Pertanika J. Sci. & Technol. 31 (3): 1587 - 1603 (2023)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

An Alternative Count Distribution for Modeling Dispersed Observations

Ademola Abiodun Adetunji[1,2] and Shamsul Rijal Muhammad Sabri[1]*

[1]*School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia*
[2]*Department of Statistics, Federal Polytechnic, Ile-Oluji, Ondo 351, Nigeria*

ABSTRACT

In most cases, count data have higher variances than means; hence using the Poisson distribution to model such observations is misleading because of the equality of the Poisson mean and variance. This study proposes a new two-parameter mixed Poisson distribution for modeling dispersed count observations. The exponential distribution is transmuted to obtain a new mixing distribution for the new proposition. Different moment-based mathematical properties of the new proposition are obtained. Applications using dispersed count observations with excess zero are made. Comparisons with related distributions for modeling dispersed observation reveal that the new distribution performs creditably well.

Keywords: Count observations, cubic rank transmutation, mixed Poisson distribution

INTRODUCTION

The classical Poisson distribution is widely utilized for modeling count data. However, since the distribution assumes equi-dispersion (variance equals mean), its general applicability is limited (Ong et al., 2021). There is model misspecification if the Poisson distribution is assumed for heterogeneous data (Asamoah, 2016) since count observations are often dispersed (Adetunji & Sabri, 2021; Omari et al., 2018). Several techniques have been developed in the literature to model over-dispersed observations (Bhati et al., 2015; Das et al., 2018; Sankaran, 1970; Zakerzadeh & Dolati, 2009). Among these diverse techniques is the mixed Poisson distribution, first proposed in the early 20th century (Greenwood & Yule, 1920). The technique is based on the assumption

ARTICLE INFO

Article history:
Received: 02 July 2022
Accepted: 05 October 2022
Published: 07 April 2023

DOI: https://doi.org/10.47836/pjst.31.3.25

E-mail addresess:
adecap4u@gmail.com (Ademola Abiodun Adetunji)
rijal@usm.my (Shamsul Rijal Muhammad Sabri)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

that there is an inherent fluctuation in the parameter of the Poisson distribution, and hence a continuous distribution is assumed for the parameter. To this end, several mixing distributions have been introduced to obtain mixed Poisson distributions. In most cases, the mixing distributions arise from the exponential distribution paradigm. Karlis and Xekalaki (2005) provide a detailed survey on the properties of mixed Poisson distribution, while Ong et al. (2021) provide recent research on the distribution.

The classical exponential distribution is suitable when the system for its application assumes a constant failure rate. This limitation has motivated researchers to improve its flexibility to model systems with diverse hazard rates. It is often achieved by adding an extra shape parameter (compounding the classical form) using different generalization techniques (Karina et al., 2019). Applications of compounded exponential distributions in reliability theory pervade almost all works of life, including but not limited to economic, reliability, environmental, industrial, and engineering spaces (Aguilar et al., 2019; Mohammed et al., 2015; Rasekhi et al., 2017).

Mixture distributions offer flexibility in modeling observations (Karlis & Xekalaki, 2005). This study uses the one-parameter cubic rank transmutation map (Al-kadim, 2018) to compound the exponential distribution to obtain a new mixing distribution. The obtained mixing distribution is then used to obtain a new mixed Poisson distribution. Karlis and Xekalaki (2005) provide some properties of mixed Poisson distributions.

MATERIALS AND METHOD

Cubic Rank Transmuted Exponential Distribution

The distribution function of a continuous random variable *X* that follows an exponential distribution with scale parameter α is given as in Equation 1:

$$G_x = 1 - e^{-\alpha x}, \qquad \alpha > 0 \qquad [1]$$

Given G_x as in Equation 1, the distribution function of the Cubic Rank Transmutation (CRT) map due to Al-kadim (2018) with an extra shape parameter β is given as in Equation 2:

$$F(x) = (1+\beta)G_x - 2\beta G_x^2 + \beta G_x^3, \qquad -1 \le \beta \le 1 \qquad [2]$$

Hence, the Cubic Rank Transmuted Exponential Distribution (CRTED) is obtained by inserting Equation 1 into Equation 2. Therefore, the CDF and PDF of the CRTED using the transmutation map of Al-kadim (2018) are obtained as in Equations 3 and 4:

$$F_x = 1 - e^{-\alpha x} + \beta e^{-2\alpha x} - \beta e^{-3\alpha x}, \qquad -1 \le \beta \le 1; \alpha > 0 \qquad [3]$$

$$f_x = \alpha e^{-\alpha x}(1 - 2\beta e^{-\alpha x} + 3\beta e^{-2\alpha x}), \qquad -1 \le \beta \le 1; \alpha > 0 \tag{4}$$

When the extra parameter assumes zero ($\beta = 0$), the CRTED becomes the exponential distribution. The corresponding survival and hazard rate functions (HRF) are obtained as in Equations 5 and 6:

$$S_x = e^{-\alpha x} - \beta e^{-2\alpha x} + \beta e^{-3\alpha x} \tag{5}$$

$$h_x = \frac{\alpha(1 - 2\beta e^{-\alpha x} + 3\beta e^{-2\alpha x})}{1 - \beta e^{-\alpha x} + \beta e^{-2\alpha x}} \tag{6}$$

