STATISTICS IN TRANSITION new series, June 2024
Vol. 25, No. 2, pp. 103–120, https://doi.org/10.59170/stattrans-2024-017
Received – 05.05.2023; accepted – 19.02.2024

On the Poisson-transmuted exponential distribution and its application to frequency of claim in actuarial science

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

Abstract

This study proposes a new discrete distribution in the mixed Poisson paradigm to obtain a distribution that provides a better fit to skewed and dispersed count observation with excess zero. The cubic transmutation map is used to extend the exponential distribution, and the obtained continuous distribution is assumed for the parameter of the Poisson distribution. Various moment-based properties of the new distribution are obtained. The Nelder-Mead algorithm provides the fastest convergence iteration under the maximum likelihood estimation technique. The shapes of the proposed new discrete distribution are similar to those of the mixing distribution. Frequencies of insurance claims from different countries are used to assess the performance of the new proposition (and its zero-inflated form). Results show that the new distribution outperforms other competing ones in most cases. It is also revealed that the natural form of the new distribution outperforms its zero-inflated version in many cases despite having observations with excess zero counts.

Key words: mixed Poisson-exponential distribution, skewed count data, dispersed observation, zero-inflated model, claim frequency.

1. Introduction

Distributions that were well regarded in the past have been improved recently, especially with the advent of software and programming languages that make mathematical computations easier. The situation has provided an enabling platform for introducing more complex distributions that better-fit observations.

In compounding baseline distributions, different techniques have been developed. Among them are: quadratic transmutation (Shaw & Buckley, 2007); beta-extended generalized (Alexander *et al.*, 2012); exponentiated generalized (Yousof *et al.*, 2015); transmuted exponentiated generalized (Shehata *et al.*, 2021); alpha-power transformation (Mahdavi & Kundu, 2017); and various forms of cubic rank

[1] Corresponding author. School of Mathematical Sciences, Universiti Sains Malaysia, 11800, Penang, Malaysia. E-mail: rijal@usm.my.

[2] Department of Statistics, Federal Polytechnic, 35101, Ile-Oluji, Nigeria.

transmutations (Al-kadim, 2018; Aslam *et al.*, 2018; Granzotto *et al.*, 2017; Rahman *et al.*, 2019).

The classical Poisson distribution is popular among distributions for discrete observations. The distribution assumes equidispersion (equality of mean and variance). In many cases, count data are overdispersed (Nikoloulopoulos & Karlis, 2008), especially the frequency of claims in actuary science (Adcock *et al.*, 2015; Adetunji & Sabri, 2021; Omari *et al.*, 2018) with attending excess zero counts. Assuming the Poisson for such observation may lead to model misspecification. This has led to the development of several techniques to model skewed and overdispersed count data, such as the negative binomial. The negative binomial distributions come from the mixed Poisson paradigm when the gamma distribution is assumed for the parameter of the Poisson (Greenwood & Yule, 1920). Several other continuous distributions with positive supports have been assumed for the Poisson parameter (Al-Awadhi & Ghitany, 2001; Bhati et al., 2015, 2017; Das et al., 2018; Gómez-Déniz & Calderín-Ojeda, 2016; Mahmoudi & Zakerzadeh, 2010).

If a discrete random variable $X \sim Poisson(\lambda)$ and $\lambda \sim \pi(\lambda)$ where $\pi(\lambda)$ is the probability distribution function (PDF) with positive supports, a mixed Poisson distribution (with $\pi(\lambda)$ as the mixing distribution) is obtained when the unconditional distribution for $\boldsymbol{X}$ is expressed from equation (1).

$$P_x = \int_0^\infty f(x|\lambda)\, \pi(\lambda)\, d\lambda \tag{1}$$

With a scale parameter θ, the distribution function (CDF) of a random variable λ with the classical exponential distribution is defined in equation (2) as:

$$G(\lambda) = 1 - e^{-\theta\lambda} \tag{2}$$

Leveraging on the skewness and unimodality properties of transmuted exponential distribution (which also characterize observations in claim frequency (Karlis & Xekalaki, 2005)), this research extends the classical exponential distribution using the cubic transmutation map (Al-kadim, 2018). If the baseline CDF is as given in equation (2) and ρ is the transmutation parameter, the CDF of the cubic rank transmutation map due to (Al-kadim, 2018) is defined in equation (3) as:

$$\pi(\lambda) = (1+\rho)G(\lambda) - 2\rho\big(G(\lambda)\big)^2 + \rho\big(G(\lambda)\big)^3,\ \lambda > 0, |\rho| \leq 1 \tag{3}$$

The classical exponential distribution serves as baseline distribution in diverse fields of study. In probability theory and applications, extended exponential distribution has enjoyed patronage (Yang et al., 2021). This spans different fields of study, including environmental, economic, reliability, and industrial (Aguilar *et al.*, 2019; Rasekhi *et al.*, 2017).

2. Materials and Methods

Cubic Rank Transmuted Exponential Distribution (CRTED)

The distribution is obtained when the transmutation map in equation (3) is used to extend the exponential distribution given in equation (2). This gives the Cubic Rant Transmuted Exponential Distribution (CRTED) with the CDF and PDF given in equations (4) and (5) as:

$$F(\lambda) = 1 - e^{-\theta\lambda} + \rho e^{-2\theta\lambda} - \rho e^{-3\theta\lambda} \tag{4}$$

$$f(\lambda) = \theta e^{-\theta\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right) \tag{5}$$

The corresponding survival and hazard rate functions are respectively given in equation (6) and (7) as:

$$S(\lambda) = e^{-\theta\lambda} - \rho e^{-2\theta\lambda} + \rho e^{-3\theta\lambda} \tag{6}$$

$$h(\lambda) = \frac{\theta\left(1-2\rho e^{-\theta\lambda}+3\rho e^{-2\theta\lambda}\right)}{1-\rho e^{-\theta\lambda}+\rho e^{-2\theta\lambda}} \tag{7}$$

