BorneoScience
The Journal of Science & Technology
Volume 44 Issue 2 : September 2023
Online ISSN : 2231-9085 | ISSN : 1394-4339

# ZERO-INFLATED POISSON TRANSMUTED WEIGHTED EXPONENTIAL DISTRIBUTION: PROPERTIES AND APPLICATIONS

**[1*]Shamsul Rijal Muhammed Sabri and [1,2]Ademola Abiodun Adetunji**
[1]School of Mathematical Sciences, Universiti Sains Malaysia, Minden, 11800 Penang, Malaysia.
[2]Department of Statistics, Federal Polytechnic, Ile-Oluji, 35101, Nigeria
Corresponding Author: [*]rijal@usm.my

Received 17th March 2023; accepted 12th September 2023
Available online 30th November 2023
Doi: https://doi.org/10.51200/bsj.v44i2.4702

**ABSTRACT.** *Count observations with high frequencies of zero counts abound in diverse fields. In actuary science, for example, insurance claims are often underreported, leading to a higher frequency of zero counts. This invariably reduces the mean and leads to over-dispersion. Different techniques to model such occurrences exist. This study uses the cubic rank transmutation map to compound the weighted exponential distribution and obtain a new count distribution in the mixed Poisson paradigm. The zero-inflated form of the proposition is obtained along with some mathematical properties. Simulated skewness, kurtosis, and dispersion index for the new distributions reveal they are suitable for model dispersed observation with positive skewness. Five datasets with high frequencies of zero counts are used to assess the performance of the new distribution along with some popular count distributions by using the maximum likelihood for parameter estimation. Results show that the natural form of the new proposition performs creditably better than its zero-inflated forms, even when there is a higher proportion of zero counts. The classical negative binomial distribution is also observed to outperform its zero-inflated form in most cases. In contrast, the zero-inflated Poisson distribution better fits the datasets than the classical Poisson distribution.*

**KEYWORDS:** Mixed Poisson distribution, mixing distribution, rank transmutation map, weighted exponential distribution, excess zero counts.

## INTRODUCTION

Results obtained from count data modelling may be misleading if there are too many or too few zero counts. Zero-deflation arises when the zero frequency in a dataset is lower than the expected frequency, while zero inflation occurs when there are too many zero counts. The former is rare, while the latter is more observed (Conceição *et al*., 2017). The Poisson distribution is often considered for modelling count data (Tajuddin & Ismail, 2022; Wagh & Kamalja, 2017). The distribution assumes equality of mean and variance for observations (Ong *et al*., 2021). Observations with a relatively higher frequency of zero counts are usually dispersed (Sellers & Raim, 2016). Since most count data are dispersed ( Adetunji &

Sabri, 2021), there is model misspecification when the classical Poisson distribution is assumed for dispersed observation (Asamoah, 2016). In order to overcome challenges that characterize the Poisson distribution, several methods that can handle dispersed count observations with excess zero have been proposed (Das *et al*., 2018; Ong *et al*., 2021). One of the techniques often utilized for modelling such observation is the mixed Poisson, first proposed in the early 20th century (Greenwood & Yule, 1920) when the gamma distribution is assumed for the parameter of the classical Poisson distribution resulting in the negative binomial distribution. The process of obtaining a mixed Poisson distribution involves assuming a continuous distribution (called the mixing distribution) with positive supports for the parameter of the Poisson distribution.

Several choices of mixing distributions have been proposed to improve the flexibility and general applicability of the mixed Poisson paradigm. A detailed survey on the choice of the mixing distribution is provided in Ong *et al*. (2021), while Karlis & Xekalaki (2005) gave several properties of the mixed Poisson distribution.

In this study, the cubic rank transmutation map (Rahman *et al*., 2019) is used to extend the weighted exponential distribution (Gupta & Kundu, 2009) and obtain a new mixing distribution assumed for the parameter of the Poisson distribution. The mixing distribution is used to obtain a new mixed Poisson distribution and its zero-inflated form.

Zero-inflated distribution is often used to model observations with several zeros by assigning an extra probability to zero occurrences (Lambert, 1992). The distribution is applied when the frequency of excess zeros is assumed to have come from two processes; the first, where zero counts are obtained by chance like the ones, twos, etc., and the second are obtained when some data are constrained to be zeros (Lambert, 1992; Shahmandi *et al*., 2020). The claim frequency in actuary science is an example of such modelling since observed zeros could have been from two scenarios. A policyholder may have no claim (no case of an accident) reflecting a true zero. Also, there are situations when the policyholder may be involved in a minor accident and hence, may have no urge to report such for a claim, mainly due to the usually cumbersome processes and procedures involved in getting the claim.

## MATERIALS AND METHOD

### Weighted Exponential Distribution

The weighted exponential distribution (Gupta & Kundu, 2009) with the CDF defined in equation (1) as:

$$G(t) = 1 - \frac{1}{a} e^{-\theta t}\left(a + 1 - e^{-a\theta t}\right) \qquad (1)$$

Gupta and Kundu (2009) showed that the shape of equation (1) is identical to that of other two-parameter continuous distributions like generalized exponential, Weibull, and gamma, hence can be used as their alternative. The 3-parameter form of the distribution was proposed by Altun (2019). Zamani *et al*. (2014) used the mixing distribution to obtain a new mixed Poisson distribution with applications in claim frequency.

## Cubic Rank Transmutation Map

Several extensions of different baseline distributions have been proposed to improve flexibility and general applicability. Among many recently proposed techniques are: the Quadratic Transmutation (QT) map (Shaw & Buckley, 2007); and Cubic Rank Transmutations (CRT) map (Al-kadim, 2018; Aslam *et al*., 2018; Rahman *et al*., 2019). Given a baseline distribution with distribution function $G(t)$, the cubic rank transmutation map (Rahman *et al*., 2019) has the CDF given in equation (2) as:

$$F(t) = (1-p)G(t) + 3p\big(G(t)\big)^2 - 2p\big(G(t)\big)^3 \tag{2}$$

Inserting equation (1) into equation (2) gives the distribution function for the Cubic Transmuted Weighted Exponential Distribution (CTWED) in equation (3) as:

$$F(t) = \left(\frac{(a+1)e^{-\theta t} - e^{-(1+a)\theta t} - a}{a^3}\right)\left(pae^{-(1+a)\theta t} - 4p(a+1)e^{-(2+a)\theta t} + 2pe^{-2(1+a)\theta t} - pa(a+1)e^{-\theta t} + 2p(a+1)^2 e^{-2\theta t} - a^2\right) \tag{3}$$

The corresponding PDF is obtained in equation (4) as:

$$f(t) = \frac{\theta(a+1)}{a^3}\big[-6p(2a+3)e^{-(2a+3)\theta t} + 6p(a+1)(a+3)e^{-(a+3)\theta t} + a^2(p-1)e^{-(a+1)\theta t} - 6pa(a+2)e^{-(a+2)\theta t} + 6pe^{-3(a+1)\theta t} + 6pae^{-2(a+1)\theta t} - 6p(a+1)^2\, e^{-3\theta t} + 6pa(a+1)e^{-2\theta t} - a^2(p-1)e^{-\theta t}\big] \tag{4}$$

