TRENDS IN SCIENCES 2022; 19(5): 2898 RESEARCH ARTICLE
https://doi.org/10.48048/tis.2022.2898

The Poisson-Transmuted Janardan Distribution for Modelling Count Data

Winai Bodhisuwan[1] and Sirinapa Aryuyuen[2,*]

[1]*Department of Statistics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand*
[2]*Department of Mathematics and Computer Science, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, Pathum Thani 12110, Thailand*

([*]Corresponding author's e-mail: sirinapa_a@rmutt.ac.th)

Received: 10 October 2020, Revised: 17 May 2021, Accepted: 27 May 2021

Abstract

In this paper, we introduce a new mixed Poisson distribution, called the Poisson-transmuted Janardan distribution. The Poisson-Janardan and Poisson-Lindley distributions are sub-model of the proposed distribution. Some mathematical properties of the proposed distribution, including the moments, moment generating function, probability generating function and generation of a Poisson-transmuted Janardan random variable, are presented. The parameter estimation is discussed based on the method of moments and the maximum likelihood estimation. In addition, we illustrated the application of the proposed distribution by fitting with 4 real data sets and comparing it with some other distributions based on the Kolmogorov-Smirnov test for criteria.

Keywords: Mixed Poisson distribution, Poisson-transmuted Janardan distribution, Count data, Over-dispersion, Probability function

Introduction

Count data are used to describe many phenomena in different fields from insurance and economics to biometrics and the social sciences. Count data analysis can use traditional models such as the Poisson distribution to describe the data if the variance and mean are equal. However, in practice, it has often been found that count data exhibit over-dispersion (variance greater than mean). Hence, there is a demand to modify the Poisson model when such problems are encountered. In the last 2 decades, many researchers have developed new distributions to count data analysis. One such method has been wide that has been widely used for observed modeling situations is a mixture of distributions [1,2].

A new mixed Poisson distribution is a mixture of distributions for the count data. A traditional mixed Poisson distribution where the mean of the Poisson variable is distributed as a gamma random variable is the Poisson-gamma (or the negative binomial) distribution, derived by [3]. It has become an increasingly popular alternative distribution to the Poisson distribution. However, the Poisson-gamma distribution may not be appropriate for some over-dispersed count data. Some well-known mixed Poisson distributions are the Poisson-Lindley by [4], the Poisson-inverse Gaussian by [5], the Poisson-Pareto by [6], the Poisson-Lomax by [7], generalized Poisson-Lindley by [8], the Poisson-weight exponential by [9], the Poisson-Janardan by [10], the Poisson-Lindley beta by [11], the Poisson-generalized Lindley by [12], etc. Recently, the Poisson-transmuted exponential was introduced by [1]. The existing mixed Poisson distributions can be used for improving count data modeling.

In this paper, a new mixed Poisson distribution is obtained by mixing the Poisson distribution and the transmuted Janardan (TJ) distribution. The TJ distribution was proposed by [13], which is a continuous distribution. It has various structural properties that are the explicit expressions for the reliability and hazard rate functions, order statistics, the moments and the moment the generating function. The TJ distribution is suggested for several applications in the modeling of real data. We describe the theoretical development of the mixed Poisson distribution. A new mixed Poisson distribution, so-called the Poisson-transmuted Janardan distribution, is proposed. In addition, some properties of the proposed distribution are shown, such as reliability and hazard rate functions, order statistics, moments and moment generating function. The different methods of parameter estimation are derived, i.e., the method of moments and the maximum likelihood estimation. The simulation study of the parameter estimation based on the MME and MLE is shown. Moreover, an application of the proposed distribution is illustrated. Finally, some discussion and conclusions are included.

Methods

Let X be a random variable which is distributed the Poisson distribution with a parameter λ. The probability mass function (pmf) of X, $g_{Pois}(x)$, is;

$$g_{Pois}(x)=\frac{e^{-\lambda}\lambda^{x}}{x!};x=0,1,2,...,\lambda>0. \tag{1}$$

A new mixed Poisson distribution arises from the Poisson distribution when its parameter λ follows the TJ distribution. The probability density function (pdf) of the TJ distribution [13] is;

$$g_{TJ}(\lambda)=\frac{\theta^{2}(\alpha\lambda+1)e^{-2\theta\lambda/\alpha}}{\alpha(\alpha^{2}+\theta)^{2}}\left[2\beta(\alpha^{2}+\theta+\alpha\theta\lambda)-(\beta-1)(\alpha^{2}+\theta)e^{\theta\lambda/\alpha}\right], \tag{2}$$

where $\alpha,\theta>0$ and $-1\le\beta\le1$. For $\beta=0$ the TJ distribution reduces to the Janardan distribution, as proposed by [14]. The cumulative density function (cdf) of the Janardan distribution, $G_{J}(\lambda)$, is;

$$G_{J}(\lambda)=1-\frac{\alpha(\alpha^{2}+\theta)+\alpha^{2}\theta\lambda}{\alpha(\alpha^{2}+\theta)}e^{-\frac{\theta\lambda}{\alpha}}, \tag{3}$$

and the inverse cdf is;

$$G_{J}^{-1}(v)=\frac{\alpha}{\theta}\left\{-\frac{\alpha^{2}+\theta}{\alpha^{2}}-W_{-1}\left[-\frac{\alpha^{2}+\theta}{\alpha^{2}}(1-v)e^{-\frac{\alpha^{2}+\theta}{\alpha^{2}}}\right]\right\}, \tag{4}$$

where v is a value of a uniform random variable on [0,1], and $W_{-1}(.)$ denotes the negative branch of the Lambert W function [15], i.e., $W(z)=e^{W(z)}=z,$ where z is a complex number. We apply the **Lambert W** function in R [16] to solve the $W_{-1}(.)$. Since the TJ distribution is the quadratic rank transmutation. Thus, we obtained the inverse cdf of the TJ distribution by using the cdf of the quadratic rank transmutation [17], that is;

$$G_{T\lambda}(\lambda)=(1+\beta)G_{\lambda}(\lambda)-\beta\left[G_{\lambda}(\lambda)\right]^{2};-1\le\beta\le1,$$

where $G_{\lambda}(\lambda)$ is the cdf of the baseline distribution. Then the inverse $G_{T\lambda}(\lambda)$ is