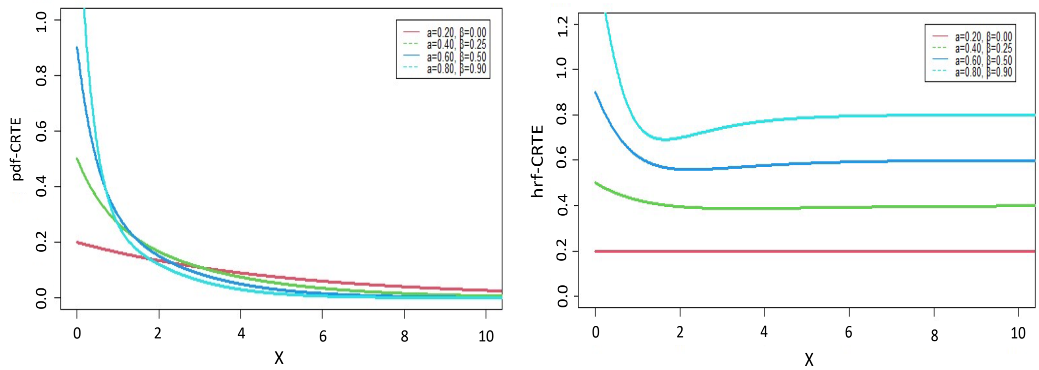

Figure 1. Shapes of the PDF and HRF of the CRTED

Figure 1 shows that the PDF of the CRTED is monotonically decreasing for different parameter combinations. Except when the CRTED becomes the exponential distribution ($\beta = 0$), the HRF plot has an inverted J-shape.

Mathematical Properties of the CTRED

Proposition 1: If a random variable X has a CRTED, the rth moment is defined as in Equation 7:

$$E(x^r) = \left(1 - \frac{\beta}{2^r} + \frac{\beta}{3^r}\right)\frac{r!}{\alpha^r} \tag{7}$$

Proof:

$$E(x^r) = \int_0^\infty x^r f_x \, dx$$

$$= \int_0^\infty x^r \left(\alpha e^{-\alpha x} - 2\alpha\beta e^{-2\alpha x} + 3\alpha\beta e^{-3\alpha x}\right) dx$$
$$= \alpha \int_0^\infty x^r e^{-\alpha x}\, dx - 2\alpha\beta \int_0^\infty x^r e^{-2\alpha x}\, dx + 3\alpha\beta \int_0^\infty x^r e^{-3\alpha x}\, dx$$
$$= \frac{r!}{\alpha^r} - \frac{\beta r!}{(2\alpha)^r} + \frac{\beta r!}{(3\alpha)^r}$$
$$= \left(1 - \frac{\beta}{2^r} + \frac{\beta}{3^r}\right)\frac{r!}{\alpha^r}$$

Proposition 2: If a random variable *X* has a CRTED, the Moment Generating Function (MGF) is obtained as in Equation 8:

$$E(e^{tx}) = \frac{\alpha}{\alpha - t} - \frac{2\alpha\beta}{2\alpha - t} + \frac{3\alpha\beta}{3\alpha - t} \quad [8]$$

Proof:

$$E(e^{tx}) = \int_0^\infty e^{tx} f_x\, dx$$
$$= \int_0^\infty e^{tx} \left(\alpha e^{-\alpha x} - 2\alpha\beta e^{-2\alpha x} + 3\alpha\beta e^{-3\alpha x}\right) dx$$
$$= \int_0^\infty \alpha e^{-(\alpha - t)x} - 2\alpha\beta e^{-(2\alpha - t)x} + 3\alpha\beta e^{-(3\alpha - t)x}\, dx$$
$$= \alpha \int_0^\infty e^{-(\alpha - t)x}\, dx - 2\alpha\beta \int_0^\infty e^{-(2\alpha - t)x}\, dx + 3\alpha\beta \int_0^\infty e^{-(3\alpha - t)x}\, dx$$
$$= \frac{\alpha}{\alpha - t} - \frac{2\alpha\beta}{2\alpha - t} + \frac{3\alpha\beta}{3\alpha - t}$$

Hence, the mean and variance of CRTED are as in Equations 9 and 10:

$$E(X) = \frac{6 - \beta}{6\alpha} \quad [9]$$

$$Var(X) = \frac{36 + 2\beta - \beta^2}{36\alpha^2} \quad [10]$$

Mixed Poisson CRTED

Proposition 3: If a conditional random variable $N|X \sim Poisson\,(X)$ is such that X ~ CRTED (α, β), the PMF of the unconditional random variable N has a mixed Poisson-CRTED if its PMF is defined as Equation 11:

$$f_n = \frac{\alpha}{(1+\alpha)^{n+1}} - \frac{2\alpha\beta}{(1+2\alpha)^{n+1}} + \frac{3\alpha\beta}{(1+3\alpha)^{n+1}} \qquad [11]$$

Proof:

$$f_n = \int_0^\infty \frac{x^n e^{-x}}{n!}.\alpha e^{-\alpha x}(1 - 2\beta e^{-\alpha x} + 3\beta e^{-2\alpha x})dx$$

$$= \frac{\alpha}{n!}\int_0^\infty x^n e^{-(1+\alpha)x}(1 - 2\beta e^{-\alpha x} + 3\beta e^{-2\alpha x})dx$$

$$= \frac{\alpha}{n!}\int_0^\infty \left(x^n e^{-(1+\alpha)x} - 2\beta x^n e^{-(1+2\alpha)x} + 3\beta x^n e^{-(1+3\alpha)x}\right)dx$$

$$= \frac{\alpha}{n!}\int_0^\infty \left(\frac{1}{(1+\alpha)^{n+1}}u^n e^{-u} - \frac{2\beta}{(1+2\alpha)^{n+1}}u^n e^{-u} + \frac{3\beta}{(1+3\alpha)^{n+1}}u^n e^{-u}\right)du$$

$$= \frac{\alpha}{(1+\alpha)^{n+1}} - \frac{2\alpha\beta}{(1+2\alpha)^{n+1}} + \frac{3\alpha\beta}{(1+3\alpha)^{n+1}}$$

Special Case: Equation 11 becomes the mixed Poisson Exponential Distribution when the shape parameter $\beta = 0$

Generally, mixed Poisson distributions assign more probability to 0 observations (Shaked, 1980). As shown in Figure 2, the shapes of the PMF of the new distribution are unimodal (Holgate, 1970), with the ability to model data with excess zero. The unimodality exhibited by a mixed Poisson distribution is an attribute derived from the mixing distribution (Karlis & Xekalaki, 2005).