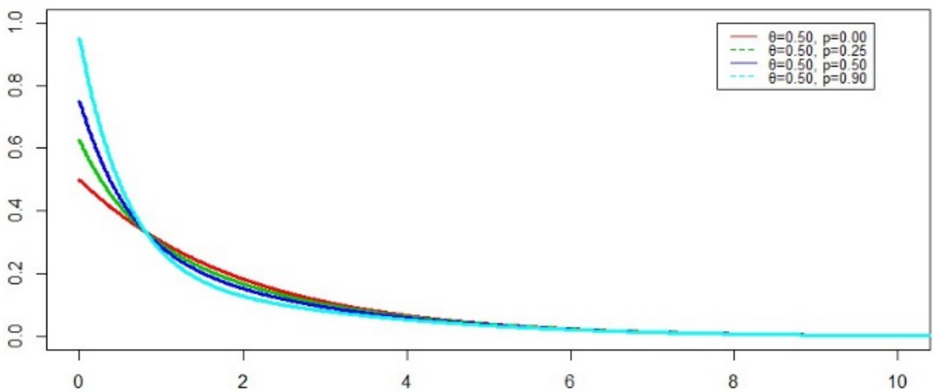

Figure 1: Shapes of the PDF CRTED

Figure 1 shows the shapes of the PDF of the CRTED for different values of θ and ρ. The figure shows that the PDF is a monotonically decreasing function.

r^{th} Moment of the CRTED

Proposition 1: If a random variable λ has a CRTED, the r^{th} moment is defined in equation (8) as:

$$E(\lambda^r) = \left(1 - \frac{\wp}{2^r} + \frac{\wp}{3^r}\right)\frac{r!}{\theta^r} \tag{8}$$

Proof:

$$E(\lambda^r) = \int_0^\infty \lambda^r f(\lambda)\, d\lambda = \int_0^\infty \lambda^r \left(\theta e^{-\theta\lambda} - 2\rho\theta e^{-2\theta\lambda} + 3\rho\theta e^{-3\theta\lambda}\right) d\lambda$$

$$= \theta \int_0^\infty \lambda^r e^{-\theta\lambda}\, d\lambda - 2\rho\theta \int_0^\infty \lambda^r e^{-2\theta\lambda}\, d\lambda + 3\rho\theta \int_0^\infty \lambda^r e^{-3\theta\lambda}\, d\lambda$$

$$= \frac{r!}{\theta^r} - \frac{\rho r!}{(2\theta)^r} + \frac{\rho r!}{(3\theta)^r} = \left(1 - \frac{\rho}{2^r} + \frac{\rho}{3^r}\right)\frac{r!}{\theta^r}$$

*Mo*ment Generating Function **of the CRTED**

Proposition 2: If a random variable λ has a CRTED, the MGF is defined in equation (9) as:

$$E\left(e^{t\lambda}\right) = \frac{\theta}{\theta - t} - \frac{2\rho\theta}{2\theta - t} + \frac{3\rho\theta}{3\theta - t} \tag{9}$$

Proof:

$$E\left(e^{t\lambda}\right) = \int_0^\infty e^{t\lambda}\left(\theta e^{-\theta\lambda} - 2\rho\theta e^{-2\theta\lambda} + 3\rho\theta e^{-3\theta\lambda}\right) d\lambda$$

$$= \int_0^\infty \theta e^{-(\theta - t)\lambda} - 2\rho\theta e^{-(2\theta - t)\lambda} + 3\rho\theta e^{-(3\theta - t)\lambda}\, d\lambda$$

$$= \theta \int_0^\infty e^{-(\theta - t)\lambda}\, d\lambda - 2\rho\theta \int_0^\infty e^{-(2\theta - t)\lambda}\, d\lambda + 3\rho\theta \int_0^\infty e^{-(3\theta - t)\lambda}\, d\lambda$$

$$= \frac{\theta}{\theta - t} - \frac{2\rho\theta}{2\theta - t} + \frac{3\rho\theta}{3\theta - t}$$

3. Poisson Crted

PMF of the Poisson CRTED

Proposition 3: If a discrete random variable $X \sim Poisson(\lambda)$ and $\lambda \sim CRTED(\rho, \theta)$ then the PMF of ***X*** has a Poisson-CRTED if its PMF is defined in equation (10) as:

$$P_x = \frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}} \tag{10}$$

Proof

Since $f(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$, $\pi(\lambda) = \theta e^{-\theta\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right)$

$$P_x = \int_0^\infty f(x|\lambda)\, \pi(\lambda)\, d\lambda$$

$$= \int_0^\infty \frac{\lambda^x e^{-\lambda}}{x!} . \theta e^{-\theta\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right) d\lambda$$

$$= \frac{\theta}{x!} \int_0^\infty \lambda^x e^{-(1+\theta)\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right) d\lambda$$

$$= \frac{\theta}{x!} \int_0^\infty \left(\frac{1}{(1+\theta)^{x+1}} u^x e^{-u} - \frac{2\rho}{(1+2\theta)^{x+1}} u^x e^{-u} + \frac{3\rho}{(1+3\theta)^{x+1}} u^x e^{-u}\right) du$$

$$= \frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}}$$

Shapes of the PMF of the Poisson CRTED in Figure 2 show that the distribution is unimodal with the ability to model positively skewed observation with excess zero counts.

