applications. *AIJR Abstracts*, page 85.

40.Shukla, K. K., Shanker, R., Tiwari, M. K., and Ababneh, F. (2023). Size-biased poisson-pranav distribution and its applications. *International Journal of Agricultural & Statistical Sciences*, 19(2):863–872.

41.Tharshan, R. and Wijekoon, P. (2024). Poisson-modification of quasi Lindley regression model for over-dispersed count responses. *Communications in Statistics-Simulation and Computation*, 53(3):1258–1273.