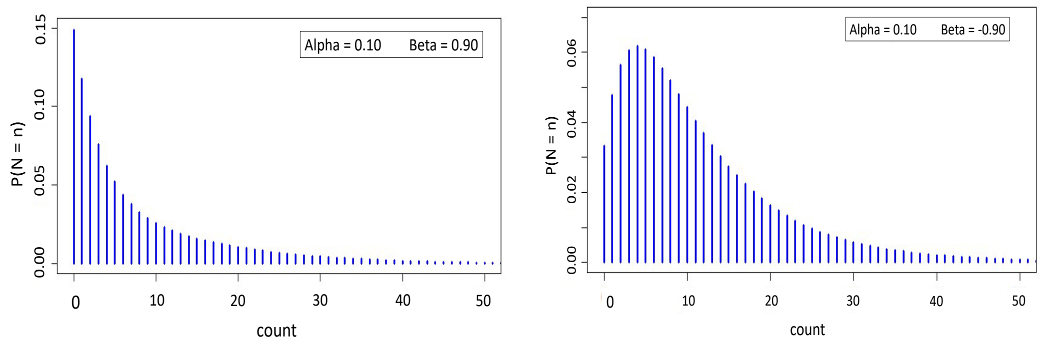

Figure 2. Shapes of the PMF of the mixed Poisson CRTED

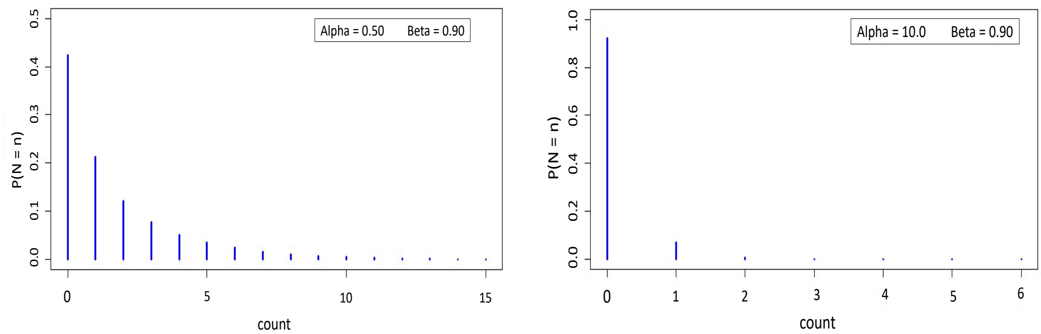

Figure 2. (Continue)

The Proof of Legitimacy of the PMF of Mixed Poisson CRTED

It is shown by proving that sum over f_n is equal to 1, i.e., $\sum_{n=0}^{\infty} f_n = 1$
Therefore,

$$\sum_{n=0}^{\infty}\left(\frac{\alpha}{(1+\alpha)^{n+1}}-\frac{2\alpha\beta}{(1+2\alpha)^{n+1}}+\frac{3\alpha\beta}{(1+3\alpha)^{n+1}}\right)$$

$$= \alpha\sum_{n=0}^{\infty}\frac{1}{(1+\alpha)^{n+1}} - 2\alpha\beta\sum_{n=0}^{\infty}\frac{1}{(1+2\alpha)^{n+1}} + 3\alpha\beta\sum_{n=0}^{\infty}\frac{1}{(1+3\alpha)^{n+1}}$$

$$= \alpha\left(\frac{1}{(1+\alpha)^{1}}+\cdots\right) - 2\alpha\beta\left(\frac{1}{(1+2\alpha)^{1}}+\cdots\right) + 3\alpha\beta\left(\frac{1}{(1+3\alpha)^{1}}+\cdots\right)$$

$$= \alpha\left(\frac{1}{\alpha}\right) - 2\alpha\beta\left(\frac{1}{2\alpha}\right) + 3\alpha\beta\left(\frac{1}{3\alpha}\right) = 1-\beta+\beta = 1$$

THE CDF OF THE MIXED POISSON CRTED

Proposition 4: If a discrete random variable *N* has the mixed Poisson CRTED, the distribution function is obtained as in Equation 12:

$$F_n = 1-\left(\frac{1}{(1+\alpha)^{n+1}}-\frac{\beta}{(1+2\alpha)^{n+1}}+\frac{\beta}{(1+3\alpha)^{n+1}}\right) \qquad [12]$$

Proof:

$$F_n = P(N \le n)$$
$$= 1 - P(N > n)$$
$$= 1 - \sum_{k=n+1}^{\infty} P_k$$
$$= 1 - \sum_{k=n+1}^{\infty}\left(\frac{\alpha}{(1+\alpha)^{k+1}}-\frac{2\alpha\beta}{(1+2\alpha)^{k+1}}+\frac{3\alpha\beta}{(1+3\alpha)^{k+1}}\right)$$

$$= 1 - \left(\sum_{k=n+1}^{\infty} \frac{\alpha}{(1+\alpha)^{k+1}} - \sum_{k=n+1}^{\infty} \frac{2\alpha\beta}{(1+2\alpha)^{k+1}} + \sum_{k=n+1}^{\infty} \frac{3\alpha\beta}{(1+3\alpha)^{k+1}}\right)$$

$$= 1 - \left(\frac{1}{(1+\alpha)^{n+1}} - \frac{\beta}{(1+2\alpha)^{n+1}} + \frac{\beta}{(1+3\alpha)^{n+1}}\right)$$