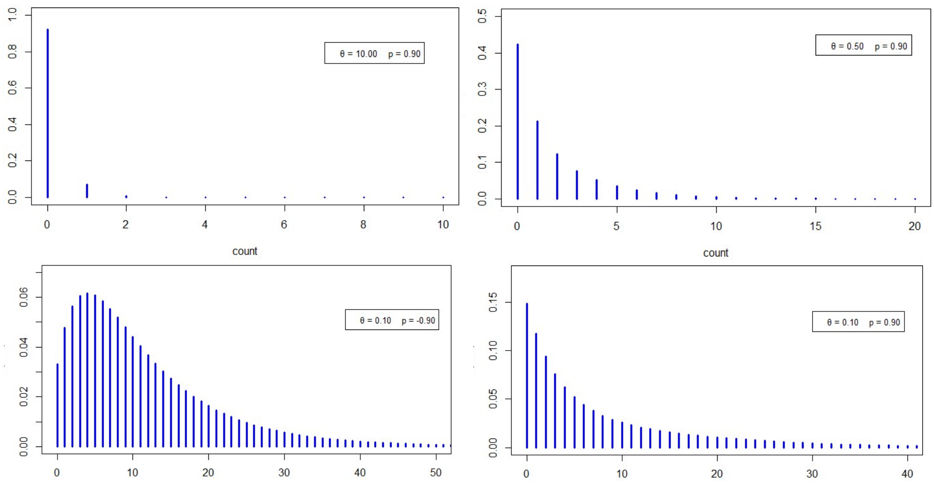

Figure 2: Shapes of the PMF of the Poisson CRTED

Mathematical Properties of the Poisson CRTED

3.1. Probability Generating Function: The PGF of a random variable $\boldsymbol{X}$ with the Poisson-CRTED is obtained in equation (11).

$$P_x(z) = \int_0^\infty e^{\lambda(z-1)}\, \pi(\lambda) d\lambda$$

$$= \int_0^\infty e^{\lambda(z-1)} \theta e^{-\theta\lambda}\left(1 - 2\rho e^{-\theta\lambda} + 3\rho e^{-2\theta\lambda}\right) d\lambda$$

$$= \theta \int_0^\infty \left(e^{-(1-z+\theta)\lambda} - 2\rho e^{-(1-z+2\theta)\lambda} + 3\rho e^{-(1-z+3\theta)\lambda}\right) d\lambda$$

Therefore,

$$P_x(z) = \frac{\theta}{1+\theta-z} - \frac{2\rho\theta}{1+2\theta-z} + \frac{3\rho\theta}{1+3\theta-z} \tag{11}$$

3.2. Moment Generating Function: The MGF is obtained in equation (12) by replacing $\boldsymbol{z}$ with e^t in (11).

$$M_X(t) = \frac{\theta}{1+\theta-e^t} - \frac{2\rho\theta}{1+2\theta-e^t} + \frac{3\rho\theta}{1+3\theta-e^t} \tag{12}$$

3.3. Mean and Variance: The mean and variance of the Poisson-CRTED are respectively obtained in equations (13) and (14) as:

$$E(X) = \frac{6-\rho}{6\theta} \tag{13}$$

$$Var(X) = \frac{36+2\rho-\rho^2-6\rho\theta+36\theta}{36\theta^2} \tag{14}$$

The Coefficient of Variation and the Dispersion Index are given in equations (15) and (16).

$$CV(X) = \frac{\sqrt{36+2\rho-\rho^2-6\rho\theta+36\theta}}{6-\rho} \tag{15}$$

$$DI(X) = \frac{36+2\rho-\rho^2-6\rho\theta+36\theta}{6\theta(6-\rho)} \tag{16}$$

3.4. Skewness and Kurtosis: The first four raw moments for a random variable with the Poisson-CRTED are presented in equations (17) to (20).

$$E(X) = \frac{6-\rho}{6\theta} \tag{17}$$

$$E(X^2) = \frac{36-5\rho+18\theta-3\rho\theta}{18\theta^2} \tag{18}$$

$$E(X^3) = \frac{216-19\rho+(216-30\rho)\theta+(36-6\rho)\theta^2}{36\theta^3} \tag{19}$$

$$E(X^4) = \frac{1296-65\rho+(54-9\rho)\theta^3+(756-105\rho)\theta^2+(1944-171\rho)\theta}{54\theta^4} \tag{20}$$

The skewness and kurtosis for the Poisson CRTED are given in equations (21) and (22) as:

$$S_k(X) = \frac{2(216+33\rho+3\rho^2-\rho^3+(324+18\rho-9\rho^2)\theta+(108-18\rho)\theta^2)}{(36+2\rho-\rho^2+(36-6\rho)\theta)^{\frac{3}{2}}} \tag{21}$$

$$K(X) = \frac{1296(1+\theta)(9+9\theta+\theta^2)-3\rho^4+12\rho^3(1-3\theta)+24\rho^2(2-3\theta-6\theta^2)+8\rho(201-27\theta^3-99\theta^2+189\theta)}{(\rho^2+(6\theta-2)\rho-36\theta-36)^2} \tag{22}$$

Tables 1–3 show simulated Skewness, Kurtosis, and Dispersion Index for selected parameters of the Poisson CRTED

Table 1: Skewness for some parameters of the Poisson-CRTED

Specification	$\theta = 0.1$	$\theta = 2.0$	$\theta = 10$
$\rho = -0.9$	1.920	2.056	3.300
$\rho = -0.5$	1.938	2.162	3.433
$\rho = 0.0$	2.002	2.309	3.618
$\rho = 0.5$	2.108	2.478	3.827
$\rho = 0.9$	2.222	2.632	4.016

Table 2: Kurtosis for some parameters of the Poisson-CRTED

Specification	$\theta = 0.1$	$\theta = 2.0$	$\theta = 10$
$\rho = -0.9$	8.928	8.908	15.366
$\rho = -0.5$	8.877	9.479	16.476
$\rho = 0.0$	9.009	10.333	18.091
$\rho = 0.5$	9.360	11.387	20.028
$\rho = 0.9$	9.813	12.415	21.879

Table 3: Dispersion Index for some parameters of the Poisson-CRTED

Specification	$\theta = 0.1$	$\theta = 2.0$	$\theta = 10$
$\rho = -0.9$	9.065	1.403	1.081
$\rho = -0.5$	9.910	1.446	1.089
$\rho = 0.0$	11.000	1.500	1.100
$\rho = 0.5$	12.136	1.557	1.111
$\rho = 0.9$	13.088	1.604	1.121

Remarks:

i. For fixed ρ, both skewness and kurtosis increase as θ increases.
ii. For fixed ρ, the dispersion index decreases as θ increases.
iii. For fixed θ, both skewness and kurtosis increase as ρ increases.
iv. For fixed θ, the dispersion index slowly increases as ρ increases.