$$G_{T\lambda}^{-1}(u)=G_{\lambda}^{-1}\left\{\frac{1}{2\beta}\left[1+\beta-\sqrt{(\beta+1)^{2}-4\beta u}\right]\right\},$$

where u is a value of a uniform random variable on [0,1]. When $G_{\lambda}(\lambda)$ is the cdf of the Janardan distribution as in eq. (3), we have the inverse cdf of the TJ distribution as;

$$G_{TJ}^{-1}(u)=G_{J}^{-1}\left\{\frac{1}{2\beta}\left[1+\beta-\sqrt{(\beta+1)^{2}-4\beta u}\right]\right\}$$

$$=\frac{\alpha}{\theta}\left\{-\frac{\alpha^2+\theta}{\alpha^2}-W_{-1}\left\{-\frac{\alpha^2+\theta}{\alpha^2}e^{-\frac{\alpha^2+\theta}{\alpha^2}}\left[1-\frac{1+\beta-\sqrt{(\beta+1)^2-4\beta u}}{2\beta}\right]\right\}\right\}. \tag{5}$$

Results and discussion

A new mixed Poisson distribution

By mixing the Poisson and TJ distributions in **Definition 1**, we obtained a new mixed Poisson distribution, and its pmf is in **Theorem 1**.

Definition 1 Let X be a random variable of the Poisson-transmuted Janardan (PTJ) distribution, will be denoted by $X\sim PTJ(\alpha,\theta,\beta)$, when $X|\lambda$ has distributed the Poisson (Pois) distribution with a positive parameter λ. Where the parameter λ has distributed the TJ distribution with the parameters $\alpha>0$, $\theta>0$ and $-1\le\beta\le 1$, i.e., $X|\lambda\sim Pois(\lambda)$ and $\lambda\sim TJ(\alpha,\theta,\beta)$.

Theorem 1 Let $X\sim PTJ(\alpha,\theta,\beta)$ then the pmf of X is;

$$f_{PTJ}(x)=\frac{\theta^2\alpha^x}{(\alpha^2+\theta)^2}\left\{\frac{2\beta}{(\alpha+2\theta)^{x+1}}\left[\frac{\alpha^2\theta(x+1)}{\alpha+2\theta}\left(1+\frac{\alpha^2(x+2)}{\alpha+2\theta}\right)+(\alpha^2+\theta)\left(1+\frac{\alpha^2(x+1)}{\alpha+2\theta}\right)\right]\right.$$
$$\left.-\frac{(\beta-1)(\alpha^2+\theta)}{(\alpha+\theta)^{x+1}}\left(1+\frac{\alpha^2(x+1)}{\alpha+\theta}\right)\right\}, \tag{6}$$

where $x=0,1,2,\ldots$ and the parameters $\alpha,\theta>0$ and $-1\le\beta\le 1$.

Proof Since $X|\lambda\sim Pois(\lambda)$ and $\lambda\sim TJ(\alpha,\theta,\beta)$, the pmf of the PTJ distribution will be obtained utilization of;

$$f_{PTJ}(x)=\int_0^\infty g_{Pois}(x|\lambda)g_{TJ}(\lambda)d\lambda. \tag{7}$$

By substituting (1) and (2) into (7), we derive the PTJ pmf;

$$f_{PTJ}(x)=\int_0^\infty\frac{e^{-\lambda}\lambda^x}{x!}\frac{\theta^2(\alpha\lambda+1)e^{-2\theta\lambda/\alpha}}{\alpha(\alpha^2+\theta)^2}\left[2\beta(\alpha^2+\theta+\alpha\theta\lambda)-(\beta-1)(\alpha^2+\theta)e^{\theta\lambda/\alpha}\right]d\lambda$$
$$=\frac{\theta^2}{\alpha(\alpha^2+\theta)^2\Gamma(x+1)}\left[2\beta\int_0^\infty(\alpha\lambda^{x+1}+\lambda^x)(\alpha^2+\theta+\alpha\theta\lambda)e^{-\left(\frac{\alpha+2\theta}{\alpha}\right)\lambda}d\lambda\right.$$
$$\left.-(\beta-1)(\alpha^2+\theta)\int_0^\infty(\alpha\lambda^{x+1}+\lambda^x)e^{-\left(\frac{\alpha+\theta}{\alpha}\right)\lambda}d\lambda\right]. \tag{8}$$

Let $u=\left(\frac{\alpha+2\theta}{\alpha}\right)\lambda$ and $v=\left(\frac{\alpha+\theta}{\alpha}\right)\lambda$, we have $d\lambda=\left(\frac{\alpha}{\alpha+2\theta}\right)du$ and $d\lambda=\left(\frac{\alpha}{\alpha+\theta}\right)dv$, respectively. The corresponding pmf in (8) is;

$$f_{PTJ}(x)=\frac{\theta^2}{\alpha(\alpha^2+\theta)^2\Gamma(x+1)}\left[2\beta\left[\alpha(\alpha^2+\theta)\left(\frac{\alpha}{\alpha+2\theta}\right)^{x+2}\Gamma(x+2)+\alpha^2\theta\left(\frac{\alpha}{\alpha+2\theta}\right)^{x+3}\Gamma(x+3)\right.\right.$$
$$\left.+(\alpha^2+\theta)\left(\frac{\alpha}{\alpha+2\theta}\right)^{x+1}\Gamma(x+1)+\alpha\theta\left(\frac{\alpha}{\alpha+2\theta}\right)^{x+2}\Gamma(x+2)\right]$$

$$-(\beta-1)(\alpha^2+\theta)\left[\alpha\left(\frac{\alpha}{\alpha+\theta}\right)^{x+2}\Gamma(x+2)+\left(\frac{\alpha}{\alpha+\theta}\right)^{x+1}\Gamma(x+1)\right]\Bigg\}.$$

Finally, we obtain the pmf of the PTJ distribution as in (6). The pmf of X satisfies the following properties: (i) $f_{PTJ}(x)\geq 0$ and (ii) $\sum_X f_{PTJ}(x)=1$ for all x.