Mathematical Properties of the Mixed Poisson CRTED

1. According to Willmot (1990), if a random variable ***N*** has a mixed Poisson distribution and the distribution function of its mixing distribution is denoted with F_x, the following results hold:

a. $$P(N \le n) = \int_0^{\infty} \frac{x^n e^{-x}}{n!} F_x \, dx$$

$$= \int_0^{\infty} \frac{x^n e^{-x}}{n!} (1 - e^{-\alpha x} + \beta e^{-2\alpha x} - \beta e^{-3\alpha x}) \, dx$$

$$= \frac{1}{n!}\left(\int_0^{\infty} x^n e^{-x} \, dx - \int_0^{\infty} x^n e^{-(1+\alpha)x} \, dx + \beta \int_0^{\infty} x^n e^{-(1+2\alpha)x} \, dx - \beta \int_0^{\infty} x^n e^{-(1+3\alpha)x} \, dx\right)$$

$$= 1 - \frac{1}{(1+\alpha)^{n+1}} + \frac{\beta}{(1+2\alpha)^{n+1}} - \frac{\beta}{(1+3\alpha)^{n+1}}$$

It is equivalent to the CDF of ***N*** as obtained in Equation 12

b. $$P(N > n) = \int_0^{\infty} \frac{x^n e^{-x}}{n!} [1 - F_x] \, dx$$

$$= \int_0^{\infty} \frac{x^n e^{-x}}{n!} (e^{-\alpha x} - \beta e^{-2\alpha x} + \beta e^{-3\alpha x}) \, dx$$

$$= \frac{1}{n!}\left(\int_0^{\infty} x^n e^{-(1+\alpha)x} \, dx - \int_0^{\infty} \beta x^n e^{-(1+2\alpha)x} \, dx + \int_0^{\infty} \beta x^n e^{-(1+3\alpha)x} \, dx\right)$$

$$= \frac{1}{(1+\alpha)^{n+1}} - \frac{\beta}{(1+2\alpha)^{n+1}} + \frac{\beta}{(1+3\alpha)^{n+1}}$$

2. **Probability Generating Function (PGF)**: Let $\pi(x)$ denotes the PDF of the mixing distribution; Karlis and Xekalaki (2005) defined the PGF of a mixed Poisson random variable N as:

$$P_n(z) = \int_0^{\infty} e^{x(z-1)} \, \pi(x) dx$$

Hence, for the mixed Poisson CRTED, the probability generating function is obtained as:

$$P_n(z) = \int_0^\infty e^{x(z-1)} \alpha(1 - 2\beta e^{-\alpha x} + 3\beta e^{-2\alpha x})dx$$

$$= \alpha \int_0^\infty \left(e^{-(1-z+\alpha)x} - 2\beta e^{-(1-z+2\alpha)x} + 3\beta e^{-(1-z+3\alpha)x}\right)dx$$

Therefore,

$$P_n(z) = \frac{\alpha}{1+\alpha-z} - \frac{2\alpha\beta}{1+2\alpha-z} + \frac{3\alpha\beta}{1+3\alpha-z} \quad [13]$$

Note: Equation 13 can be obtained from Equation 8 (the moment generating function of the mixing distribution, Cubic Rank Transmuted Exponential Distribution) by substituting ***t*** with ***z* – *1***. Hence, the moment-generating function of a mixing distribution can uniquely define the PGF of its mixed Poisson version.

3. **Moment Generating Function:** The MGF in Equation 14 is obtained by replacing z with e^t in Equation 13.

$$M_N(t) = \frac{\alpha}{1+\alpha-e^t} - \frac{2\alpha\beta}{1+2\alpha-e^t} + \frac{3\alpha\beta}{1+3\alpha-e^t} \quad [14]$$

Using the relation $E(N^r) = \frac{d^r M_N(t)}{dt^r}|_{t=0};\ r = 1,2,\ldots$ for the central moments, the first four moments about the origin are obtained respectively as in Equations 15 to 18:

$$E(N) = \frac{6-\beta}{6\alpha} \quad [15]$$

$$E(N^2) = \frac{36 - 5\beta + 18\alpha - 3\alpha\beta}{18\alpha^2} \quad [16]$$

$$E(N^3) = \frac{216 - 19\beta + (216 - 30\beta)\alpha + (36 - 6\beta)\alpha^2}{36\alpha^3} \quad [17]$$

$$E(N^4) = \frac{1296 - 65\beta + (54 - 9\beta)\alpha^3 + (756 - 105\beta)\alpha^2 + (1944 - 171\beta)\alpha}{54\alpha^4} \quad [18]$$

4**. Mean and Variance**: The mean and variance of the mixed Poisson CRTED are obtained as in Equation 19:

$$E(N) = E\big(E(N|X)\big) = E(X) = \frac{6-\beta}{6\alpha} \quad [19]$$

Hence, the mean of the mixed Poisson distribution is equal to the mean of its mixing distributions (Equation 20).

$$Var(N) = E\big(Var(N|X)\big) + V\big(E(N|X)\big) = E(X) + V(X) = \frac{36+2\beta-\beta^2-6\alpha\beta+36\alpha}{36\alpha^2} \quad [20]$$

It implies that the variance of mixed Poisson distribution is the sum of the mean of its conditional variance and the variance of its conditional mean (Karlis & Xekalaki, 2005). Hence, the variance of the mixed Poisson distribution always exceeds that of the classical Poisson distribution.