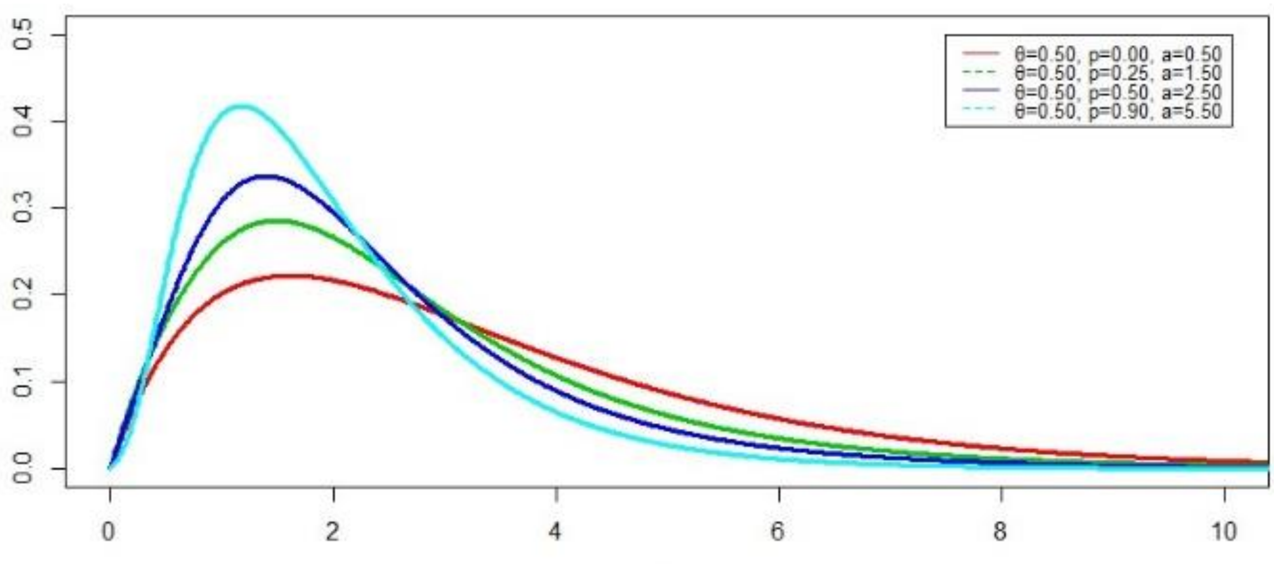


**Figure 1**. Shapes of the PDF of the CTWED

Figure 1 shows that the PDF of the CTWED is positively skewed and unimodal.

## Moment and Moment Generating Function

**Proposition 1:** If a random variable *T* has a CTWED, the r*th* moment is obtained in equation (5) as:

$$E(t^r) = \left(\frac{(a+1)r!}{a^3\theta^r}\right)\left(\frac{-6p}{(2a+3)^r} + \frac{6p(a+1)}{(a+3)^r} + \frac{a^2(p-1)}{(a+1)^{r+1}} - \frac{6pa}{(a+2)^r} + \frac{2p}{3^r(a+1)^{r+1}} + \frac{3pa}{2^r(a+1)^{r+1}} - \frac{2p(a+1)^2}{3^r} + \frac{3pa(a+1)}{2^r} - a^2(p-1)\right) \quad (5)$$

*Proof*

$$E(t^r) = \int_0^\infty t^r f(t)\, dt$$

$$= \int_0^\infty t^r \frac{\theta(a+1)}{a^3}\left[-6p(2a+3)e^{-(2a+3)\theta t} + 6p(a+1)(a+3)e^{-(a+3)\theta t} + a^2(p-1)e^{-(a+1)\theta t} - 6pa(a+2)e^{-(a+2)\theta t} + 6pe^{-3(a+1)\theta t} + 6pae^{-2(a+1)\theta t} - 6p(a+1)^2e^{-3\theta t} + 6pa(a+1)e^{-2\theta t} - a^2(p-1)e^{-\theta t}\right] dt$$

$$= \frac{\theta(a+1)}{a^3}\left[-6p(2a+3)\int_0^\infty t^r e^{-(2a+3)\theta t}dt + 6p(a+1)(a+3)\int_0^\infty t^r e^{-(a+3)\theta t}dt + a^2(p-1)\int_0^\infty t^r e^{-(a+1)\theta t}dt - 6pa(a+2)\int_0^\infty t^r e^{-(a+2)\theta t}dt + 6p\int_0^\infty t^r e^{-3(a+1)\theta t}dt + 6pa\int_0^\infty t^r e^{-2(a+1)\theta t}dt - 6p(a+1)^2\int_0^\infty t^r e^{-3\theta t}dt + 6pa(a+1)\int_0^\infty t^r e^{-2\theta t}dt - a^2(p-1)\int_0^\infty t^r e^{-\theta t}dt\right]$$

$$= \left(\frac{\theta(a+1)r!}{a^3\theta^{r+1}}\right)\left(\frac{-6p(2a+3)}{(2a+3)^{r+1}} + \frac{6p(a+1)(a+3)}{(a+3)^{r+1}} + \frac{a^2(p-1)}{(a+1)^{r+1}} - \frac{6pa(a+2)}{(a+2)^{r+1}} + \frac{6p}{3^{r+1}(a+1)^{r+1}} + \frac{6pa}{2^{r+1}(a+1)^{r+1}} - \frac{6p(a+1)^2}{3^{r+1}} + \frac{6pa(a+1)}{2^{r+1}} - a^2(p-1)\right)$$

$$= \left(\frac{(a+1)r!}{a^3\theta^r}\right)\left(\frac{-6p}{(2a+3)^r} + \frac{6p(a+1)}{(a+3)^r} + \frac{a^2(p-1)}{(a+1)^{r+1}} - \frac{6pa}{(a+2)^r} + \frac{2p}{3^r(a+1)^{r+1}} + \frac{3pa}{2^r(a+1)^{r+1}} - \frac{2p(a+1)^2}{3^r} + \frac{3pa(a+1)}{2^r} - a^2(p-1)\right)$$

**Proposition 2**: If a random variable *T* has a CTWED, the MGF is obtained in equation (6) as:

$$E(e^{zt}) = \frac{\theta(a+1)}{a^3}\left(\frac{-6p(2a+3)}{2a\theta+3\theta-z} + \frac{6p(a+1)(a+3)}{a\theta+3\theta-z} + \frac{a^2(p-1)}{a\theta+\theta-z} - \frac{6pa(a+2)}{a\theta+2\theta-z} + \frac{6p}{3a\theta+3\theta-z} + \frac{6pa}{2a\theta+2\theta-z} - \frac{6p(a+1)^2}{3\theta-z} + \frac{6pa(a+1)}{2\theta-z} - \frac{a^2(p-1)}{\theta-z}\right) \quad (6)$$

*Proof*

$$E(e^{zt}) = \int_0^\infty e^{zt} f(t)\; dt$$

$$= \int_0^\infty e^{zt}\frac{\theta(a+1)}{a}\left[-6p(2a+3)e^{-(2a+3)\theta t} + 6p(a+1)(a+3)e^{-(a+3)\theta t} + a^2(p-1)e^{-(a+1)\theta t} - 6pa(a+2)e^{-(a+2)\theta t} + 6pe^{-3(a+1)\theta t} + 6pae^{-2(a+1)\theta t} - 6p(a+1)^2e^{-3\theta t} + 6pa(a+1)e^{-2\theta t} - a^2(p-1)e^{-\theta t}\right]\; dt$$

$$= \frac{\theta(a+1)}{a^3}\int_0^\infty \left(-6p(2a+3)e^{-(2a\theta+3\theta-z)t} + 6p(a+1)(a+3)e^{-(a\theta+3\theta-z)t} + a^2(p-1)e^{-(a\theta+\theta-z)t} - 6pa(a+2)e^{-(a\theta+2\theta-z)t} + 6pe^{-(3a\theta+3\theta-z)t} + 6pae^{-(2a\theta+2\theta-z)t} - 6p(a+1)^2\, e^{-(3\theta-z)t} + 6pa(a+1)e^{-(2\theta-z)t} - a^2(p-1)e^{-(\theta-z)t}\right)dt$$

$$= \frac{\theta(a+1)}{a^3}\left(\frac{-6p(2a+3)}{2a\theta+3\theta-z} + \frac{6p(a+1)(a+3)}{a\theta+3\theta-z} + \frac{a^2(p-1)}{a\theta+\theta-z} - \frac{6pa(a+2)}{a\theta+2\theta-z} + \frac{6p}{3a\theta+3\theta-z} + \frac{6pa}{2a\theta+2\theta-z} - \frac{6p(a+1)^2}{3\theta-z} + \frac{6pa(a+1)}{2\theta-z} - \frac{a^2(p-1)}{\theta-z}\right)$$