The pmf behavior of the proposed distribution is shown in **Figure 1**. The PTJ pmf has a decreasing function for $\alpha<1$, see **Figures 1(a)** and **1(b)**. If $\alpha\geq 1$ then the pmf of the PTJ distribution has a unimodal distribution, see **Figures 1(c)** and **1(d)**. The PTJ distribution has 2 sub-model as follows.

Corollary 1 Let $X\sim PTJ(\alpha,\theta,\beta)$, if $\beta=0$ then PTJ distribution reduce to the Poisson-Janardan (PJ) distribution, which the PJ distribution is introduced by [10]. The pmf of the PJ distribution is;

$$f_{PJ}(x)=\left(\frac{\theta}{\alpha+\theta}\right)^2\left(\frac{\alpha}{\alpha+\theta}\right)^x\left[1+\frac{\alpha(\alpha x+1)}{\alpha^2+\theta}\right];x=0,1,2,\ldots,\alpha>0,\theta>0. \tag{9}$$

Corollary 2 Let $X\sim PTJ(\alpha,\theta,\beta)$, if $\beta=0$ and $\alpha=1$ then the PTJ distribution reduces to the Poisson-Lindley (PL) distribution, which the PL distribution is introduced by [4]. The pmf of the PL distribution is;

$$f_{PL}(x)=\frac{\theta^2(2+\theta+x)}{(1+\theta)^{x+3}};x=0,1,2,\ldots,\theta>0. \tag{10}$$

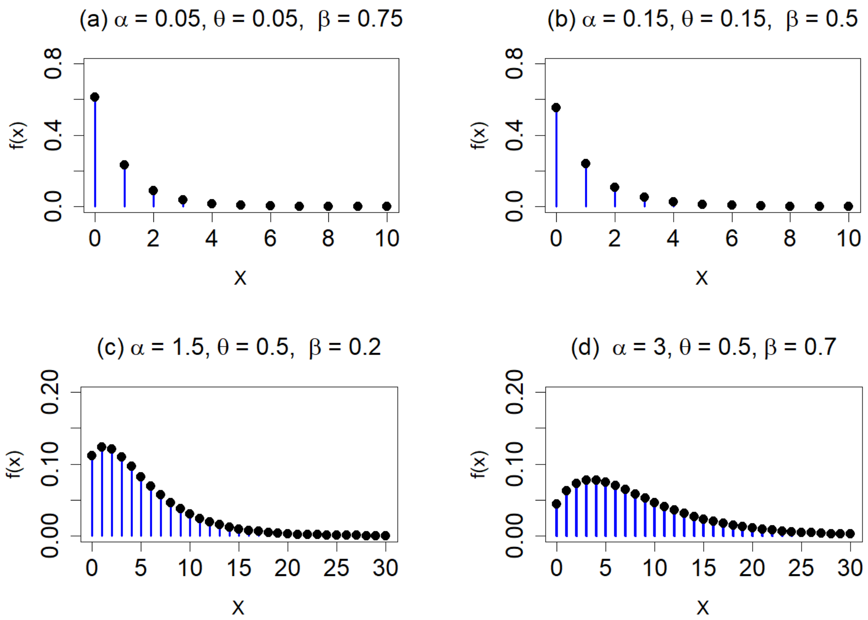

Figure 1 The PJT pmf plots with some specified value of parameters α, θ and β.

Properties of the PTJ distribution

Some properties of the proposed distribution including, the moments, moment generating function, probability generating function and the generating of a PTJ random variable, are introduced.

Proposition 1 If $X\sim PTJ(\alpha,\theta,\beta)$, then the k*th* factorial moment of X is;

$$\mu_k' = \frac{1}{(\alpha^2+\theta)^2}\left(\frac{\alpha}{\theta}\right)^k \left\{(\alpha^2+\theta)\left[\alpha^2\Gamma(k+2)+\theta\Gamma(k+1)\right] - \frac{\beta}{4}\left[\alpha^4\left(4\Gamma(k+2)-\frac{\Gamma(k+3)+2\Gamma(k+2)}{2^k}\right)\right.\right.$$
$$\left.\left. +4\theta^2\Gamma(k+1)\left(1-\frac{1}{2^k}\right)+4\alpha^2\theta\left(\Gamma(k+2)+\Gamma(k+1)-\frac{\Gamma(k+2)+\Gamma(k+1)}{2^k}\right)\right]\right\},$$
(11)

where $k = 1,2,3,\ldots,$ and the parameters $\alpha, \theta > 0$ and $-1 \le \beta \le 1$.

Proof The k*th* moment about the origin of the PTJ distribution can be obtained as;

$$\mu_k' = E\left[E\left(X^k \mid \lambda\right)\right] = \int_0^\infty \left[\sum_{x=0}^\infty \frac{x^k e^{-\lambda}\lambda^x}{\Gamma(x+1)}\right]\frac{\theta^2(\alpha\lambda+1)e^{-\frac{2\theta\lambda}{\alpha}}}{\alpha(\alpha^2+\theta)^2}\left[2\beta(\alpha^2+\theta+\alpha\theta\lambda)-(\beta-1)(\alpha^2+\theta)e^{\frac{\theta\lambda}{\alpha}}\right]d\lambda,$$

since $\sum_{x=0}^\infty \frac{x^k e^{-\lambda}\lambda^x}{\Gamma(x+1)} = \lambda^k$, we obtain;

$$\mu_k' = \frac{\theta^2}{\alpha(\alpha^2+\theta)^2}\left\{2\beta\left[(\alpha^2+\theta)\int_0^\infty(\alpha\lambda^{k+1}+\lambda^k)e^{-\frac{2\theta\lambda}{\alpha}}d\lambda + \alpha\theta\int_0^\infty(\alpha\lambda^{k+2}+\lambda^{k+1})e^{-\frac{2\theta\lambda}{\alpha}}d\lambda\right]\right.$$
$$\left. -(\beta-1)(\alpha^2+\theta)\int_0^\infty(\alpha\lambda^{k+1}+\lambda^k)e^{-\frac{\theta\lambda}{\alpha}}d\lambda\right\}$$
$$= \frac{\theta^2}{\alpha(\alpha^2+\theta)^2}\left\{2\beta\left[\alpha^2\theta\left(\frac{\alpha}{2\theta}\right)^{k+3}\Gamma(k+3)+\alpha(\alpha^2+2\theta)\left(\frac{\alpha}{2\theta}\right)^{k+2}\Gamma(k+2)+(\alpha^2+\theta)\Gamma(k+1)\right.\right.$$
$$\left.\left.\times\left(\frac{\alpha}{2\theta}\right)^{k+1}\right]-(\beta-1)(\alpha^2+\theta)\left[\alpha\left(\frac{\alpha}{\theta}\right)^{k+2}\Gamma(k+2)+\left(\frac{\alpha}{\theta}\right)^{k+1}\Gamma(k+1)\right]\right\}.$$