Maximum Likelihood Estimation of the Poisson-CRTED

Assuming $x_1, x_2, \dots, x_n$ are a random sample of size n from the Poisson-CRTED, the log-likelihood function of the distribution is obtained in equation (23) as:

$$\mathcal{L} = \prod_{i=1}^{n} P_{(x_i)} = \prod_{i=1}^{n} \left(\frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}} \right)$$

$$\ell = log\, \mathcal{L} = \sum_{i=1}^{n} log \left(\frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}} \right) \tag{23}$$

Equation (23) gives a non-linear model that can be solved numerically with different algorithms from the optimr (Nash *et al.*, 2019) from the R language (R-Core Team, 2020).

4. Zero-Inflated Poisson-Crted

Proposition 4: If a discrete random variable $\boldsymbol{X}$ has a Poisson-CRTED with PMF denoted by P_x, if the inflation parameter is denoted with τ, the PMF of the zero-inflated Poisson-CRTED is given in equation (24).

$$P_x^{ZI} = \begin{cases} \tau + (1-\tau)P_0, & x = 0 \\ (1-\tau)P_x, & x = 1,2,3,\dots \end{cases}$$

This is obtained as:

$$P_x^{ZI} = \begin{cases} \tau + (1-\tau)\left(\frac{\theta}{(1+\theta)} - \frac{2\rho\theta}{(1+2\theta)} + \frac{3\rho\theta}{(1+3\theta)}\right), \ x = 0 \\ (1-\tau)\left(\frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}}\right), \ x = 1,2,3,\ldots \end{cases} \tag{24}$$

Proof

Since the PMF of the Poisson CRTED is given as: $P_x = \frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}}$ and $P_0 = \frac{\theta}{(1+\theta)} - \frac{2\rho\theta}{(1+2\theta)} + \frac{3\rho\theta}{(1+3\theta)}$, hence the result.

Mathematical Properties of the Zero-Inflated Poisson CRTED

The PGF of the Zero-Inflated Poisson CRTED denoted by $P_x^{ZI}(z)$ is obtained using: $P_x^{ZI}(z) = (1-\tau)P_x(z)$ where $P_x(z)$ is the PGF of the Poisson CRTED in equation (11). Hence, the PGF of the ZI-Poisson CRTED is given in equation (25) as:

$$P_x^{ZI}(z) = (1-\tau)\left(\frac{\theta}{1+\theta-z} - \frac{2\rho\theta}{1+2\theta-z} + \frac{3\rho\theta}{1+3\theta-z}\right) \tag{25}$$

The MGF is therefore expressed in equation (26) as:

$$M_X^{ZI}(t) = (1-\tau)\left(\frac{\theta}{1+\theta-e^t} - \frac{2\rho\theta}{1+2\theta-e^t} + \frac{3\rho\theta}{1+3\theta-e^t}\right) \tag{26}$$

If the r^{th} moment of the Poisson CRTED is denoted by $E(X^r)$, then the r^{th} moment of the ZI-Poisson CRTED is defined as:
$m_r = E(X_{ZI}^r) = (1-\tau)E(X^r)$

Hence, the first four moments of the ZI-Poisson CRTED are given in equations (27) – (30).

$$m_1 = (1-\tau)\frac{6-\rho}{6\theta} \tag{27}$$

$$m_2 = (1-\tau)\frac{36-5\rho+18\theta-3\rho\theta}{18\theta^2} \tag{28}$$

$$m_3 = (1-\tau)\frac{216-19\rho+(216-30\rho)\theta+(36-6\rho)\theta^2}{36\theta^3} \tag{29}$$

$$m_4 = (1-\tau)\frac{1296-65\rho+(54-9\rho)\theta^3+(756-105\rho)\theta^2+(1944-171\rho)\theta}{54\theta^4} \tag{30}$$

MLE of the Parameters of the ZI-Poisson CRTED

If a random variable ***X*** is assumed to follow the ZI-Poisson CRTED with its PMF P_x indexed with (θ, ρ), and τ as the zero-inflation parameter, then the likelihood function is defined as:

$$\mathcal{L}(\tau, \Theta) = \prod_{n_0}(\tau + (1-\tau)P_0)\prod_{n_1}\left((1-\tau)P(x>0)\right)$$

where n_0 is the frequency of zero counts in the dataset; n_1 is the frequency of non-zero counts; $n = (n_0 + n_1)$, P_0 is the realization of at $x = 0$. The log-likelihood function is obtained as follows:

$$\ell = n_0 \ln(\tau + (1-\tau)P_0) + n_1 \ln(1-\tau) + \left(\sum_{n_1} \ln(P_x)\right)$$

$$= n_0 \ln\left(\tau + (1-\tau)\left(\frac{\theta}{(1+\theta)} - \frac{2\rho\theta}{(1+2\theta)} + \frac{3\rho\theta}{(1+3\theta)}\right)\right) + n_1 \ln(1-\tau) + \sum_{n_1} \ln\left(\left(\frac{\theta}{(1+\theta)^{x+1}} - \frac{2\rho\theta}{(1+2\theta)^{x+1}} + \frac{3\rho\theta}{(1+3\theta)^{x+1}}\right)\right)$$

$$\frac{\partial\ell}{\partial\tau} = \frac{n_0(1-P_0)}{\tau + (1-\tau)P_0} - \frac{n_1}{(1-\tau)}$$

$$\hat{\tau} = \frac{n_0}{n_1} - \frac{n_1}{n}\left(\frac{P_0}{1-P_0}\right)$$

The MLE for parameters (τ, θ, ρ) is obtained numerically by solving $\frac{\partial\ell}{\partial\tau} = 0$, $\frac{\partial\ell}{\partial\theta} = 0$ and $\frac{\partial\ell}{\partial\rho} = 0$. Among competing algorithms for optimizations that come with the optimr packages (Nash *et a*l., 2019) in R language (R-Core Team, 2020), the Nelder-Mead provides the fastest iterations for convergence and the least log-likelihood values.

Competing Distributions

The new propositions are assessed by comparing their performances with the (i) Poisson, (ii) Zero Inflated Poisson (ZIP), (iii) Negative Binomial (Neg. Bin.), and (iv) Zero Inflated Negative Binomial (ZI-Neg. Bin.) distributions.