## Mixed Poisson CTWED

**Proposition 3:** Given that a random variable $X \sim Poisson(T)$ and $T \sim CTWED\ (\theta, p, a)$. A discrete random variable *X* has a mixed Poisson-CTWED (PCTWED) if its PMF is defined in equation (7) as:

$$P_x = \frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta)^{x+1}} - \frac{6p(2a+3)}{(1+2a\theta+3\theta)^{x+1}} + \frac{a^2(p-1)}{(1+a\theta+\theta)^{a+1}} - \frac{6pa(a+2)}{(1+a\theta+2\theta)^{x+1}} + \frac{6p}{(1+3a\theta+3\theta)^{x+1}} + \frac{6pa}{(1+2a\theta+2\theta)^{x+1}} - \frac{6p(a+1)^2}{(1+3\theta)^{x+1}} + \frac{6pa(a+1)}{(1+2\theta)^{x+1}} - \frac{a^2(p-1)}{(1+\theta)^{x+1}}\right) \quad (7)$$

*Proof*

$$P_x = \int_0^\infty \frac{t^x e^{-t}}{x!}\frac{\theta(\alpha+1)}{\alpha^3}\left[-6p(2a+3)e^{-(2a+3)\theta t} + 6p(a+1)(a+3)e^{-(a+3)\theta t} + a^2(p-1)e^{-(a+1)\theta t} - 6pa(a+2)e^{-(a+2)\theta t} + 6pe^{-3(a+1)\theta t} + 6pae^{-2(a+1)\theta t} - 6p(a+1)^2\, e^{-3\theta t} + 6pa(a+1)e^{-2\theta t} - a^2(p-1)e^{-\theta t}\right] dt$$

$$= \frac{\theta(a+1)}{a^3 x!}\int_0^\infty t^x\left(6p(a+1)(a+3)e^{-(1+a\theta+3\theta)t} - 6p(2a+3)e^{-(1+2a\theta+3\theta)t} + a^2(p-1)e^{-(1+a\theta+\theta)t} - 6pa(a+2)e^{-(1+a\theta+2\theta)t} + 6pe^{-(1+3a\theta+3\theta)t} + 6pae^{-(1+2a\theta+2\theta)t} - 6p(a+1)^2\, e^{-(1+3\theta)t} + 6pa(a+1)e^{-(1+2\theta)t} - a^2(p-1)e^{-(1+\theta)t}\right) dt$$

$$= \frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta)^{x+1}} - \frac{6p(2a+3)}{(1+2a\theta+3\theta)^{x+1}} + \frac{a^2(p-1)}{(1+a\theta+\theta)^{a+1}} - \frac{6pa(a+2)}{(1+a\theta+2\theta)^{x+1}} + \frac{6p}{(1+3a\theta+3\theta)^{x+1}} + \frac{6pa}{(1+2a\theta+2\theta)^{x+1}} - \frac{6p(a+1)^2}{(1+3\theta)^{x+1}} + \frac{6pa(a+1)}{(1+2\theta)^{x+1}} - \frac{a^2(p-1)}{(1+\theta)^{x+1}}\right)$$

**Sub-Model:** Equation (7) becomes the PMF of the mixed Poisson Weighted Exponential Distribution (Zamani *et al*., 2014) when the $p = 0$.

Figure 2 shows different shapes of the PCTWED for different parameter values. The figure reveals that the distribution is suitable for unimodal distribution and observations with excess zeros, resembling the shapes of the PDF of the CTWED in Figure 1.

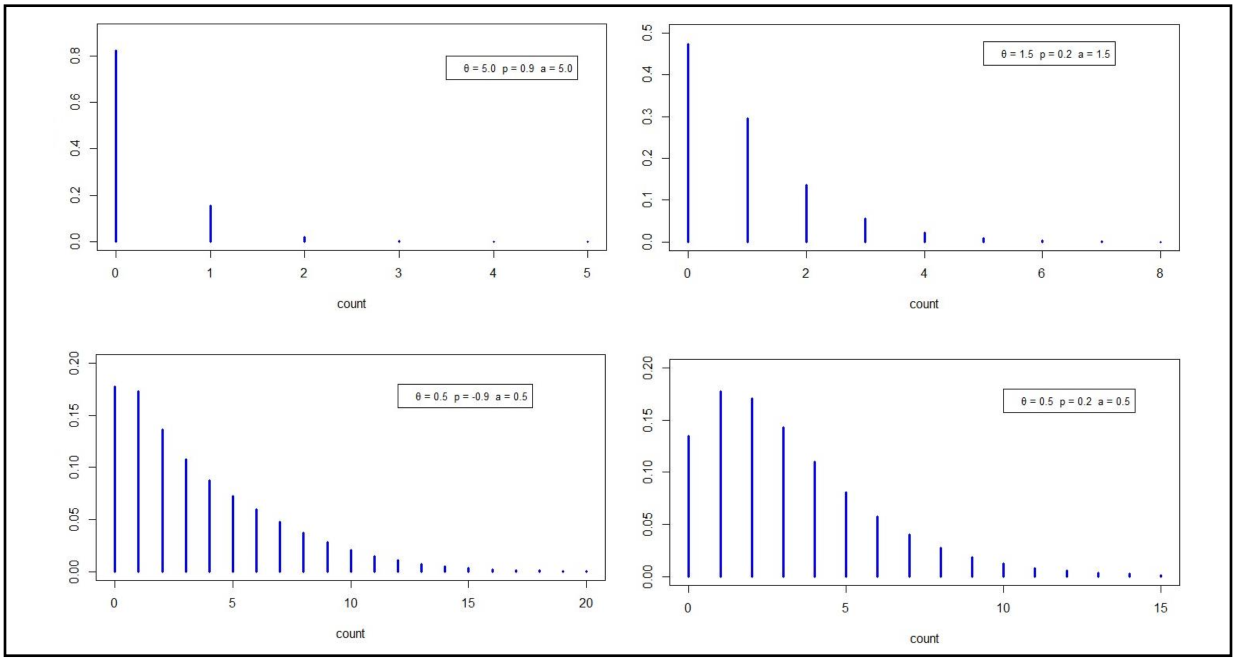


**Figure 2.** Shapes of the PMF of the PCTWED

## The CDF of the PCTWED

**Proposition 4:** The CDF of a random variable $X$ with the PCTWED is defined in equation (8) as:

$$F(x) = 1 - \left(\frac{6p(a+1)^2}{a^3(1+a\theta+3\theta)^{x+1}} - \frac{6p(a+1)}{a^3(1+2a\theta+3\theta)^{x+1}} + \frac{(p-1)}{a(1+a\theta+\theta)^{x+1}} - \frac{6p(a+1)}{a^2(1+a\theta+2\theta)^{x+1}} + \frac{2p}{a^3(1+3a\theta+3\theta)^{x+1}} + \frac{3p}{a^2(1+2a\theta+2\theta)^{x+1}} - \frac{2p(a+1)^3}{a^3(1+3\theta)^{x+1}} + \frac{3p(a+1)^2}{a^2(1+2\theta)^{x+1}} - \frac{(p-1)(a+1)}{a(1+\theta)^{x+1}}\right) \quad (8)$$