Finally, the k*th* moment about the origin of the PTJ distribution is expressed as in (11). By taking $k = 1,2,3$ and 4 in (11), the first 4 factorial moments about the origin of the PTJ distribution can be obtained;

$$\mu_1' = \frac{1}{(\alpha^2+\theta)^2}\left(\frac{\alpha}{\theta}\right)\left[\left(\alpha^2+\theta\right)\left(2\alpha^2+\theta\right)-\frac{\beta}{4}\left(3\alpha^4+2\theta^2+6\alpha^2\theta\right)\right], \quad (12)$$

$$\mu_2' = \frac{1}{(\alpha^2+\theta)^2}\left(\frac{\alpha}{\theta}\right)^2\left[\left(\alpha^2+\theta\right)\left(6\alpha^2+2\theta\right)-\frac{\beta}{4}\left(15\alpha^4+6\theta^2+24\alpha^2\theta\right)\right], \quad (13)$$

$$\mu_3' = \frac{1}{(\alpha^2+\theta)^2}\left(\frac{\alpha}{\theta}\right)^3\left[\left(\alpha^2+\theta\right)\left(24\alpha^2+6\theta\right)-\frac{\beta}{4}\left(75\alpha^4+21\theta^2+105\alpha^2\theta\right)\right], \quad (14)$$

$$\mu_4' = \frac{1}{(\alpha^2+\theta)^2}\left(\frac{\alpha}{\theta}\right)^4\left[\left(\alpha^2+\theta\right)\left(120\alpha^2+24\theta\right)-\frac{\beta}{4}\left(420\alpha^4+90\theta^2+540\alpha^2\theta\right)\right]. \quad (15)$$

From the first 4 factorial moments of the PTJ distribution, the mean, variance, index of dispersion (ID), skewness (Sk) and kurtosis (Ku) of X are respectively given by;

$$E(X)=\frac{\alpha}{4\theta(\alpha^2+\theta)^2}\left[4(\alpha^2+\theta)(2\alpha^2+\theta)-\beta(3\alpha^4+2\theta^2+6\alpha^2\theta)\right],$$

$$V(X)=\frac{\alpha^2}{\theta^2(\alpha^2+\theta)^2}\left\{\left[2(\alpha^2+\theta)(3\alpha^2+\theta)-\frac{3\beta}{4}(5\alpha^4+2\theta^2+8\alpha^2\theta)\right]\right.$$
$$\left.-\left[\frac{4(\alpha^2+\theta)(2\alpha^2+\theta)-\beta(3\alpha^4+2\theta^2+6\alpha^2\theta)}{2(\alpha^2+\theta)}\right]^2\right\},$$

$$ID=\frac{\alpha}{\theta}\left[\frac{8(\alpha^2+\theta)(3\alpha^2+\theta)-3\beta(5\alpha^4+2\theta^2+8\alpha^2\theta)}{4(\alpha^2+\theta)(2\alpha^2+\theta)-\beta(3\alpha^4+2\theta^2+6\alpha^2\theta)}-\frac{4(\alpha^2+\theta)(2\alpha^2+\theta)-\beta(3\alpha^4+2\theta^2+6\alpha^2\theta)}{4(\alpha^2+\theta)^2}\right],$$

$$Sk=\frac{\alpha^3}{\theta^3(\alpha^2+\theta)^2\sigma_X^3}\left\{\left[6(\alpha^2+\theta)(4\alpha^2+\theta)-\frac{3\beta}{4}(25\alpha^4+7\theta^2+35\alpha^2\theta)\right]-\frac{3}{(\alpha^2+\theta)^2}\left[2(\alpha^2+\theta)\right.\right.$$
$$\left.\times(3\alpha^2+\theta)-\frac{3\beta}{4}(5\alpha^4+2\theta^2+8\alpha^2\theta)\right]\left[(\alpha^2+\theta)(2\alpha^2+\theta)-\frac{\beta}{4}(3\alpha^4+2\theta^2+6\alpha^2\theta)\right]$$
$$\left.+\frac{2}{(\alpha^2+\theta)^4}\left[(\alpha^2+\theta)(2\alpha^2+\theta)-\frac{\beta}{4}(3\alpha^4+2\theta^2+6\alpha^2\theta)\right]^3\right\},$$

$$Ku=\frac{\alpha^4}{\theta^4(\alpha^2+\theta)^2\sigma_X^4}\left\{\left[24(\alpha^2+\theta)(5\alpha^2+\theta)-\frac{30\beta}{4}(14\alpha^4+3\theta^2+18\alpha^2\theta)\right]-\frac{4}{(\alpha^2+\theta)^2}\left[6(\alpha^2+\theta)\right.\right.$$
$$\left.\times(4\alpha^2+\theta)-\frac{3\beta}{4}(25\alpha^4+7\theta^2+35\alpha^2\theta)\right]\left[(\alpha^2+\theta)(2\alpha^2+\theta)-\frac{\beta}{4}(3\alpha^4+2\theta^2+6\alpha^2\theta)\right]$$
$$+\frac{6}{(\alpha^2+\theta)^4}\left[(\alpha^2+\theta)(6\alpha^2+2\theta)-\frac{3\beta}{4}(5\alpha^4+2\theta^2+8\alpha^2\theta)\right]\left[(\alpha^2+\theta)(2\alpha^2+\theta)\right.$$
$$\left.\left.-\frac{\beta}{4}(3\alpha^4+2\theta^2+6\alpha^2\theta)\right]^2-\frac{3}{(\alpha^2+\theta)^6}\left[(\alpha^2+\theta)(2\alpha^2+\theta)-\frac{\beta}{4}(3\alpha^4+2\theta^2+6\alpha^2\theta)\right]^4\right\},$$

where $\sigma_X=\sqrt{V(X)}$ is the standard deviation of X. **Table 1** presents some values of these statistics for various parameters as in **Figure 1**.