Using the relationships of the moment about the origin (De Jong & Heller, 2008), skewness and kurtosis measures are respectively defined as Equation 21:

$$S_k(N) = \frac{E(N^3) - 3E(N^2)E(N) + 2\big(E(N)\big)^3}{\big(Var(N)\big)^{\frac{3}{2}}}$$

$$Kurt(N) = \frac{E(N^4) - 4E(N^3)E(N) + 6E(N^2)\big(E(N)\big)^2 - 3\big(E(N)\big)^4}{\big(Var(N)\big)^2}$$

Therefore, the skewness and kurtosis for the mixed Poisson CRTED are as in Equations 21 and 22:

$$S_k(N) = \frac{2\,(216 + 33\beta + 3\beta^2 - \beta^3 + (324 + 18\beta - 9\beta^2)\alpha + (108 - 18\beta)\alpha^2)}{(36 + 2\beta - \beta^2 + (36 - 6\beta)\alpha)^{\frac{3}{2}}} \quad [21]$$

$$Kurt(N) = \frac{1296(1+\alpha)\big(9+9\alpha+\alpha^2\big) - 3\beta^4 + 12\beta^3(1-3\alpha) + 24\beta^2\big(2-3\alpha-6\alpha^2\big) + 8\beta\big(201-27\alpha^3-99\alpha^2+189\alpha\big)}{(\beta^2+(6\alpha-2)\beta-36\alpha-36)^2} \quad [22]$$

The Dispersion Index (DI) measures the degree of dispersion in observations. When DI > 1, there is over-dispersion. There is under-dispersion when DI < 1 and equi-dispersion when DI = 1. The measure is given as in Equation 23:

$$DI = \frac{Var(N)}{E(N)}$$

Hence, the Dispersion Index for the mixed Poisson CRTED is:

$$DI(N) = \frac{36+2\beta-\beta^2-6\alpha\beta+36\alpha}{6\alpha(6-\beta)} \quad [23]$$

Table 1 shows simulated Skewness, Kurtosis, and DI for some parameters of the mixed Poisson CRTED.

Table 1

Skewness, kurtosis, and DI for some parameters for mixed Poisson CRTED

	Skewness			Kurtosis			Dispersion Index		
	α = 1.0	α = 2.0	α = 10	α = 0.1	α = 2.0	α = 10	α = 0.1	α = 2.0	α = 10
β = -0.9	1.92	2.06	3.30	8.93	8.91	15.37	9.07	1.40	1.08
β = -0.5	1.94	2.16	3.43	8.88	9.48	16.48	9.91	1.45	1.09
β = 0.0	2.00	2.31	3.62	9.01	10.33	18.09	11.00	1.50	1.10
β = 0.5	2.11	2.48	3.83	9.36	11.39	20.03	12.14	1.56	1.11
β = 0.9	2.22	2.63	4.02	9.81	12.42	21.88	13.09	1.60	1.12

Remarks:

For fixed β, both skewness and kurtosis increase as parameter α increases.
For fixed β, the dispersion index decreases as α increases.
For fixed α, skewness and kurtosis gradually increase as parameter β increases.
For fixed α, the dispersion index slowly increases at β increases.

Maximum Likelihood Estimation

Given a random sample of size k from the mixed Poisson CRTED with PMF, as stated in Equation 11, the log-likelihood function for the distribution is obtained as in Equation 24.

$$f_n = \frac{\alpha}{(1+\alpha)^{n+1}} - \frac{2\alpha\beta}{(1+2\alpha)^{n+1}} + \frac{3\alpha\beta}{(1+3\alpha)^{n+1}}$$

$$\mathcal{L} = \prod_{i=1}^{k} f_{n_i} = \prod_{i=1}^{k}\left(\frac{\alpha}{(1+\alpha)^{n+1}} - \frac{2\alpha\beta}{(1+2\alpha)^{n+1}} + \frac{3\alpha\beta}{(1+3\alpha)^{n+1}}\right) \qquad [24]$$

$$\ell = log\,\mathcal{L} = \sum_{i=1}^{k} log\left(\frac{\alpha}{(1+\alpha)^{n+1}} - \frac{2\alpha\beta}{(1+2\alpha)^{n+1}} + \frac{3\alpha\beta}{(1+3\alpha)^{n+1}}\right)$$

The estimators for (α, β) denoted with $(\hat{\alpha}, \hat{\beta})$ are the solutions of Equation 24. These solutions contain non-linear equations, which make analytical solutions quite tedious. Using the *MaxLik* functions in the R language (R-Core Team, 2020), the Newton-Raphson algorithm is used to obtain the estimates.

RESULTS AND DISCUSSION

Competing Distributions

The performances of the new distributions are compared with four related discrete distributions. (1) Poisson distribution, (2) Negative Binomial (NB) distribution, (3) Mixed Poisson Exponential (MPE) distributions, and (4) Mixed Poisson Lindley (MPL) distribution. These distributions are considered because they are all mixed probability distributions (the Negative Binomial is obtained when the gamma distribution is used as the mixing distribution, see Greenwood & Yule, 1920). In particular, the mixed Poisson Exponential distribution is a special case of the new proposition. Table 2 gives the PMF of the considered related distributions.