*Proof*

$$F(X) = P(X \le x)$$

$$= 1 - P(X > x)$$

$$= 1 - \sum_{k=x+1}^{\infty} P_k$$

$$= 1 - \sum_{k=x+1}^{\infty} \frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta)^{k+1}} - \frac{6p(2a+3)}{(1+2a\theta+3\theta)^{k+1}} + \frac{a^2(p-1)}{(1+a\theta+\theta)^{k+1}} - \frac{6pa(a+2)}{(1+a\theta+2\theta)^{k+1}} + \frac{6p}{(1+3a\theta+3\theta)^{k+1}} + \frac{6pa}{(1+2a\theta+2\theta)^{k+1}} - \frac{6p(a+1)^2}{(1+3\theta)^{k+1}} + \frac{6pa(a+1)}{(1+2\theta)^{k+1}} - \frac{a^2(p-1)}{(1+\theta)^{k+1}}\right)$$

$$= 1 - \frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{\theta(a+3)(1+a\theta+3\theta)^{x+1}} - \frac{6p(2a+3)}{\theta(2a+3)(1+2a\theta+3\theta)^{x+1}} + \frac{a^2(p-1)}{\theta(a+1)(1+a\theta+\theta)^{x+1}} - \frac{6pa(a+2)}{\theta(a+2)(1+a\theta+2\theta)^{x+1}} + \frac{6p}{3\theta(a+1)(1+3a\theta+3\theta)^{x+1}} + \frac{6pa}{2\theta(a+1)(1+2a\theta+2\theta)^{x+1}} - \frac{6p(a+1)^2}{3\theta(1+3\theta)^{x+1}} + \frac{6pa(a+1)}{2\theta(1+2\theta)^{x+1}} - \frac{a^2(p-1)}{2\theta(1+\theta)^{x+1}}\right)$$

**Zero-Inflated Poisson**

$$= 1 - \left(\frac{6p(a+1)^2}{a^3(1+a\theta+3\theta)^{x+1}} - \frac{6p(a+1)}{a^3(1+2a\theta+3\theta)^{x+1}} + \frac{(p-1)}{a(1+a\theta+\theta)^{x+1}} - \frac{6p(a+1)}{a^2(1+a\theta+2\theta)^{x+1}} + \frac{2p}{a^3(1+3a\theta+3\theta)^{x+1}} + \frac{3p}{a^2(1+2a\theta+2\theta)^{x+1}} - \frac{2p(a+1)^3}{a^3(1+3\theta)^{x+1}} + \frac{3p(a+1)^2}{a^2(1+2\theta)^{x+1}} - \frac{(p-1)(a+1)}{a(1+\theta)^{x+1}}\right)$$

## Mathematical Properties of the PCTWED

**Proposition 5:** If $f(t)$ is the PDF of the mixing distribution for a discrete random $X$, the Probability Generating Function (PGF) is obtained as:

$$P_x(z) = \int_0^\infty e^{t(z-1)} f(t)dt$$

$$= \int_0^\infty e^{t(z-1)} \frac{\theta(a+1)}{a^3}\Big[6p(a+1)(a+3)e^{-(a+3)\theta t} - 6p(2a+3)e^{-(2a+3)\theta t} + a^2(p-1)e^{-(a+1)\theta t} - 6pa(a+2)e^{-(a+2)\theta t} + 6pe^{-3(a+1)\theta t} + 6pae^{-2(a+1)\theta t} - 6p(a+1)^2 e^{-3\theta t} + 6pa(a+1)e^{-2\theta t} - a^2(p-1)e^{-\theta t}\Big]dt$$

$$= \frac{\theta(a+1)}{a^3}\int_0^\infty \Big(6p(a+1)(a+3)e^{-(1+a\theta+3\theta-z)t} - 6p(2a+3)e^{-(1+2a\theta+3\theta-z)t} + a^2(p-1)e^{-(1+a\theta+\theta-z)t} - 6pa(a+2)e^{-(1+a\theta+2\theta-z)t} + 6pe^{-(1+3a\theta+3\theta-z)t} + 6pae^{-(1+2a\theta+\theta-z)t} - 6p(a+1)^2 e^{-(1+3\theta-z)t} + 6pa(a+1)e^{-(1+2\theta-z)t} - a^2(p-1)e^{-(1+\theta-z)t}\Big)dt$$

Therefore, the PGF is obtained in equation (9) as

$$P_x(z) = \frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta-z)} - \frac{6p(2a+3)}{(1+2a\theta+3\theta-z)} + \frac{a^2(p-1)}{(1+a\theta+\theta-z)} - \frac{6pa(a+2)}{(1+a\theta+2\theta-z)} + \frac{6p}{(1+3a\theta+3\theta-z)} + \frac{6pa}{(1+2a\theta+\theta-z)} - \frac{6p(a+1)^2}{(1+3\theta-z)} + \frac{6pa(a+1)}{(1+2\theta-z)} - \frac{a^2(p-1)}{(1+\theta-z)}\right) \quad (9)$$

The Moment Generating Function (MGF) is obtained in equation (10) by replacing ***z*** with $e^z$ in equation (9).

$$M_X(z) = \frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta-e^z)} - \frac{6p(2a+3)}{(1+2a\theta+3\theta-e^z)} + \frac{a^2(p-1)}{(1+a\theta+\theta-e^z)} - \frac{6pa(a+2)}{(1+a\theta+2\theta-e^z)} + \frac{6p}{(1+3a\theta+3\theta-e^z)} + \frac{6pa}{(1+2a\theta+\theta-e^z)} - \frac{6p(a+1)^2}{(1+3\theta-e^z)} + \frac{6pa(a+1)}{(1+2\theta-e^z)} - \frac{a^2(p-1)}{(1+\theta-e^z)}\right) \quad (10)$$

The first four central moments are obtained in equations (11) - (14) as:

$$E(X) = \frac{(12-2p)a^4+(102-15p)a^3+(318-34p)a^2+(432-38p)a+216-19p}{6\theta(a+1)(a+2)(a+3)(2a+3)} \quad (11)$$

$$E(X^2)=\left(\frac{\theta(1+a)}{a^3}\right)\left(\frac{12p(a+1)(a+3)}{(a\theta+3\theta)^3}+\frac{6p(a+1)(a+3)}{(a\theta+3\theta)^2}-\frac{12p(2a+3)}{(2a\theta+3\theta)^3}-\frac{6p(2a+3)}{(2a\theta+3\theta)^2}+\frac{2a^2(p-1)}{(a\theta+\theta)^3}+\frac{a^2(p-1)}{(a\theta+\theta)^2}-\frac{12pa(a+2)}{(a\theta+2\theta)^3}-\frac{6pa(a+2)}{(a\theta+2\theta)^2}+\frac{12p}{(3a\theta+3\theta)^3}+\frac{6p}{(3a\theta+3\theta)^2}+\frac{12pa}{(2a\theta+2\theta)^3}+\frac{6pa}{(2a\theta+2\theta)^2}-\frac{8p(1+a)^2-27pa(1+a)+36a^2(p-1)}{18\theta^3}-\frac{4p(1+a)^2-9pa(1+a)+6a^2(p-1)}{6\theta^2}\right) \quad (12)$$