5. Simulation Studies

The new proposition is assessed in this study by simulating count observations with similar characteristics to the claim frequency in actuaries. The simulated data are based on the assumption that a policyholder may not report more than 4 claims in a reference period (usually a year). The algorithm utilized for the simulation is:

i. Specify the proportion of zero in each sample (P_0 = 0.5, 0.7, 0.9)
ii. Simulate a random sample of size n (n = 50, 250, 500, and 1000)
iii. Obtain parameter estimates and log-likelihood for each assumed distribution.

Table 4: Results for simulated data when $P_0 = 0.50$

Sample Size	*Distribution*	*Parameter Estimates*	*–LL*
n = 50	*PCRTED*	$\hat{\theta} = 1.341;\ \hat{\rho} = -0.771$	62.259
	ZI-PCRTED	$\hat{\theta} = 20.060;\ \hat{\rho} = 5.867;\ \hat{\tau} = -716.5$	65.594
	Poisson	$\hat{\theta} = 0.840$	65.213
	ZI-Poisson	$\hat{\theta} = 1.247;\ \hat{\tau} = 0.327$	61.232
	Neg. Bin.	$\hat{\theta} = 1.543;\ \hat{\rho} = 0.647$	63.034
	ZI-Neg. Bin.	$\hat{\theta} = 0.00004;\ \hat{\rho} = 0.357;\ \hat{\tau} = -117.2$	63.188
			65.079

Table 4: Results for simulated data when $P_0 = 0.50$ (cont.)

Sample Size	*Distribution*	*Parameter Estimates*	*–LL*
$n = 250$	*PCRTED*	$\hat{\theta} = 1.209;\ \hat{\rho} = -0.484$	323.713
	ZI-PCRTED	$\hat{\theta} = 1.256;\ \hat{\rho} = 0.696;\ \hat{\tau} = 0.372$	387.764
	Poisson	$\hat{\theta} = 0.892$	339.839
	ZI-Poisson	$\hat{\theta} = 1.379;\ \hat{\tau} = 0.353$	324.602
	Neg. Bin.	$\hat{\theta} = 1.357;\ \hat{\rho} = 0.603$	326.427
	ZI-Neg. Bin.	$\hat{\theta} = 0.763;\ \hat{\rho} = 0.518;\ \hat{\tau} = 0.377$	387.993
$n = 500$	*PCRTED*	$\hat{\theta} = 1.330;\ \hat{\rho} = -0.647$	628.953
	ZI-PCRTED	$\hat{\theta} = 1.338;\ \hat{\rho} = -0.006;\ \hat{\tau} = 0.677$	718.784
	Poisson	$\hat{\theta} = 0.832$	653.206
	ZI-Poisson	$\hat{\theta} = 1.244;\ \hat{\tau} = 0.331$	631.018
	Neg. Bin.	$\hat{\theta} = 1.414;\ \hat{\rho} = 0.630$	629.484
	ZI-Neg. Bin.	$\hat{\theta} = 1.072;\ \hat{\rho} = 0.588;\ \hat{\tau} = 0.681$	718.719
$n = 1000$	*PCRTED*	$\hat{\theta} = 1.377;\ \hat{\rho} = -1.220$	1284.447
	ZI-PCRTED	$\hat{\theta} = 1.365;\ \hat{\rho} = -0.852;\ \hat{\tau} = 0.838$	1404.054
	Poisson	$\hat{\theta} = 0.873$	1322.324
	ZI-Poisson	$\hat{\theta} = 1.192;\ \hat{\tau} = 0.268$	1293.749
	Neg. Bin.	$\hat{\theta} = 1.785;\ \hat{\rho} = 0.672$	1285.411
	ZI-Neg. Bin.	$\hat{\theta} = 1.562;\ \hat{\rho} = 0.651;\ \hat{\tau} = 0.838$	1403.110

Table 5: Results for simulated data when $P_0 = 0.70$

Sample Size	*Distribution*	*Parameter Estimates*	*–LL*
$n = 50$	*PCRTED*	$\hat{\theta} = 1.891;\ \hat{\rho} = 0.788$	43.395
	ZI-PCRTED	$\hat{\theta} = 5.929;\ \hat{\rho} = 5.435;\ \hat{\tau} = -25.544$	43.421
	Poisson	$\hat{\theta} = 0.460$	47.622
	ZI-Poisson	$\hat{\theta} = 0.926;\ \hat{\tau} = 0.503$	45.397
	Neg. Bin.	$\hat{\theta} = 0.760;\ \hat{\rho} = 0.623$	45.425
	ZI-Neg. Bin.	$\hat{\theta} = 0.0003;\ \hat{\rho} = 0.449;\ \hat{\tau} = -10.40$	46.220
$n = 250$	*PCRTED*	$\hat{\theta} = 1.896;\ \hat{\rho} = 0.314$	236.621
	ZI-PCRTED	$\hat{\theta} = 2.039;\ \hat{\rho} = 2.445;\ \hat{\tau} = -0.047$	287.128
	Poisson	$\hat{\theta} = 0.500$	248.799
	ZI-Poisson	$\hat{\theta} = 0.912;\ \hat{\tau} = 0.452$	239.519
	Neg. Bin.	$\hat{\theta} = 0.911;\ \hat{\rho} = 0.646$	238.655
	ZI-Neg. Bin.	$\hat{\theta} = 0.362;\ \hat{\rho} = 0.552;\ \hat{\tau} = -0.032$	287.579
$n = 500$	*PCRTED*	$\hat{\theta} = 1.910;\ \hat{\rho} = 0.574$	461.388
	ZI-PCRTED	$\hat{\theta} = 1.969;\ \hat{\rho} = 1.559;\ \hat{\tau} = 0.560$	541.627
	Poisson	$\hat{\theta} = 0.474$	483.392
	ZI-Poisson	$\hat{\theta} = 0.919;\ \hat{\tau} = 0.484$	462.576
	Neg. Bin.	$\hat{\theta} = 0.822;\ \hat{\rho} = 0.634$	462.559
	ZI-Neg. Bin.	$\hat{\theta} = 0.539;\ \hat{\rho} = 0.587;\ \hat{\tau} = 0.564$	542.152
$n = 1000$	*PCRTED*	$\hat{\theta} = 1.879;\ \hat{\rho} = 0.776$	910.515
	ZI-PCRTED	$\hat{\theta} = 1.907;\ \hat{\rho} = 1.246;\ \hat{\tau} = 0.791$	1021.272
	Poisson	$\hat{\theta} = 0.464$	958.230
	ZI-Poisson	$\hat{\theta} = 0.938;\ \hat{\tau} = 0.506$	912.204
	Neg. Bin.	$\hat{\theta} = 0.755;\ \hat{\rho} = 0.619$	913.020
	ZI-Neg. Bin.	$\hat{\theta} = 0.617;\ \hat{\rho} = 0.597;\ \hat{\tau} = 0.792$	1022.125