Table 2
Some related distributions

Distribution	**PMF**
Poisson	$\frac{e^{-\alpha}\alpha^{n}}{n!}$
Negative Binomial	$\left(\frac{n+\beta-1}{n}\right)\left(\frac{\alpha}{1+\alpha}\right)^{n}\left(\frac{1}{1+\alpha}\right)^{\beta}$
Mixed Poisson Exponential	$\frac{\alpha}{(1+\alpha)^{n+1}}$
Mixed Poisson Lindley`	$\frac{\alpha^{2}(n+\alpha+2)}{(1+\alpha)^{n+3}}$

Count Data

The performance of the proposed distribution is assessed using six-count observations (Table 3). The first dataset is found in De Jong and Heller (2008). It consists of a one-year insurance policy for third-party claims for Australian vehicle owners. The second data set is the frequency of insurance claims from Belgium in 1993, earlier utilized by Zamani et al. (2014). The third data contains the frequency of claims for automobile injury obtained by the General Insurance Association in Singapore for 1993. The data was earlier used by Frees (2010) and Frees and Valdez (2008). The fourth dataset is the frequency of mistakes in copying groups of random digits first reported by Kemp and Kemp (1965) and in several other propositions (Sah & Mishra, 2019; Samutwachirawong, 2021; Shanker & Mishra, 2014; Shanker & Mishra, 2016) involving new discrete distributions. The fifth dataset (yeast cell counts per square) was used to assess the mixed Poisson Lindley distribution and other related distributions (Shanker & Hagos, 2015). The sixth dataset is the number of epileptic seizures previously used on weighted generalized Poisson distribution (Chakraborty, 2010) and mixed Poisson Transmuted Exponential distribution (Samutwachirawong, 2021).

Table 3
Summary of Count Datasets

Observation	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
0	63232	57178	6996	35	128	126
1	4333	5617	455	11	37	80
2	271	446	28	8	18	59

Table 3 *(Continue)*

Observation	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
3	18	50	4	4	3	42
4	2	8		2	1	24
5						8
6						5
7						4
8						3
Percentage of 0	93.19	90.33	93.49	58.33	68.45	35.90
Dispersion Index	1.08	1.08	1.09	1.61	1.32	1.87
Skewness	4.07	3.52	4.27	1.31	1.75	1.31
Kurtosis	18.50	14.59	20.88	0.72	2.80	1.71

Although the negative binomial distribution provides a good fit for the dataset, the proposed distribution best fits dataset 1, as shown in Table 4, with the least – 2 log-likelihood (36099.26) and chi-square statistic (0.85). The mixed Poisson CRTED also best fits dataset 2 with the least – 2 log-likelihood (44127.64) and chi-square value (10.92), as shown in Table 5.

Table 6 shows that the negative binomial distribution better fits the new proposition with the lowest – 2 log-likelihood and chi-square values. The new position competes quite well with very close values to the minimum obtained when a negative binomial is assumed. From Table 7, the mixed Poisson CRTED has the least chi-square value and the joint least – 2 log-likelihood value (along with the mixed Poisson Lindley distribution). The new proposition also performs creditable well for datasets 5 and 6 with relatively lower – 2 log-likelihood and chi-square values.

As shown in Tables 8 and 9, the results of the fifth and sixth datasets reveal that the new proposition provides a relatively better fit with the least chi-square values (2.55 and 5.64) for the fifth and sixth datasets and lower – 2 log-likelihood.

Table 4

Performance of mixed Poisson CRTED using Dataset 1

Obser.	Freq.	Expected Frequency				
		Poisson-CRTED	Poisson	Neg. Bin	MPE	MPL
0	63232	63232.39	63091.61	63230.60	63252.72	63255.08
1	4333	4330.00	4593.07	4330.57	4291.00	4289.95

Table 4 *(Continue)*

Obser.	Freq.	Expected Frequency				
		Poisson-CRTED	Poisson	Neg. Bin	MPE	MPL
2	271	275.15	167.19	276.48	291.10	290.01
3	18	17.29	4.06	17.22	19.75	19.55
4	2	1.09	0.07	1.06	1.34	1.31
ML Estimate		$\hat{\alpha}$ = 14.6190	$\hat{\alpha}$ = 0.0728	$\hat{\alpha}$ =1.1569	$\hat{\alpha}$ = 13.7408	$\hat{\alpha}$ = 14.6315
		$\hat{\beta}$ = -0.3817		$\hat{\beta}$ = 0.9408		
-2 LL		36099.26	36203.00	36099.36	36100.90	36100.76
Chi-Square		0.85	177.66	0.98	2.29	2.17

Table 5

Performance of mixed Poisson CRTED using Dataset 2

Obser.	Freq.	Expected Frequency				
		Poisson-CRTED	Poisson	Neg. Bin	MPE	MPL
0	57178	57203.04	56949.763	57188.34	57247.47	57243.92
1	5167	5569.68	6019.590	5581.31	5472.99	5479.09
2	446	481.59	318.135	485.28	523.23	521.46
3	50	40.88	11.209	40.47	50.02	49.39
4	8	3.49	0.296	3.30	4.78	4.66
ML Estimate		$\hat{\alpha}$ = 10.4502	$\hat{\alpha}$ = 0.1057	$\hat{\alpha}$ = 1.2791	$\hat{\alpha}$ = 9.460	$\hat{\alpha}$ = 10.2973
		$\hat{\beta}$ = -0.6081		$\hat{\beta}$ = 0.9237		
-2 LL		44127.64	44301.08	44128.62	44136.36	44135.66
Chi-Square		10.92	413.84	12.33	17.44	16.87

Table 6

Performance of mixed Poisson CRTED using Dataset 3

Obser.	Freq.	Expected Frequency				
		Poisson-CRTED	Poisson	Neg. Bin	MPE	MPL
0	6996	6996.35	6977.80	6996.71	6994.08	6994.03
1	455	452.97	487.75	452.52	456.97	457.12
2	28	31.33	17.05	31.38	29.86	29.79
3	4	2.19	0.40	2.22	1.95	1.94