Table 1 Statistics for the PTJ distribution with some various values of parameters α, θ and β.

Figure	α	θ	β	E(X)	V(X)	ID	Sk	Ku
a	0.05	0.05	0.75	0.66	0.53	0.81	2.69	15.03
b	0.15	0.15	0.50	0.85	0.86	1.01	2.39	11.80
c	1.50	0.50	0.20	5.01	16.14	3.22	1.56	6.66
d	3.00	0.50	0.70	8.54	43.08	5.05	1.80	8.50

Proposition 2 If $X\sim PTJ(\alpha,\theta,\beta)$ then the moment generating function (mgf) of X is;

$$M_X(t)=\left(\frac{\theta}{\alpha^2+\theta}\right)^2\left\{\frac{2\beta}{2\theta-\alpha(e^t-1)}\left[\frac{\alpha^2\theta}{2\theta-\alpha(e^t-1)}\left(1+\frac{2\alpha^2}{2\theta-\alpha(e^t-1)}\right)\right.\right.$$
$$\left.\left.+(\alpha^2+\theta)\left(1+\frac{\alpha^2}{2\theta-\alpha(e^t-1)}\right)\right]-\frac{(\beta-1)(\alpha^2+\theta)}{\theta-\alpha(e^t-1)}\left(1+\frac{\alpha^2}{\theta-\alpha(e^t-1)}\right)\right\}, \quad (16)$$

where $t < \log\left(\frac{\theta}{\alpha}+1\right)$ and the parameters $\alpha,\ \theta > 0$ and $-1 \le \beta \le 1$.

Proof The mgf of a mixed Poisson distribution can be obtained from;

$$M_X(t) = E\left[E\left(e^{tX} \mid \lambda\right)\right] = \int_0^\infty \left[\sum_{x=0}^\infty \frac{e^{tx}e^{-\lambda}\lambda^x}{\Gamma(x+1)}\right] g_{TJ}(\lambda)d\lambda,$$

since $\sum_{x=0}^\infty \frac{e^{tx}e^{-\lambda}\lambda^x}{\Gamma(x+1)} = e^{\lambda(e^t-1)}$ and $g_{TJ}(\lambda)$ as in (2), we obtain;

$$\begin{aligned}M_X(t) &= \frac{\theta^2}{\alpha(\alpha^2+\theta)^2}\left\{2\beta\int_0^\infty (\alpha\lambda+1)(\alpha^2+\theta+\alpha\theta\lambda)e^{-\left[\frac{2\theta}{\alpha}-(e^t-1)\right]\lambda}d\lambda\right.\\ &\quad \left.-(\beta-1)(\alpha^2+\theta)\int_0^\infty (\alpha\lambda+1)e^{-\left[\frac{\theta}{\alpha}-(e^t-1)\right]\lambda}d\lambda\right\},\\ &= \frac{\theta^2}{\alpha(\alpha^2+\theta)^2}\left\{2\beta\left[\frac{\alpha^2\Gamma(2)\alpha(\alpha^2+\theta)}{\left(2\theta-\alpha(e^t-1)\right)^2}+\frac{\alpha^3\Gamma(3)\alpha^2\theta}{\left(2\theta-\alpha(e^t-1)\right)^3}+\frac{\alpha\Gamma(1)(\alpha^2+\theta)}{2\theta-\alpha(e^t-1)}+\frac{\alpha^2\Gamma(2)\alpha\theta}{\left(2\theta-\alpha(e^t-1)\right)^2}\right]\right.\\ &\quad \left.-(\beta-1)(\alpha^2+\theta)\left[\frac{\alpha^2\Gamma(2)\alpha}{\left(\theta-\alpha(e^t-1)\right)^2}+\frac{\alpha\Gamma(1)}{\theta-\alpha(e^t-1)}\right]\right\}.\end{aligned}$$

Finally, we obtain $M_X(t)$ in (16).

Proposition 3 If $X \sim PTJ(\alpha,\theta,\beta)$ then the probability generating function (pgf) of X is;

$$\begin{aligned}H_X(s) &= \left(\frac{\theta}{\alpha^2+\theta}\right)^2\left\{\frac{2\beta}{2\theta+\alpha(1-s)}\left[\frac{\alpha^2\theta}{2\theta+\alpha(1-s)}\left(1+\frac{2\alpha^2}{2\theta+\alpha(1-s)}\right)\right.\right.\\ &\quad \left.\left.+(\alpha^2+\theta)\left(1+\frac{\alpha^2}{2\theta+\alpha(1-s)}\right)\right]-\frac{(\beta-1)(\alpha^2+\theta)}{\theta+\alpha(1-s)}\left(1+\frac{\alpha^2}{\theta+\alpha(1-s)}\right)\right\},\end{aligned} \qquad (17)$$

where the parameters $\alpha,\theta > 0$ and $-1 \le \beta \le 1$.

Proof The pgf of a mixed Poisson distribution can be obtained by utilizing the pgf of the Poisson distribution as follows;

$$H_X(s) = E\left[E\left(s^X \mid \lambda\right)\right] = \int_0^\infty \left[\sum_{x=0}^\infty \frac{s^x e^{-\lambda}\lambda^x}{\Gamma(x+1)}\right] g_{TJ}(\lambda)d\lambda,$$

since $\sum_{x=0}^\infty \frac{s^x e^{-\lambda}\lambda^x}{\Gamma(x+1)} = e^{-(1-s)\lambda}$ and $g_{TJ}(\lambda)$ as in (2), we have;

$$H_X(s)=\frac{\alpha^{-1}\theta^2}{(\alpha^2+\theta)^2}\left\{2\beta\int_0^{\infty}(\alpha\lambda+1)(\alpha^2+\theta+\alpha\theta\lambda)e^{-\left[\frac{2\theta}{\alpha}+(1-s)\right]\lambda}d\lambda-(\beta-1)(\alpha^2+\theta)\int_0^{\infty}(\alpha\lambda+1)e^{-\left[\frac{\theta}{\alpha}+(1-s)\right]\lambda}d\lambda\right\}$$