Table 6: Results for simulated data when $P_0 = 0.90$

Sample Size	*Distribution*	*Parameter Estimates*	*–LL*
$n = 50$	*PCRTED*	$\hat{\theta} = 3.055;\ \hat{\rho} = 3.282$	21.126
	ZI-PCRTED	$\hat{\theta} = 11.610;\ \hat{\rho} = 5.759;\ \hat{\tau} = -34.25$	25.230
	Poisson	$\hat{\theta} = 0.160$	24.740
	ZI-Poisson	$\hat{\theta} = 1.027;\ \hat{\tau} = 0.844$	21.039
	Neg. Bin.	$\hat{\theta} = 0.144;\ \hat{\rho} = 0.473$	21.880
	ZI-Neg. Bin.	$\hat{\theta} = 0.00005;\ \hat{\rho} = 0.392; \hat{\tau} = -14.41$	26.255
$n = 250$	*PCRTED*	$\hat{\theta} = 1.811;\ \hat{\rho} = 4.072$	122.289
	ZI-PCRTED	$\hat{\theta} = 4.097;\ \hat{\rho} = 5.615;\ \hat{\tau} = -11.987$	125.321
	Poisson	$\hat{\theta} = 0.208$	156.709
	ZI-Poisson	$\hat{\theta} = 1.493;\ \hat{\tau} = 0.861$	121.234
	Neg. Bin.	$\hat{\theta} = 0.106;\ \hat{\rho} = 0.338$	122.435
	ZI-Neg. Bin.	$\hat{\theta} = 0.00001;\ \hat{\rho} = 0.303; \hat{\tau} = -69.32$	126.856
$n = 500$	*PCRTED*	$\hat{\theta} = 1.922;\ \hat{\rho} = 4.033$	238.213
	ZI-PCRTED	$\hat{\theta} = 3.752;\ \hat{\rho} = 5.471;\ \hat{\tau} = -5.372$	261.363
	Poisson	$\hat{\theta} = 0.200$	302.900
	ZI-Poisson	$\hat{\theta} = 1.440;\ \hat{\tau} = 0.861$	236.515
	Neg. Bin.	$\hat{\theta} = 0.108;\ \hat{\rho} = 0.350$	240.191
	ZI-Neg. Bin.	$\hat{\theta} = 0.022;\ \hat{\rho} = 0.321;\ \hat{\tau} = -2.541$	266.676
$n = 1000$	*PCRTED*	$\hat{\theta} = 2.061;\ \hat{\rho} = 3.918$	462.265
	ZI-PCRTED	$\hat{\theta} = 2.500;\ \hat{\rho} = 4.501;\ \hat{\tau} = 0.311$	543.589
	Poisson	$\hat{\theta} = 0.193$	587.124
	ZI-Poisson	$\hat{\theta} = 1.372;\ \hat{\tau} = 0.859$	464.867
	Neg. Bin.	$\hat{\theta} = 0.112;\ \hat{\rho} = 0.367$	473.291
	ZI-Neg. Bin.	$\hat{\theta} = 0.066;\ \hat{\rho} = 0.351;\ \hat{\tau} = 0.325$	541.587

When 50% of simulated data is zero, Table 4 reveals that the new proposition (PCRTED) provides the best fit in most cases across different sample sizes. The ZI-Poisson has a relatively better fit for small samples. It is also observed that the natural forms of the PCRTED and Neg. Bin. provide better fits than their respective zero-inflated forms.

Table 5 shows various statistics when the simulated data have a 70% proportion of zero counts. The ZI-Poisson has a relatively better fit for small samples. It is also observed that the natural forms of the PCRTED and Neg. Bin. provide better fits than their respective zero-inflated forms.

Table 5 shows various statistics when the simulated data have 70% proportion of zero counts. The PCRTED, ZI-Poisson, and Neg. Bin. have relatively better fits than other competing distributions.

When 90% of the simulated dataset are zero counts, the ZI-Poisson better fits smaller samples, while the PCRTED best fits larger samples, as shown in Table 6.

Generally, the –LL statistics increase as the sample size increases for a different proportion of zero counts, while it reduces as the proportion of zero counts increases.

6. Applications

Four sets of claim frequency from different countries are assessed for model comparisons in the study. The first dataset represents claim frequency from automobile injuries from the General Insurance Association of Singapore in 1993 (Frees, 2010; Frees & Valdez, 2008). The second observation is the frequency of third-party claims for Australian vehicle owners (De Jong & Heller, 2008). The third dataset is the insurance claim from Belgium in 1993 (Denuit, 1997; Zamani & Ismail, 2014). The last set of observations considered in the study represents the claim frequency of 10,814 policyholders for the automobile portfolio in a Turkish insurance company between 2012 and 2014 (Meytrianti *et al.*, 2019).

All datasets considered are positively skewed and dispersed with varying proportions of zero counts that often characterize the frequency of claims in actuarial science, as shown in Table 7. This is suggestive that the classical Poisson distribution may not provide an adequate fit.