Table 6 *(Continue)*

Obser.	Freq.	Expected Frequency				
		Poisson-CRTED	Poisson	Neg. Bin	MPE	MPL
ML Estimate		$\hat{\alpha}$ =13.5556 $\hat{\beta}$ = 0.3169	$\hat{\alpha}$ = 0.0699	$\hat{\alpha}$ = 0.8740 $\hat{\beta}$ = 0.9260	$\hat{\alpha}$ = 14.3052	$\hat{\alpha}$ = 15.1903
-2 LL		3864.80	3882.36	3864.76	3864.94	3864.96
Chi-Square		1.86	41.96	1.79	2.28	2.32

Table 7

Performance of mixed Poisson CRTED using Dataset 4

Obser.	Freq.	Expected Frequency				
		Poisson-CRTED	Poisson	Neg. Bin	MPE	MPL
0	35	34.19	27.41	33.95	33.64	33.06
1	11	14.27	21.47	14.49	14.78	15.27
2	8	6.31	8.41	6.39	6.49	6.74
3	4	2.85	2.20	2.85	2.85	2.88
4	2	1.30	0.43	1.28	1.25	1.21
ML Estimate		$\hat{\alpha}$ = 1.2250 $\hat{\beta}$ = 0.2544	$\hat{\alpha}$ = 0.7833	$\hat{\alpha}$ = 0.9381 $\hat{\beta}$ = 0.5449	$\hat{\alpha}$ = 1.2766	$\hat{\alpha}$ = 1.7434
-2 LL		146.70	155.10	146.74	146.76	146.70
Chi-Square		2.06	14.44	2.15	2.28	2.50

Table 8

Performance of mixed Poisson CRTED using Dataset 5

Obser.	Freq.	Expected Frequency				
		Poisson-CRTED	Poisson	Neg. Bin	MPE	MPL
0	128	127.26	118.06	126.73	128.09	127.41
1	37	41.43	54.30	42.08	40.35	41.11
2	18	12.76	12.49	12.84	12.71	12.86
3	3	3.87	1.91	3.80	4.00	3.94
4	1	1.17	0.22	1.11	1.26	1.18
ML Estimate		$\hat{\alpha}$ = 2.2633 $\hat{\beta}$ = -0.2481	$\hat{\alpha}$ = 0.4599	$\hat{\alpha}$ = 1.1950 $\hat{\beta}$ = 0.7221	$\hat{\alpha}$ = 2.1744	$\hat{\alpha}$ = 2.7416
-2 LL		340.14	347.66	340.05	340.19	339.98
Chi-Square		2.55	12.16	2.88	2.78	2.72

Table 9

Performance of mixed Poisson CRTED using Dataset 6

Obser.	Freq.	Expected Frequency				
		Poisson-CRTED	Poisson	Neg. Bin	MPE	MPL
0	126	122.60	74.93	120.23	137.96	128.68
1	80	92.65	115.71	93.00	83.74	87.13
2	59	57.85	89.34	59.17	50.82	55.27
3	42	33.75	45.99	34.94	30.85	33.64
4	24	19.19	17.75	19.83	18.72	19.90
5	8	10.82	5.48	10.99	11.36	11.53
6	5	6.10	1.41	5.98	6.90	6.58
7	4	3.45	0.31	3.22	4.19	3.70
8	3	1.96	0.06	1.72	2.54	2.06
ML Estimate		$\hat{\alpha}$ = 0.7285	$\hat{\alpha}$ = 1.5442	$\hat{\alpha}$ = 1.5501	$\hat{\alpha}$ = 0.6476	$\hat{\alpha}$ = 0.9734
		$\hat{\beta}$ = -0.7753		$\hat{\beta}$ = 0.5010		
-2 LL		1189.44	1272.09	1189.88	1196.79	1190.36
Chi-Square		5.64	37.20	6.51	9.65	5.72

CONCLUSION

Using cubic rank transmuted exponential distribution as the mixing distribution (a newly introduced continuous distribution), this paper proposed a new two-parameter mixed Poisson distribution. The mixed Poisson exponential distribution, a special form of the geometric distribution, is a special case of the new proposition. Shapes of the new distribution are unimodal, with the ability to model skewed data with excess zero.

Various mathematical properties of the new proposition, like the measures based on the moment (skewness, kurtosis, and dispersion index), are obtained. Parameters of the distribution are obtained using the maximum likelihood estimation. Six moderately dispersed count observations with varying percentages of zero are used to assess the performance of the new proposition along with other related mixed Poisson distributions (Poisson, negative binomial, mixed Poisson exponential, and mixed Poisson Lindley). Using the – 2 log-likelihood and the chi-square goodness of fit to compare performances, the new proposition performs creditably well and can be used as an alternative for the competing distributions. In some cases, the – 2 Log Likelihood values obtained for the proposed distribution are closer to some competing distributions; however, the chi-square goodness of fit test reveals that the new proposition provides the best fit to the assessed datasets.

ACKNOWLEDGMENT

A sincere appreciation goes to the Federal Government of Nigeria through the Tertiary Education Trust Fund (TETFund) for the Doctoral Scholarship. The authors would like to express their sincere appreciation to all staff of the School of Mathematical Sciences, Universiti Sains Malaysia.

REFERENCES

Adetunji, A. A., & Sabri, S. R. M. (2021). Modelling claim frequency in insurance using count models. *Asian Journal of Probability and Statistics*, *14*(4), 14-20. https://doi.org/10.9734/ajpas/2021/v14i430334

Aguilar, G. A. S., Moala, F. A., & Cordeiro, G. M. (2019). Zero-truncated poisson exponentiated gamma distribution : Application and estimation methods. *Journal of Statistical Theory and Practice, 13*(4), 1-20 . https://doi.org/10.1007/s42519-019-0059-2

Al-kadim, K. A. (2018). Proposed generalized formula for transmuted distribution. *Journal of Babylon University, Pure and Applied Sciences*, *26*(4), 66-74.