$$=\frac{\theta^2}{\alpha(\alpha^2+\theta)^2}\left\{2\beta\left\{\frac{\alpha^2\Gamma(2)\alpha(\alpha^2+\theta)}{[2\theta+\alpha(1-s)]^2}+\frac{\alpha^3\Gamma(3)\alpha^2\theta}{[2\theta+\alpha(1-s)]^3}+\frac{\alpha\Gamma(1)(\alpha^2+\theta)}{2\theta+\alpha(1-s)}+\frac{\alpha^2\Gamma(2)\alpha\theta}{[2\theta+\alpha(1-s)]^2}\right\}\right.$$

$$\left.-(\beta-1)\frac{\alpha\Gamma(1)(\alpha^2+\theta)}{\theta+\alpha(1-s)}\left[1+\frac{\alpha^2}{\theta+\alpha(1-s)}\right]\right\}.$$

Finally, we obtain $H_X(s)$ in (17). Alternatively, the pgf of the PTJ distribution can be obtained by setting $s=e^t$ as the mgf of X as in (16).

Generating a random variable of the PTJ distribution

Now we can generate a variate X from the PJT distribution with the parameters α, θ and β by the following algorithm.

Algorithm 1

Step 1 Generate λ with the inverse cdf of the TJ distribution in (5) as follows;

1) Generate U from the uniform on interval [0, 1].

2) Set $c=\frac{\alpha^2+\theta}{\alpha^2}$ and λ as follows:

$$\lambda=\frac{\alpha}{\theta}\left\{-c-W_{-1}\left\{-ce^{-c}\left[1-\frac{1}{2\beta}\left(1+\beta-\sqrt{(\beta+1)^2-4\beta u}\right)\right]\right\}\right\}.$$

Step 2 Generate X from the $\text{Pois}(\lambda)$.

Parameter estimation of the PTJ distribution

Parametric inference deals with the estimation of an unknown parameter of a chosen distribution. The experimenter assumes that $\tilde{x}=(x_1,\ldots,x_n)$ is a random sample $(X_1,\ldots,X_n)$ such that X_i are independent and identically distributed random variables when $X_i\sim\text{PTJ}(\alpha,\theta,\beta)$. In this study, the method of moments and the maximum likelihood estimation are introduced to estimate parameters α, θ and β.

Method of moments

The method of moment estimators (MMEs), i.e., $\hat{\alpha}$, $\hat{\theta}$ and $\hat{\beta}$ of parameters α, θ and β, respectively can be obtained by solving the following equations, i.e., $m_k=\mu'_k$;

$$m_1=\frac{1}{(\alpha^2+\theta)^2}\left(\frac{\alpha}{\theta}\right)\left[\left(\alpha^2+\theta\right)\left(2\alpha^2+\theta\right)-\frac{\beta}{4}\left(3\alpha^4+2\theta^2+6\alpha^2\theta\right)\right], \tag{18}$$

$$m_2=\frac{1}{(\alpha^2+\theta)^2}\left(\frac{\alpha}{\theta}\right)^2\left[\left(\alpha^2+\theta\right)\left(6\alpha^2+2\theta\right)-\frac{\beta}{4}\left(15\alpha^4+6\theta^2+24\alpha^2\theta\right)\right], \tag{19}$$

$$m_3=\frac{1}{(\alpha^2+\theta)^2}\left(\frac{\alpha}{\theta}\right)^3\left[\left(\alpha^2+\theta\right)\left(24\alpha^2+6\theta\right)-\frac{\beta}{4}\left(75\alpha^4+21\theta^2+105\alpha^2\theta\right)\right], \tag{20}$$

where m_k is the sample moments, i.e., $m_k=\frac{1}{n}\sum_{i=1}^{n}x_i^k, k=1,2,3.$

Since Eqs. (18) and (19) are no closed-form expression thus we use a numerical method by using the gmm package in the R stats package [18] to solve a non-linear model [19].

Maximum likelihood estimation

Let $X_1,\ldots,X_n$ be a random observation of size n from the PTJ distribution with the pmf in (4). The log-likelihood function for the parameter vector of $\tilde{\omega}=(\alpha,\theta,\beta)$ can be written as;

$$\log L(\tilde{\omega}) = 2n\log\theta - 2n\log(\alpha^2+\theta) + \sum_{i=1}^{n}\log\alpha^{x_i}\left\{\frac{2\beta}{(\alpha+2\theta)^{x_i+1}}\left[\frac{\alpha^2\theta(x_i+1)}{\alpha+2\theta}\left(1+\frac{\alpha^2(x_i+2)}{\alpha+2\theta}\right)\right.\right.$$
$$\left.\left.+(\alpha^2+\theta)\left(1+\frac{\alpha^2(x_i+1)}{\alpha+2\theta}\right)\right] - \frac{(\beta-1)(\alpha^2+\theta)}{(\alpha+\theta)^{x_i+1}}\left[1+\frac{\alpha^2(x_i+1)}{\alpha+\theta}\right]\right\}.$$

To estimate the unknown parameters α, θ and β, we take the partial derivatives with respect to α, θ and β and equate them to 0, i.e.,

$$\frac{\partial\log L(\tilde{\omega})}{\partial\alpha}=0, \frac{\partial\log L(\tilde{\omega})}{\partial\theta}=0, \frac{\partial\log L(\tilde{\omega})}{\partial\beta}=0. \tag{21}$$

Since the Eq. (21) cannot be derived in closed forms, thus we used the numerical method of the 3-dimensional Newton-Raphson type procedure. We obtain the solutions of the maximum likelihood estimators (MLEs) of α, θ and β from the Eq. (21) by using the nlm function in the R stats package [18].

Simulation

We run a simulation study for parameter estimation of the PTJ parameters. The simulations are described as:

1) The sample sizes are taken as n = 20, 40, 60, 80, 100, 150, 300 and 500.

2) The data are generated from the PTJ(α, θ, β) where the true parameter values are set as 4-cases as in **Table 1**.

3) Each situation is repeated $T = 1000$ times.