$$E(X^3)=\left(\frac{\theta(a+1)}{a^3}\right)\left(\frac{36p}{(3a\theta+3\theta)^4}+\frac{36p}{(3a\theta+3\theta)^3}+\frac{6p}{(3a\theta+3\theta)^2}-\frac{36p(2a+3)}{(2a\theta+3\theta)^4}-\frac{36p(2a+3)}{(2a\theta+3\theta)^3}-\frac{6p(2a+3)}{(2a\theta+3\theta)^2}+\frac{36pa}{(2a\theta+2\theta)^4}+\frac{36pa}{(2a\theta+2\theta)^3}+\frac{6pa}{(2a\theta+2\theta)^2}+\frac{36p(a+1)(a+3)}{(a\theta+3\theta)^4}+\frac{36p(a+1)(a+3)}{(a\theta+3\theta)^3}+\frac{6p(a+1)(a+3)}{(a\theta+3\theta)^2}-\frac{36pa(a+2)}{(a\theta+2\theta)^4}-\frac{36pa(a+2)}{(a\theta+2\theta)^3}-\frac{6pa(a+2)}{(a\theta+2\theta)^2}+\frac{6a^2(p-1)}{(a\theta+\theta)^4}+\frac{6a^2(p-1)}{(a\theta+\theta)^3}+\frac{a^2(p-1)}{(a\theta+\theta)^2}-\frac{216a^2(p-1)-81pa(a+1)+16p(a+1)^2}{36\theta^4}-\frac{36a^2(p-1)-27pa(a+1)+8p(a+1)^2}{6\theta^3}-\frac{4p(a+1)^2-9pa(a+1)+6a^2(p-1)}{6\theta^2}\right) \quad (13)$$

$$E(X^4)=\left(\frac{\theta(a+1)}{a^3}\right)\left(\frac{144p}{(3a\theta+3\theta)^5}+\frac{216p}{(3a\theta+3\theta)^4}+\frac{84p}{(3a\theta+3\theta)^3}+\frac{6p}{(3a\theta+3\theta)^2}-\frac{144p(2a+3)}{(2a\theta+3\theta)^5}-\frac{216p(2a+3)}{(2a\theta+3\theta)^4}-\frac{84p(2a+3)}{(2a\theta+3\theta)^3}-\frac{6p(2a+3)}{(2a\theta+3\theta)^2}+\frac{144pa}{(2a\theta+2\theta)^5}+\frac{216pa}{(2a\theta+2\theta)^4}+\frac{84pa}{(2a\theta+2\theta)^3}+\frac{6pa}{(2a\theta+2\theta)^2}+\frac{144p(a+1)(a+3)}{(a\theta+3\theta)^5}+\frac{216p(a+1)(a+3)}{(a\theta+3\theta)^4}+\frac{84p(a+1)(a+3)}{(a\theta+3\theta)^3}+\frac{6p(a+1)(a+3)}{(a\theta+3\theta)^2}-\frac{144pa(a+2)}{(a\theta+2\theta)^5}-\frac{216pa(a+2)}{(a\theta+2\theta)^4}-\frac{84pa(a+2)}{(a\theta+2\theta)^3}-\frac{6pa(a+2)}{(a\theta+2\theta)^2}+\frac{24a^2(p-1)}{(a\theta+\theta)^5}+\frac{36a^2(p-1)}{(a\theta+\theta)^4}+\frac{14a^2(p-1)}{(a\theta+\theta)^3}+\frac{a^2(p-1)}{(a\theta+\theta)^2}-\frac{32p(a+1)^2+1296a^2(p-1)-243pa(a+1)}{54\theta^5}-\frac{16p(a+1)^2+216a^2(p-1)-81pa(a+1)}{6\theta^4}-\frac{56p(a+1)^2+252a^2(p-1)-189pa(a+1)}{18\theta^3}-\frac{4p(a+1)^2+6a^2(a-1)-9pa(a+1)}{6\theta^2}\right) \quad (14)$$

## Skewness and Kurtosis

The skewness and kurtosis for the PCTWED are obtained from the central moments (De Jong & Heller, 2008) as:

$$S_k=\frac{E(X^3)-3E(X^2)E(X)+2\left(E(X)\right)^3}{\left(Var(X)\right)^{\frac{3}{2}}}$$

$$Kurt=\frac{E(X^4)-4E(X^3)E(X)+6E(X^2)\left(E(X)\right)^2-3\left(E(X)\right)^4}{\left(Var(X)\right)^2}$$

Tables 1 – 3 show simulated Skewness, Kurtosis, and Dispersion Index for some parameters distribution.

**Table 1.** Skewness for some parameters of the PCTWED

| | $a=0.5$ | | | $a=2.5$ | | | $a=7.5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\theta$ =0.1 | $\theta$ =2.0 | $\theta$ =10.0 | $\theta$ =0.1 | $\theta$ =2.0 | $\theta$ =10.0 | $\theta$ =0.1 | $\theta$ =2.0 | $\theta$ =10.0 |
| *p = -0.9* | 1.68 | 2.44 | 4.06 | 1.76 | 3.08 | 5.82 | 1.94 | 4.27 | 8.80 |
| *p = -0.5* | 1.82 | 2.48 | 4.15 | 1.89 | 3.13 | 5.97 | 2.03 | 4.36 | 9.04 |
| *p = 0.0* | 2.00 | 2.50 | 4.25 | 2.02 | 3.18 | 6.17 | 2.10 | 4.48 | 9.38 |
| *p = 0.5* | 2.10 | 2.46 | 4.35 | 2.05 | 3.20 | 6.38 | 2.06 | 4.59 | 9.76 |
| *p = 0.9* | 1.77 | 2.34 | 4.41 | 1.74 | 3.18 | 6.55 | 1.80 | 4.67 | 10.09 |

**Table 2.** Kurtosis for some parameters of the PCTWED

| | $a = 0.5$ | | | $a = 2.5$ | | | $a = 7.5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\theta$ =0.1 | $\theta$ =2.0 | $\theta$ =10.0 | $\theta$ =0.1 | $\theta$ =2.0 | $\theta$ =10.0 | $\theta$ =0.1 | $\theta$ =2.0 | $\theta$ =10.0 |
| $p$ = -0.9 | 6.60 | 10.61 | 21.61 | 7.06 | 14.46 | 39.12 | 7.96 | 23.38 | 82.65 |
| $p$ = -0.5 | 7.56 | 10.98 | 22.30 | 7.90 | 14.83 | 40.80 | 8.62 | 24.15 | 86.98 |
| $p$ = 0.0 | 9.02 | 11.25 | 23.06 | 9.09 | 15.13 | 43.03 | 9.39 | 25.06 | 93.01 |
| $p$ = 0.5 | 10.40 | 10.97 | 23.59 | 9.93 | 15.01 | 45.33 | 9.59 | 25.76 | 99.84 |
| $p$ = 0.9 | 8.72 | 9.82 | 23.68 | 8.17 | 14.34 | 47.16 | 7.96 | 26.02 | 105.97 |

**Table 3.** Dispersion Index for some parameters of the PCTWED

| | $a = 0.5$ | | | $a = 2.5$ | | | $a = 7.5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\theta$ =0.1 | $\theta$ =2.0 | $\theta$ =10.0 | $\theta$ =0.1 | $\theta$ =2.0 | $\theta$ =10.0 | $\theta$ =0.1 | $\theta$ =2.0 | $\theta$ =10.0 |
| $p$ = -0.9 | 9.86 | 1.44 | 1.09 | 4.80 | 1.19 | 1.04 | 2.56 | 1.08 | 1.02 |
| $p$ = -0.5 | 8.99 | 1.40 | 1.08 | 4.42 | 1.17 | 1.03 | 2.41 | 1.07 | 1.01 |
| $p$ = 0.0 | 7.67 | 1.33 | 1.07 | 3.86 | 1.14 | 1.03 | 2.18 | 1.06 | 1.01 |
| $p$ = 0.5 | 6.00 | 1.25 | 1.05 | 3.14 | 1.11 | 1.02 | 1.88 | 1.04 | 1.01 |
| $p$ = 0.9 | 4.35 | 1.17 | 1.03 | 2.44 | 1.07 | 1.01 | 1.59 | 1.03 | 1.01 |

**Remarks:**

For fixed $p$ and $a$, skewness and kurtosis increase while the dispersion index decreases as $\theta$ increases. When $p$ and $\theta$ are fixed, skewness and kurtosis increase while the dispersion index reduces as $a$ increases. For fixed $a$ and $\theta$, skewness, kurtosis increase while the dispersion index decreases as $p$ increases in most cases. Both skewness and kurtosis peak at $p = 0.5$.