Table 7: Descriptive Statistics of the Claim Frequency

Dataset	*Dispersion Index*	*Skewness*	*Kurtosis*	*% of Zero*
Dataset I	1.09	4.27	20.88	93.49
Dataset II	1.06	4.07	18.50	93.19
Dataset III	1.09	3.52	14.59	90.33
Dataset IV	1.26	2.56	7.71	79.01

7. Results and Discussion

The results from analyzing the four datasets assuming the new proposition (and zero-inflated form) and the competing distributions are presented in Tables 8 to 11.

Table 8: Parameter estimates for claim frequency from Singapore (dataset I)

Obser.	*Freq.*	*PCRTED*	*ZI-PCRTED*	*Poisson*	*ZI-Poisson*	*Neg. Bin.*	*ZI-Neg. Bin.*
0	6996	6996.26	7000.79	6977.80	6995.92	6996.71	7006.80
1	455	452.95	449.63	487.75	452.76	452.52	444.59
2	28	31.42	31.25	17.05	32.69	31.38	28.85
3	4	2.20	1.30	0.40	1.57	2.22	2.49
	$\hat{\theta}$	13.52	58.41	0.07	0.14	0.87	0.01
	$\hat{\tau}$	0.33	5.68		0.52	0.93	0.87
	$\hat{\rho}$		-75.60				-73.71
–LL		1931.40	1939.36	1941.18	1933.17	1932.38	1938.20
Chi-Square		1.05	5.98	41.96	4.43	1.79	1.80

From Table 8, the new proposition (PCRTED) provides the best fit judging by the smallest –LL and chi-square statistics values. It is also observed that the PCRTED and Neg. Bin. provide better fit to the datasets when compared with their respective zero-inflated forms.

Table 9: Parameter estimates for claim frequency from Australia (dataset II)

Obser.	*Freq.*	*PCRTED*	*ZI-PCRTED*	*Poisson*	*ZI-Poisson*	*Neg. Bin.*	*ZI-Neg. Bin.*
0	63232	63232.31	63227.36	63091.61	63230.49	63230.60	63317.89
1	4333	4330.09	4321.38	4593.07	4325.83	4330.57	4252.49
2	271	275.14	298.19	167.19	286.59	276.48	261.98
3	18	17.29	8.91	4.06	12.66	17.22	21.45
4	2	1.09	0.15	0.07	0.42	1.06	1.97
	$\hat{\theta}$	14.62	85.73	0.07	0.13	1.16	0.01
	$\hat{\tau}$	-0.38	5.79		0.45	0.94	0.88
	$\hat{\rho}$		-179.55				-76.06
–LL		18048.63	18081.29	18101.50	18052.20	18049.68	18105.58
Chi-Square		0.75	34.56	177.66	9.07	0.98	2.51

Table 9 shows that PCRTED is the best fit for the dataset. The negative binomial distribution follows this, while the Poisson distribution does not fit well.

Table 10: Parameter estimates for claim frequency from Belgium (dataset III)

Obser.	*Freq.*	*PCRTED*	*ZI-PCRTED*	*Poisson*	*ZI-Poisson*	*Neg. Bin.*	*ZI-Neg. Bin.*
0	57178	57180.35	57196.18	56949.76	57177.48	57188.34	57249.63
1	5617	5594.29	5584.82	6019.59	5584.80	5581.31	5558.90
2	446	480.24	497.42	318.14	504.87	485.28	438.37
3	50	40.40	20.12	11.21	30.43	40.47	45.91
4	8	3.42	0.46	0.30	1.38	3.30	5.40
	$\hat{\theta}$	10.52	63.36	0.11	0.18	1.28	0.01
	$\hat{\tau}$	-0.67	5.78		0.42	0.92	0.84
	$\hat{\rho}$		-179.11				-71.13
–LL		22063.75	22123.32	22150.54	22075.30	22064.31	22136.57
Chi-Square		1.97	174.23	413.84	51.55	12.33	2.44

The PCRTED also best fits the third dataset with the least value of *–LL* (22063.75) and the chi-square statistic (1.97) from Table 10. The performance of the PCRTED is significantly better than its zero-inflated form (ZI-PCRTED).

Table 11: Parameter estimates for claim frequency from Turkey (dataset IV)

Obser.	*Freq.*	*PCRTED*	*ZI-PCRTED*	*Poisson*	*ZI-Poisson*	*Neg. Bin.*	*ZI-Neg. Bin.*
0	8544	8544.62	8544.30	8292.42	8544.19	8543.47	8561.78
1	1796	1793.77	1765.70	2201.64	1759.23	1795.62	1807.66
2	370	375.96	428.09	292.27	430.75	375.71	331.89
3	81	78.77	69.38	25.87	70.31	78.50	81.03
4	23	16.50	6.07	1.72	8.61	16.39	22.23
	$\hat{\theta}$	3.77	16.19	0.27	0.49	1.01	0.01
	$\hat{\tau}$	-0.01	5.67		0.46	0.79	0.63
	$\hat{\rho}$		-76.17				-82.43
–LL		7028.72	7066.88	7153.16	7038.91	7029.71	7057.06
Chi-Square		2.72	57.60	484.41	35.02	2.84	4.51

With the lowest –LL (7028.72) and chi-square (2.72), the PCRTED best fits the fourth dataset. Both PCRTED and Neg. Bin. provide a better fit for the dataset than their zero-inflated versions.

8. Conclusion

This study proposed a new continuous lifetime distribution with positive support using a one-parameter cubic transmutation map to extend the exponential distribution. The new distribution is assumed for the Poisson parameter in the mixed Poisson paradigm. A new mixed Poisson distribution (the Poisson Cubic Rank Transmuted Exponential Distribution, PCRTED) is proposed along with its zero-inflated version. Various moment-based mathematical properties of the PCRTED are obtained. The shapes of the new discrete distribution proposed are similar to that of the continuous mixing distribution.

With a focus on the frequency of claims in insurance, datasets with varying proportions of zero counts are simulated at different sample sizes. The performance of the new proposition is compared with the classical Poisson and negative binomial distributions (with their zero-inflated forms). The real-life application of the new proposition is assessed on claim frequency from different countries. Results show that the new proposition better fits various datasets than competing distributions using both –LL and chi-square goodness of fit statistics as the selection criteria. It is also found that the natural form of the new distribution outperforms its zero-inflated version in many cases despite having observations with higher-than-expected frequencies of zero counts.