4) Formulas used for calculating the bias and root mean square error (RMSE) values of $\hat{\alpha}$, $\hat{\theta}$ and $\hat{\beta}$ are given by ;

$$\text{Bias}(\hat{\alpha})=\frac{1}{T}\left(\hat{\alpha}_t-\alpha\right),\ \text{Bias}(\hat{\theta})=\frac{1}{T}\left(\hat{\theta}_t-\theta\right),\ \text{Bias}(\hat{\beta})=\frac{1}{T}\left(\hat{\beta}_t-\beta\right),$$
$$\text{RMSE}(\hat{\alpha})=\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\alpha}_t-\alpha\right)^2},\ \text{RMSE}(\hat{\theta})=\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\theta}_t-\theta\right)^2},\ \text{and}\ \text{RMSE}(\hat{\beta})=\sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\beta}_t-\beta\right)^2}.$$

The results from simulated data sets are reported in **Tables 2 - 5**. The average bias and RMSE values of the MMEs and MLEs are decreased when the sample size increases for all cases. **Tables 2 - 5** indicate that the MLEs have the estimated value close to the true parameters more than the MMEs in all cases. However, the results in **Tables 2** and **3**, data sets with under-dispersion and equi-dispersion, indicate that the MMEs give a poor estimated value, but the MLEs still give the estimated value close to the true parameters.

Table 2 The bias and RMSE values of the MMEs and MLEs of the PTJ distribution with parameters $\alpha = 0.05, \theta = 0.05$ and $\beta = 0.75$ ($\mathrm{ID} = 0.81$: Under-dispersion).

		MMEs			MLEs		
n		α	θ	β	α	θ	β
20	Bias	0.9373	2.3346	−0.5100	0.0533	0.0449	−0.2052
	RMSE	2.3341	6.2416	0.6289	0.1772	0.1593	0.3783
40	Bias	0.8306	2.1763	−0.5058	0.0506	0.0457	−0.1312
	RMSE	1.9283	5.5675	0.6523	0.1148	0.1064	0.2450
60	Bias	0.7692	2.0095	−0.5284	0.0311	0.0295	−0.1252
	RMSE	1.7479	4.9456	0.6768	0.0927	0.0862	0.2216
80	Bias	0.6787	1.7916	−0.4955	0.0215	0.0182	−0.1001
	RMSE	1.5221	4.5015	0.6495	0.0731	0.0688	0.1792
100	Bias	0.6331	1.6497	−0.4693	0.0097	0.0066	−0.0851
	RMSE	1.4628	4.2898	0.6318	0.0553	0.0547	0.1520
150	Bias	0.4726	1.1386	−0.4788	0.0098	0.0061	−0.0738
	RMSE	1.0630	2.8360	0.6439	0.0441	0.0425	0.1259
300	Bias	0.3315	0.6865	−0.4269	0.0044	−0.0052	−0.0335
	RMSE	0.6594	1.4879	0.5845	0.0262	0.0235	0.0682
500	Bias	0.1995	0.5269	−0.4171	0.0040	−0.0005	−0.0080
	RMSE	0.4052	0.9885	0.5529	0.0235	0.0191	0.0370

Table 3 The bias and RMSE values of the MMEs and MLEs of the PTJ distribution with parameters $\alpha = 0.15, \theta = 0.15$ and $\beta = 0.5$ ($\mathrm{ID} = 1.01$: equi-dispersion).

		MMEs			MLEs		
n		α	θ	β	α	θ	β
20	Bias	1.0827	2.4119	−0.3092	0.1021	0.1108	−0.1012
	RMSE	2.9373	6.4841	0.4890	0.7697	0.7872	0.3399
40	Bias	1.0539	2.3359	−0.2828	0.0548	0.0481	−0.0101
	RMSE	2.8191	6.1521	0.5045	0.4617	0.4658	0.2607
60	Bias	0.7614	1.6701	−0.2905	0.0507	0.0427	0.0114
	RMSE	1.9507	4.4498	0.5031	0.3992	0.4309	0.2248
80	Bias	0.5905	1.2806	−0.2398	0.0421	0.0350	0.0599
	RMSE	1.7703	3.8296	0.4734	0.3713	0.4067	0.2133
100	Bias	0.6062	1.2481	−0.1811	0.0371	0.0261	0.0728
	RMSE	1.7036	3.7201	0.4306	0.3510	0.3276	0.2112
150	Bias	0.4756	0.8735	−0.1464	0.0289	0.0099	0.0753
	RMSE	1.5260	2.6292	0.3851	0.3190	0.2868	0.2127
300	Bias	0.2702	0.4416	−0.0984	0.0078	−0.0043	0.0421
	RMSE	0.8611	1.2048	0.3404	0.2737	0.2512	0.1163
500	Bias	0.1392	0.2259	−0.0346	−0.0027	−0.0062	0.0274
	RMSE	0.4721	0.6915	0.2839	0.2301	0.2005	0.1180

Table 4 The bias and RMSE values of the MMEs and MLEs of the PTJ distribution with parameters $\alpha = 1.5, \theta = 0.5$ and $\beta = 0.2$ ($\text{ID} = 3.22$: over-dispersion).

		MMEs			MLEs		
n		α	θ	β	α	θ	β
20	Bias	0.7590	0.3195	−0.1082	0.1094	0.0532	0.0563
	RMSE	5.4854	2.0761	0.4476	0.4951	0.2165	0.4536
40	Bias	0.4788	0.2162	−0.1659	0.0662	0.0332	−0.0007
	RMSE	4.5362	1.6587	0.4815	0.4853	0.2108	0.4736
60	Bias	0.5930	0.2336	−0.1270	0.0105	−0.0081	0.0375
	RMSE	4.2059	1.5319	0.4692	0.4389	0.2031	0.4260
80	Bias	0.5433	0.3195	−0.1082	0.1094	0.0532	0.0563
	RMSE	4.0495	2.0761	0.4476	0.4951	0.2165	0.4536
100	Bias	0.4241	0.1364	−0.1059	0.0367	0.0248	0.1248
	RMSE	3.9047	1.2648	0.4660	0.4258	0.1693	0.3915
150	Bias	0.4946	0.1608	−0.0813	0.0130	0.0209	0.1692
	RMSE	3.4496	1.1713	0.4578	0.4229	0.1686	0.3600
300	Bias	0.2993	0.0965	−0.0589	0.0520	0.0049	0.0777
	RMSE	2.5698	0.8905	0.4220	0.3629	0.0773	0.1770
500	Bias	0.1705	0.0767	−0.0469	0.0289	−0.0295	0.0279
	RMSE	2.1090	0.8026	0.2849	0.1821	0.0791	0.1599

Table 5 The bias and RMSE values of the MMEs and MLEs of the PTJ distribution with parameters $\alpha = 3, \theta = 0.5$ and $\beta = 0.7$ ($\text{ID} = 5.05$: over-dispersion).