## Maximum Likelihood Estimation of the PCTWED

Given random samples of size $n$ drawn from the PCTWED with $(\theta, p, a)$ as defined in equation (7), the log-likelihood function for the distribution is obtained in equation (15) as:

$$\ell = \sum_{i=1}^{n} log\left(\frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta)^{x+1}} - \frac{6p(2a+3)}{(1+2a\theta+3\theta)^{x+1}} + \frac{a^2(p-1)}{(1+a\theta+\theta)^{x+1}} - \frac{6pa(a+2)}{(1+a\theta+2\theta)^{x+1}} + \frac{6p}{(1+3a\theta+3\theta)^{x+1}} + \frac{6pa}{(1+2a\theta+2\theta)^{x+1}} - \frac{6p(a+1)^2}{(1+3\theta)^{x+1}} + \frac{6pa(a+1)}{(1+2\theta)^{x+1}} - \frac{a^2(p-1)}{(1+\theta)^{x+1}}\right)\right) \quad (15)$$

Estimators for $(\theta, p, a)$ denoted with $(\hat{\theta}, \hat{p}, \hat{a})$ are the solutions for the log-likelihood function. The equations form a non-linear equation system that can only be solved numerically. This research uses the optimr function (Nash *et al.*, 2019) in the **R language** (R-Core Team, 2020) is used.

## Zero-Inflated PCTWED

**Proposition 6.** If a discrete random variable $X$ has a PCTWED with PMF $P_x$ and its realization at $x = 0$ as $P_0$. If the zero-inflation parameter is denoted with $\pi$, then a discrete random variable $X_Z$ has a zero-inflated PCTWED if its PMF is defined in equation (16) as:

$$P_x^Z = \begin{cases} \pi + (1-\pi)P_0, & x = 0 \\ (1-\pi)P_x, & x = 1,2,3,\ldots \end{cases} \qquad (16)$$

where

$$P_x = \frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta)^{x+1}} - \frac{6p(2a+3)}{(1+2a\theta+3\theta)^{x+1}} + \frac{a^2(p-1)}{(1+a\theta+\theta)^{a+1}} - \frac{6pa(a+2)}{(1+a\theta+2\theta)^{x+1}} + \frac{6p}{(1+3a\theta+3\theta)^{x+1}} + \frac{6pa}{(1+2a\theta+2\theta)^{x+1}} - \frac{6p(a+1)^2}{(1+3\theta)^{x+1}} + \frac{6pa(a+1)}{(1+2\theta)^{x+1}} - \frac{a^2(p-1)}{(1+\theta)^{x+1}}\right)$$

and

$$P_0 = \frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta)} - \frac{6p(2a+3)}{(1+2a\theta+3\theta)} + \frac{a^2(p-1)}{(1+a\theta+\theta)} - \frac{6pa(a+2)}{(1+a\theta+2\theta)} + \frac{6p}{(1+3a\theta+3\theta)} + \frac{6pa}{(1+2a\theta+2\theta)} - \frac{6p(a+1)^2}{(1+3\theta)} + \frac{6pa(a+1)}{(1+2\theta)} - \frac{a^2(p-1)}{(1+\theta)}\right)$$

*Proof*

The proof is obtained by appropriate substitution. Hence the result.

## Mathematical Properties of the Zero-Inflated PCTWED

If the PGF of the PCTWED is denoted with $P_x(z)$, then the PGF of the Zero-Inflated PCTWED denoted by $P_x^Z(z)$ is obtained using: $P_x^Z(z) = (1-\pi)P_x(z)$

Hence, the PGF of the Zero-Inflated PCTWED is given in equation (17) as:

$$P_x^Z(z) = (1-\pi)\left(\frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta-z)} - \frac{6p(2a+3)}{(1+2a\theta+3\theta-z)} + \frac{a^2(p-1)}{(1+a\theta+\theta-z)} - \frac{6pa(a+2)}{(1+a\theta+2\theta-z)} + \frac{6p}{(1+3a\theta+3\theta-z)} + \frac{6pa}{(1+2a\theta+\theta-z)} - \frac{6p(a+1)^2}{(1+3\theta-z)} + \frac{6pa(a+1)}{(1+2\theta-z)} - \frac{a^2(p-1)}{(1+\theta-z)}\right)\right) \qquad (17)$$

The corresponding MGF is therefore expressed in equation (18) as:

$$M_x^Z(z) = (1-\pi)\left(\frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta-e^t)} - \frac{6p(2a+3)}{(1+2a\theta+3\theta-e^t)} + \frac{a^2(p-1)}{(1+a\theta+\theta-e^t)} - \frac{6pa(a+2)}{(1+a\theta+2\theta-e^t)} + \frac{6p}{(1+3a\theta+3\theta-e^t)} + \frac{6pa}{(1+2a\theta+\theta-e^t)} - \frac{6p(a+1)^2}{(1+3\theta-e^t)} + \frac{6pa(a+1)}{(1+2\theta-z)} - \frac{a^2(p-1)}{(1+\theta-e^t)}\right)\right) \qquad (18)$$

If the $r^{th}$ raw moment of the PCTWED is denoted by $E(X^r)$, then the $r^{th}$ raw moment of the zero-inflated PCTWED is generally defined as:

$$m_r = E(X_Z^r) = (1-\pi)E(X^r)$$

Given the first four raw moments of the PCTWED as $E(X)$, $E(X^2)$, $E(X^3)$, and $E(X^4)$ as obtained in equations (11) to (14), the first four raw moments of the zero-inflated PCTWED are given in equations (19) to (22).

$$m_1 = (1-\pi)E(X) \qquad (19)$$
$$m_2 = (1-\pi)E(X^2) \qquad (20)$$
$$m_3 = (1-\pi)E(X^3) \qquad (21)$$
$$m_4 = (1-\pi)E(X^4) \qquad (22)$$

The variance, dispersion index, skewness, and kurtosis are obtained respectively from equations (19) to (22) as:

$$Var(X_Z) = m_2 - [m_1]^2$$
$$DI = \frac{Var(X_Z)}{m_1}$$
$$S_k = \frac{m_3 - 3m_2m_1 + 2(m_1)^3}{\left(Var(X_Z)\right)^{\frac{3}{2}}}$$
$$kurt = \frac{m_4 - 4m_3m_1 + 6m_2(m_1)^2 - 3(m_1)^4}{\left(Var(X_Z)\right)^2}$$

## Maximum Likelihood Estimation of the Parameter of Zero-Inflated PCTWED

If a random variable $X$ is assumed to follow the Zero-Inflated PCTWED with PMF $P_x^Z$ indexed with parameter $(\theta, p, a, \pi)$, then the distribution parameters can be estimated using the MLE method. The likelihood function is defined as:

$$\mathcal{L}(\theta, p, a, \pi) = \prod_{n_0}(\pi + (1-\pi)P_0)\prod_{n_1}\left((1-\pi)P(x > 0)\right)$$