Asamoah, K. (2016). On the credibility of insurance claim frequency: Generalized count models and parametric estimators. *Insurance: Mathematics and Economics*, *70*, 339-353. https://doi.org/10.1016/j.insmatheco.2016.07.003

Bhati, D., Sastry, D. V. S., & Maha Qadri, P. Z. (2015). A new generalized poisson-lindley distribution: Applications and properties. *Austrian Journal of Statistics*, *44*(4), 35-51. https://doi.org/10.17713/ajs.v44i4.54

Chakraborty, S. (2010). On some distributional properties of the family of weighted generalized poisson distribution. *Communications in Statistics - Theory and Methods*, *39*(15), 2767-2788. https://doi.org/10.1080/03610920903129141

Das, K. K., Ahmed, I., & Bhattacharjee, S. (2018). A new three-parameter poisson-lindley distribution for modelling over-dispersed count data. *International Journal of Applied Engineering Research*, *13*(23), 16468-16477.

De Jong, P., & Heller, G. Z. (2008). *Generalized Linear Models for Insurance Data*. Cambridge University Press.

Frees, E. W. (2010). *Regression Modeling with Actuarial and Financial Applications*. Cambridge University Press. https://doi.org/10.1017/CBO9780511814372

Frees, E. W., & Valdez, E. A. (2008). Hierarchical insurance claims modeling. *Journal of the American Statistical Association*, *103*(484), 1457-1469. https://doi.org/10.1198/016214508000000823

Greenwood, M., & Yule, G. U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society*, *83*(2), Article 255. https://doi.org/10.2307/2341080

Holgate, P. (1970). The modality of some compound poisson distributions. *Biometrika*, *87*(3), 666-667. https://doi.org/10.1093/biomet/57.3.666

Karina, T. M., Nurrohmah, S., & Fithriani, I. (2019). Heterogeneous Weibull count distribution. *Journal of Physics: Conference Series*, *1218*(1), Article 012020. https://doi.org/10.1088/1742-6596/1218/1/012020

Karlis, D., & Xekalaki, E. (2005). Mixed poisson distributions. *International Statistical Review*, *73*(1), 35-58. https://doi.org/10.1111/j.1751-5823.2005.tb00250.x

Kemp, C. D., & Kemp, A. W. (1965). Some propoerties of the Hermite distribution. *Biometrika*, *52*(3-4), 381-394. https://doi.org/10.1093/biomet/52.3-4.381

Mohammed, Y. A., Yatim, B., & Ismail, S. (2015). Mixture model of the exponential, Gamma and Weibull distributions to analyse heterogeneous survival data. *Journal of Scientific Research and Reports, 5*(2), 132-139. https://doi.org/10.9734/JSRR/2015/15014

Omari, C. O., Nyambura, S. G., & Mwangi, J. M. W. (2018). Modeling the frequency and severity of auto insurance claims using statistical distributions. *Journal of Mathematical Finance*, *8*(1), 137-160. https://doi.org/10.4236/jmf.2018.81012

Ong, S. H., Low, Y. C., & Toh, K. K. (2021). Recent developments in mixed poisson distributions. *ASM Science Journal*, *14*(3), 1-10. https://doi.org/10.32802/asmscj.2020.464

R-Core Team. (2020). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. https://www.r-project.org/

Rasekhi, M., Alizadeh, M., Altun, E., Hamedani, G., Afify, A. Z., & Ahmad, M. (2017). The modified exponential distribution with applications. *Pakistan Journal of Statistics*, *33*(5), 383-398.

Sah, B. K., & Mishra, A. (2019). A Generalised Exponential-Lindley Mixture of Poisson Distribution. *Nepalese Journal of Statistics, 3*, 11-20. https://doi.org/10.4135/9781412952644.n350

Samutwachirawong, S. (2021). Poisson-exponential and Gamma distribution: Properties and applications. *Journal of Applied Statistics and Information Technology*, *6*(2), 17-24.

Sankaran, M. (1970). 275. Note: The discrete poisson-lindley distribution. *Biometrics*, *26*(1), 145-149. https://doi.org/10.2307/2529053

Shaked, M. (1980). On mixtures from exponential families. *Journal of Royal Statistical Society. Series B (Methodological)*, *42*(2), 192-198. https://doi.org/10.1111/j.2517-6161.1980.tb01118.x

Shanker, R., & Hagos, F. (2015). On poisson-lindley distribution and its applications to biological sciences. *Biometrics and Biostatistics International Journal*, *2*(5), 1-5.

Shanker, R., & Mishra, A. (2014). A Two-Parameter Poisson-Lindley distribution. *International Journal of Statistics and Systems*, *9*(1), 79-85.

Shanker, R., & Mishra, A. (2016). A Quasi Poisson-Lindley distribution. *Journal of the Indian Statistical Association*, *54*(1), 113-125.

Willmot, G. E. (1990). Asymptotic tail behaviour of poisson mixtures with applications. *Advances in Applied Probability*, *22*(1), 147-159. https://doi.org/10.2307/1427602

Zakerzadeh, H., & Dolati, A. (2009). Generalized Lindley distribution. *Journal of Mathematical Extension*, *3*(2), 1-17.

Zamani, H., Ismail, N., & Faroughi, P. (2014). Poisson-weighted exponential univariate version and regression model with applications. *Journal of Mathematics and Statistics*, *10*(2), 148-154. https://doi.org/10.3844/jmssp.2014.148.154