Data Availability

The details of the data used for this study have been discussed in the article.

Declaration of Interest

The authors have no conflicts of interest to report.

References

Adcock, C., Eling, M., & Loperfido, N., (2015). Skewed distributions in finance and actuarial science: a review. *European Journal of Finance*, *21*(13–14), pp. 1253–1281. https://doi.org/10.1080/1351847X.2012.720269

Adetunji, A. A., Sabri, S. R. M., (2021). Modelling Claim Frequency in Insurance Using Count Models. *Asian Journal of Probability and Statistics*, *14*(4), pp. 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*. https://doi.org/10.1007/s42519-019-0059-2

Al-Awadhi, S. A., Ghitany, M. E., (2001). Statistical Properties of Poisson–Lomax Distribution and its Application to Repeated Accidents Data. *Journal of Applied Statistical Science*, *10*(4), pp. 365–372.

Al-kadim, K. A., (2018). Proposed Generalized Formula for Transmuted Distribution. *Journal of Babylon University, Pure and Applied Sciences*, *26*(4), pp. 66–74.

Alexander, C., Cordeiro, G. M., Ortega, E. M. M., & Sarabia, J. M., (2012). Generalized Beta-Generated Distributions. *Computational Statistics & Data Analysis*, *56*, pp. 1880–1897.

Aslam, M., Hussain, Z., & Asghar, Z., (2018). Cubic Transmuted-G family of distributions and its properties. *Stochastic and Quality Control, De Gruyte*, *33*(2), pp. 103–112. https://doi.org/10.1515/eqc-2017-0027

Bhati, D., Kumawat, P., & Gómez–Déniz, E., (2017). A New Count Model Generated from Mixed Poisson Transmuted Exponential Family with an application to Health Care Data. *Communications in Statistics - Theory and Methods*, *46*(22), pp. 11060–11076. https://doi.org/10.1080/03610926.2016.1257712

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), pp. 35–51. https://doi.org/10.17713/ajs.v44i4.54

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), pp. 16468–16477. http://www.ripublication.com

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

Denuit, M., (1997). A New Distribution of Poisson-Type for the Number of Claims. *ASTIN Bulletin*, *27*(2), pp. 229–242. https://doi.org/10.2143/AST.27.2.542049

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), pp. 1457–1469. https://doi.org/10.1198/016214508000000823

Gómez-Déniz, E., Calderín-Ojeda, E., (2016). The Poisson-Conjugate Lindley Mixture Distribution. *Communications in Statistics - Theory and Methods*, *45*(10), pp. 2857–2872. https://doi.org/10.1080/03610926.2014.892134

Granzotto, D. C. T., Louzada, F., & Balakrishnan, N., (2017). Cubic Rank Transmuted Distributions: Inferential Issues and Applications. *Journal of Statistical Computation and Simulation*, *87*(14), pp. 2760–2778. https://doi.org/10.1080/00949655.2017.1344239

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), p. 255. https://doi.org/10.2307/2341080

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

Mahdavi, A., Kundu, D., (2017). A new Method for Generating Distributions with an Application to Exponential Distribution. *Communications in Statistics - Theory and Methods*, *46*(13), pp. 6543–6557. https://doi.org/10.1080/03610926. 2015.1130839

Mahmoudi, E., Zakerzadeh, H., (2010). Generalized Poisson-Lindley Distribution. *Communications in Statistics - Theory and Methods*, *39*(10), pp. 1785–1798. https://doi.org/10.1080/03610920902898514

Meytrianti, A., Nurrohmah, S., & Novita, M., (2019). An Alternative Distribution for Modelling Overdispersion Count Data: Poisson Shanker Distribution. *ICSA - International Conference on Statistics and Analytics 2019*, *1*, pp. 108–120. https://doi.org/10.29244/icsa.2019.pp108-120

Nash, J. C., Varadhan, R., & Grothendieck, G., (2019). *optimr Package: A Replacement and Extension of the "optim" Function*.

Nikoloulopoulos, A. K., & Karlis, D., (2008). On modeling count data: A comparison of some well-known discrete distributions. *Journal of Statistical Computation and Simulation*, *78*(3), pp. 437–457. https://doi.org/10.1080/10629360601010760

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), pp. 137–160. https://doi.org/10.4236/jmf.2018.81012

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

Rahman, M. M., Al-Zahrani, B., Shahbaz, S. H., & Shahbaz, M. Q., (2019). Cubic Transmuted Uniform Distribution: An Alternative to Beta and Kumaraswamy Distributions. *European Journal of Pure and Applied Mathematics*, *12*(3), 1106–1121. https://doi.org/10.29020/nybg.ejpam.v12i3.3410

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), pp. 383–398.

Shaw, W. T., Buckley, I. R. C., (2007). The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *Research Report*.

Shehata, W. A. M., Yousof, H., & Aboraya, M., (2021). A Novel Generator of Continuous Probability Distributions for the Asymmetric Left-skewed Bimodal Real-life Data with Properties and Copulas. *Pakistan Journal of Statistics and Operation Research*, *17*(4), pp. 943–961. https://doi.org/10.18187/pjsor.v17i4.3903

Yang, Y., Tian, W., & Tong, T., (2021). Generalized Mixtures of Exponential Distribution and Associated Inference. *Mathematics*, *9*(12), pp. 1–22. https://doi.org/10.3390/math9121371

Yousof, H. M., Afify, A. Z., Alizadeh, M., Butt, N. S., Hamedani, G. . G., & Ali, M. M., (2015). The Transmuted Exponentiated Generalized-G Family of Distributions. *Pakistan Journal of Statistics and Operation Researchakis*, *11*(4), pp. 441–464. https://doi.org/10.18187/pjsor.v11i4.1164

Zamani, H., Ismail, N., (2014). Functional form for the Zero Inflated Generalized Poisson Regression Model. *Communications in Statistics - Theory and Methods*, *43*, pp. 515–529.