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

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

$$= n_0 \ln\left(\pi + (1-\pi)\left(\frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta)} - \frac{6p(2a+3)}{(1+2a\theta+3\theta)} + \frac{a^2(p-1)}{(1+a\theta+\theta)} - \frac{6pa(a+2)}{(1+a\theta+2\theta)} + \frac{6p}{(1+3a\theta+3\theta)} + \frac{6pa}{(1+2a\theta+2\theta)} - \frac{6p(a+1)^2}{(1+3\theta)} + \frac{6pa(a+1)}{(1+2\theta)} - \frac{a^2(p-1)}{(1+\theta)}\right)\right)\right) + n_1 \ln(1-\pi) + \sum_{n_1} \ln\left(\frac{\theta(a+1)}{a^3}\left(\frac{6p(a+1)(a+3)}{(1+a\theta+3\theta)^{x+1}} - \frac{6p(2a+3)}{(1+2a\theta+3\theta)^{x+1}} + \frac{a^2(p-1)}{(1+a\theta+\theta)^{x+1}} - \frac{6pa(a+2)}{(1+a\theta+2\theta)^{x+1}} + \frac{6p}{(1+3a\theta+3\theta)^{x+1}} + \frac{6pa}{(1+2a\theta+2\theta)^{x+1}} - \frac{6p(a+1)^2}{(1+3\theta)^{x+1}} + \frac{6pa(a+1)}{(1+2\theta)^{x+1}} - \frac{a^2(p-1)}{(1+\theta)^{x+1}}\right)\right)$$

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

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

The MLE for parameters space $(\theta, p, a, \pi)$ are obtained numerically by solving $\frac{\partial \ell}{\partial \pi} = 0$, $\frac{\partial \ell}{\partial \theta} = 0$, $\frac{\partial \ell}{\partial p} = 0$, and $\frac{\partial \ell}{\partial a} = 0$. This is done using different algorithms that come with the **optimr** package (Nash *et al.*, 2019) in **R language** (R-Core Team, 2020).

## APPLICATIONS

### Data

Five count datasets characterized by many zeros are considered to assess the performance of the PCTWED and zero-inflated PCTWED. The new propositions are compared with the Poisson and negative binomial distributions (along with their respective zero-inflated forms). Dataset I consists of the number of claims on motorcycle insurance from WASA (a Swedish Insurance Firm) from 1994 – 1998. The data has been previously utilized (Omari *et al.*, 2018) on count distributions. The second dataset is the number of automobile insurance policies in Australia between 2004 and 2005, as previously presented in De Jong and Heller (2008). Dataset III comprises of frequency of claim insurance in a Belgium firm in 1993. The data was used to assess claim distributions (Zamani *et al.*, 2014). The fourth dataset is the frequency of claims of 10, 814 policyholders of a Turkish insurance firm between 2012 and 2014. The data were assessed on Poisson-related distributions (Meytrianti *et al*., 2019). The fifth dataset is the yeast cell counts per square, as previously examined on Poison Lindley distribution (Shanker & Hagos, 2015). All five datasets are dispersed and positively skewed, with very high percentages of zero counts (table 4).

**Table 4.** Summary of Datasets

| X | Dataset I | Dataset II | Dataset III | Dataset IV | Dataset V |
|---|---|---|---|---|---|
| 0 | 63878 | 63232 | 57178 | 8544 | 128 |
| 1 | 643 | 4333 | 5617 | 1796 | 37 |
| 2 | 27 | 271 | 446 | 370 | 18 |
| 3 | | 18 | 50 | 81 | 3 |
| 4 | | 2 | 8 | 23 | 1 |
| % of 0 | 98.96 | 93.19 | 90.33 | 79.01 | 68.45 |
| Dispersion Index | 1.07 | 1.08 | 1.08 | 1.26 | 1.32 |
| Kurtosis | 118.26 | 18.50 | 14.59 | 7.71 | 2.80 |
| Skewness | 10.46 | 4.07 | 3.52 | 2.56 | 1.75 |

## RESULTS AND DISCUSSION

With the least values of the chi-square statistic and – LL, the proposed distribution (PCTWED) provide the best fit to the first dataset (table 5) while its zero-inflated form (ZI-PCTWED) performs worst. It is also observed that the negative binomial distribution (another mixed Poisson distribution) performs better than its zero-inflated form, while the ZIP performs better than the Poisson distribution.

For the second dataset (table 6), the PCTWED also gives the best fit with the lowest values of both – LL and chi-square statistic. Like the first data, the ZiNB also performs better than the classical negative binomial distribution, while the ZIP performs better than the Poisson distribution.

Table 7 shows that the ZiNB best fits the third dataset; the PCTWED follows this. The ZI-PCTWED gives the worst fit to the data with the highest values of both – LL and chi-square.

Tables 8 and 9 show that the new proposition (PCTWED) best fits both datasets IV and V. The negative binomial provides a better fit than its zero-inflated form, while the ZIP gives a better fit than the Poisson distribution for both datasets. Dataset V has the lowest percentage of zero counts among the five datasets assessed. The ZI-PCTWED gives a relatively better fit in this dataset than the negative binomial distribution and its zero-inflated form.

**Table 5.** Data I (Swedish Claim Frequency)

| X | Freq. | PCTWED | ZI-PCTWED | Poisson | ZIP | Neg. Bin. | ZiNB |
|---|---|---|---|---|---|---|---|
| 0 | 63878 | 63878.13 | 63877.74 | 63854.63 | 63878.11 | 63877.73 | 63893.56 |
| 1 | 643 | 642.83 | 632.94 | 689.63 | 643.61 | 644.62 | 629.97 |
| 2 | 27 | 26.76 | 37.38 | 3.72 | 25.58 | 24.43 | 23.26 |
| *Estimates* | $\theta$ | 89.097 | 504.897 | 0.011 | 0.080 | 0.154 | 0.003 |
| | $p$ | -4.759 | -0.761 | | 0.864 | 0.934 | 0.926 |
| | $a$ | -4.926 | -3.590 | | | | -40.710 |
| | $\pi$ | | -54.232 | | | | |
| | *– LL* | 3840.44 | 3853.79 | 3872.00 | 3840.88 | 3841.50 | 3857.32 |
| | *Chi-Square* | 0.00 | 3.04 | 148.64 | 0.08 | 0.27 | 0.88 |

**Table 6.** Data II (Australian Claim Frequency)

| X | Freq. | PCTWED | ZI-PCTWED | Poisson | ZIP | Neg. Bin. | ZiNB |
|---|---|---|---|---|---|---|---|
| 0 | 63232 | 63230.90 | 63239.55 | 63091.61 | 63230.49 | 63230.60 | 63317.89 |
| 1 | 4333 | 4332.84 | 4310.28 | 4593.07 | 4325.83 | 4330.57 | 4252.49 |
| 2 | 271 | 273.98 | 305.84 | 167.19 | 286.59 | 276.48 | 261.98 |
| 3 | 18 | 17.14 | 0.33 | 4.06 | 12.66 | 17.22 | 21.45 |
| 4 | 2 | 1.07 | 0.00 | 0.07 | 0.42 | 1.06 | 1.97 |
| *Estimates* | $\theta$ | 15.167 | 387.325 | 0.073 | 0.133 | 1.157 | 0.007 |
| | $p$ | -0.045 | -5.445 | | 0.451 | 0.941 | 0.878 |
| | $a$ | 9.406 | -4.253 | | | | -76.060 |
| | $\pi$ | | -135.921 | | | | |
| | *– LL* | 18049.64 | 18148.72 | 18101.50 | 18052.20 | 18049.68 | 18105.58 |
| | *Chi-Square* | 0.89 | 2585.95 | 177.66 | 9.07 | 0.98 | 48.15 |

**Table 7.** Data III (Belgium Claim Frequency)

| X | Freq. | PCTWED | ZI-PCTWED | Poisson | ZIP | Neg. Bin. | ZiNB |
|---|---|---|---|---|---|---|---|
| 0 | 57178 | 57178.64 | 57187.37 | 56949.763 | 57177.48 | 57188.34 | 57249.63 |
| 1 | 5617 | 5598.70 | 5587.71 | 6019.590 | 5584.80 | 5581.31 | 5558.90 |
| 2 | 446 | 477.08 | 523.00 | 318.135 | 504.87 | 485.28 | 438.37 |
| 3 | 50 | 40.67 | 0.91 | 11.209 | 30.43 | 40.47 | 45.91 |
| 4 | 8 | 3.57 | 0.01 | 0.296 | 1.38 | 3.30 | 5.40 |
| *Estimates* | $\theta$ | 9.573 | 271.157 | 0.106 | 0.181 | 1.279 | 0.008 |
| | $p$ | 0.354 | -6.343 | | 0.415 | 0.924 | 0.843 |
| | $a$ | 13.223 | -4.312 | | | | -71.130 |
| | $\pi$ | | -111.862 | | | | |
| | *– LL* | 22063.30 | 22311.68 | 22150.54 | 22075.30 | 22064.31 | 22136.57 |
| | *Chi-Square* | 9.74 | 10122.64 | 413.84 | 51.55 | 12.33 | 2.44 |

**Table 8.** Data IV (Turkish Claim Frequency)

| X | | Freq. | PCTWED | ZI-PCTWED | Poisson | ZIP | Neg. Bin. | ZiNB |
|---|---|---|---|---|---|---|---|---|
| 0 | | 8544 | 8545.14 | 8544.04 | 8292.42 | 8544.19 | 8543.47 | 8561.78 |
| 1 | | 1796 | 1789.69 | 1771.07 | 2201.64 | 1759.23 | 1795.62 | 1807.66 |
| 2 | | 370 | 380.67 | 492.20 | 292.27 | 430.75 | 375.71 | 331.89 |
| 3 | | 81 | 78.92 | 6.34 | 25.87 | 70.31 | 78.50 | 81.03 |
| 4 | | 23 | 15.78 | 0.33 | 1.72 | 8.61 | 16.39 | 22.23 |
| *Estimates* | $\theta$ | | 4.353 | 41.119 | 0.266 | 0.490 | 1.009 | 0.006 |
| | $p$ | | -0.506 | -4.526 | | 0.458 | 0.792 | 0.635 |
| | $a$ | | 12.831 | -4.408 | | | | -82.430 |
| | $\pi$ | | | -19.730 | | | | |
| | | $-LL$ | 7029.56 | 7334.26 | 7153.16 | 7038.91 | 7029.71 | 7057.06 |
| | | *Chi-Square* | 2.68 | 2452.35 | 484.41 | 35.02 | 2.84 | 4.52 |

**Table 9.** Data V (Yeast cell counts per square)

| X | | Freq. | PCTWED | ZI-PCTWED | Poisson | ZIP | Neg. Bin. | ZiNB |
|---|---|---|---|---|---|---|---|---|
| 0 | | 128 | 127.31 | 127.48 | 118.06 | 128.00 | 126.73 | 180.68 |
| 1 | | 37 | 38.36 | 35.34 | 54.30 | 38.35 | 42.08 | 4.51 |
| 2 | | 18 | 16.72 | 21.26 | 12.49 | 15.49 | 12.84 | 1.15 |
| 3 | | 3 | 3.86 | 2.51 | 1.91 | 4.17 | 3.80 | 0.39 |
| 4 | | 1 | 0.65 | 0.35 | 0.22 | 0.84 | 1.11 | 0.15 |
| *Estimates* | $\theta$ | | 9.003 | 97.851 | 0.4609 | 0.808 | 1.195 | 0.004 |
| | $p$ | | -12.457 | -0.232 | | 0.431 | 0.722 | 0.490 |
| | $a$ | | 0.013 | -1.467 | | | | -11.720 |
| | $\pi$ | | | -31.216 | | | | |
| | | $-LL$ | 168.60 | 170.31 | 173.83 | 168.80 | 170.02 | 172.05 |
| | | *Chi-Square* | 0.53 | 1.92 | 12.16 | 0.81 | 2.88 | 517.31 |

## CONCLUSION

The skewness, kurtosis, and dispersion index for some parameter combinations are presented to assess the behaviour of the distributions. The shapes of the PCTWED show that it can be effectively utilized to model observations with an unusual frequency of zeros.

Performances of the new propositions are assessed on five count observations with varying percentages of zero counts. Comparisons are made with the Poisson and negative binomial distributions (along with the irrespective zero-inflated forms). The maximum likelihood estimation using different algorithms that come with the optimr package in R-language is used to provide estimates for the parameters of the distributions. The chi-square goodness of fits and the – LL are used for model selection. Results show that the PCTWED outperforms its zero-inflated form in particular and all the competing distributions in most cases. The classical negative binomial distribution provides a better fit for datasets with above 70% of zero counts than its zero-inflated form. In contrast, the zero-inflated Poisson outperforms the Poisson distribution in all cases.

The finding shows that mixed Poisson distributions and the negative binomial distribution (when the gamma distribution is used as the nixing distribution) are naturally suitable to model observation with higher than usual zero counts in count observation. Therefore, it is unnecessary to obtain their zero-inflated form to model observations with many zero counts.

## 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.

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

Altun, E. 2019. Weighted-Exponential Regression Model: An Alternative to the Gamma Regression Model. *International Journal of Modeling, Simulation, and Scientific Computing*, **10(6)**, 1–15.

Asamoah, K. 2016. On the Credibility of Insurance Claim Frequency: Generalized Count Models and Parametric Estimators. *Insurance: Mathematics and Economics*, **70**, 339–353.

Aslam, M., Hussain, Z., & Asghar, Z. 2018. Cubic Transmuted-G family of distributions and its properties. *Stochastic and Quality Control, De Gruyte*, **33(2)**, 103–112.

Conceição, K. S., Louzada, F., Andrade, M. G., & Helou, E. S. 2017. Zero-Modified Power Series Distribution and its Hurdle Distribution Version. *Journal of Statistical Computation and Simulation*, **87(9)**, 1842–1862.

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.

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)**, 255.

Gupta, R. D., & Kundu, D. 2009. A New Class of Weighted Exponential Distributions. *Statistics*, **43(6)**, 621–634.

Karlis, D., & Xekalaki, E. 2005. Mixed Poisson distributions. *International Statistical Review*, **73(1)**, 35–58.

Lambert, D. 1992. Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics*, **34**, 1–14.

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**, 108–120.

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

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.

Ong, S. H., Low, Y. C., & Toh, K. K. 2021. Recent Developments in Mixed Poisson Distributions. *ASM Science Journal*, **14(3)**, 1–10.

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.

Sellers, K. F., & Raim, A. 2016. A flexible zero-inflated model to address data dispersion. *Computational Statistics and Data Analysis*, **99**, 68–80.

Shahmandi, M., Wilson, P., & Thelwall. 2020. A New Algorithm for Zero-Modified Models Applied to Citation Counts. *Scientometrics*, **125(2)**, 993–1010.

Shanker, R., & Hagos, F. 2015. On Poisson-Lindley distribution and its Applications to Biological Sciences. *Biometrics and Biostatistics International Journal*, **2(5)**, 1–5.

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.*

Tajuddin, M. R. R., & Ismail, N. 2022. Frequentist and Bayesian Zero-Inflated Regression Models on Insurance Claim Frequency: A comparison study using Malaysia's Motor Insurance Data. *Malaysian Journal of Science*, **41(2)**, 16–29.

Wagh, Y. S., & Kamalja, K. K. 2017. Comparison of methods of estimation for parameters of generalized Poisson distribution through simulation study. *Communications in Statistics: Simulation and Computation*, **46(5)**, 4098–4112.

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.