		MMEs			MLEs		
n		α	θ	β	α	θ	β
20	Bias	−0.3212	−0.0049	−0.3249	0.0872	−0.0069	−0.0332
	RMSE	5.2880	1.0119	0.5170	0.8910	0.3195	0.1749
40	Bias	−0.7363	0.0756	−0.3663	0.0261	−0.0235	−0.0165
	RMSE	4.8868	0.9138	0.5416	0.8176	0.2513	0.1683
60	Bias	−0.4147	−0.0203	−0.3148	0.0846	−0.0013	−0.0334
	RMSE	4.6500	0.8671	0.5064	0.8007	0.2453	0.1609
80	Bias	−0.2510	−0.0001	0.2792	0.0461	−0.0119	−0.0323
	RMSE	4.5957	0.8636	0.4802	0.7943	0.2399	0.1416
100	Bias	−0.1103	0.0264	−0.2237	−0.0006	−0.0161	−0.0275
	RMSE	4.5562	0.8592	0.4419	0.7371	0.2157	0.1237
150	Bias	0.0557	0.0428	−0.2057	0.2429	0.0159	−0.0343
	RMSE	4.1688	0.7523	0.4302	0.6970	0.2014	0.1138
300	Bias	0.3991	0.0996	−0.1158	0.0260	−0.0113	−0.0377
	RMSE	3.7946	0.7114	0.3465	0.6038	0.1092	0.0768
500	Bias	0.5631	0.1393	−0.0956	0.0175	−0.0089	−0.0281
	RMSE	3.5453	0.7133	0.3243	0.4826	0.0819	0.0573

Application

For an application study, some real data sets are fitting distributions such as the PTJ, PJ, PL and Pois distributions. The parameters of each distribution are estimated by the method of moments and the maximum likelihood estimation. The first data set (Data I) is the number of European red mites on apple leaves, which appear in [20]. The second data set (Data II) is the number of insects of Pyrausta nubilalis in 1937 [21]. The third data (Data III) is the number of hospitalizations, per family member, per year [22]. And the number of the outbreak of strikes (Data IV) in the south of London during World War II [23]. The summary of these data is shown in **Table 6**.

Table 6 Data summary of the real data sets.

Data	n	Median	Mode	Mean	Variance	ID
The number of European red mites per leaf	150	1	0	1.1470	2.2736	1.9822
The number of insects	56	0	0	0.7500	1.3382	1.7843
The number of hospitalizations per family member	2,924	0	0	0.0985	0.1059	1.0751
The number of outbreak of strikes	156	1	1	0.9936	0.7419	0.7419

In this study, the Kolmogorov-Smirnov (KS) test is used for the criteria for the goodness of fit test of each distribution, where the model has the smaller value of KS test that is the best fit for the data. The parameter estimates and the goodness of fit test for these data sets are illustrated in **Tables 7 - 10**. The results show that the maximum likelihood estimation gives the parameter estimates that make the fitting distributions close to the empirical data greater than the fitting distribution by using the parameter estimate of the method of moments. The fitted distribution with the parameters estimated by the maximum likelihood estimation gives the lower KS value than the fitted distribution with the parameters estimated by the method of moments in all cases (**Tables 7 - 10**). In **Tables 7** and **8**, the real data sets with over-dispersion, the PTJ distribution gives a lower value of the KS test compared to other distributions such as the PJ, PL and Poisson distributions. For count data with equi-dispersion and under-dispersion, the Poisson distribution gives a lower KS value than other distributions (**Tables 9** and **10**). However, the fitted PTJ distribution with the parameters estimated by the maximum likelihood estimation gives the lower KS value than other distributions in count data with equi-dispersion (**Table 9**). We conclude that the PTJ distribution is appropriate to fit count data with over-dispersion.

Table 7 Fitting distributions of the numbers of European red motes on apple leaves (X).

X	observed	Expected values of fitting distributions							
		Maximum likelihood estimation				Method of moments			
		Pois	**PL**	**PJ**	**PTJ**	**Pois**	**PL**	**PJ**	**PTJ**
0	70	47.65	67.26	68.91	69.01	47.66	57.89	64.5	66.26
1	38	54.64	38.89	37.82	37.76	54.64	37.81	36.76	36.40
2	17	31.33	21.24	20.42	20.37	31.33	23.18	20.96	20.32
3	10	11.98	11.19	10.89	10.87	11.97	13.65	11.95	11.49
4	9	3.43	5.74	5.75	5.74	3.43	7.82	6.81	6.56
5	3	0.79	2.89	3.01	3.01	0.79	4.39	3.88	3.77
6	2	0.15	1.43	1.56	1.57	0.15	2.42	2.21	2.18
7	1	0.02	0.70	0.81	0.81	0.02	1.32	1.26	1.26
$-\log L$		242.81	222.51	222.38	222.38	-	-	-	-
AIC		487.62	447.02	448.76	450.76	-	-	-	-
BIC		490.63	450.03	454.78	459.79	-	-	-	-
$\hat{\lambda}$		1.1467	-	-	-	1.1466	-	-	-
$\hat{\alpha}$		-	-	0.3552	0.4297	-	-	0.0033	0.0653
$\hat{\theta}$		-	1.2602	0.3861	0.4674	-	1.0355	0.0025	0.0494
$\hat{\beta}$		-	-	-	0.0580	-	-	-	0.1781
KS test		0.1490	0.0239	0.0202	0.0200	0.1490	0.0820	0.0449	0.0356
p-value		< 0.0001	0.8560	0.9154	0.9165	< 0.0001	0.5572	0.9265	0.9540

Table 8 Fitting distributions of the numbers of insects (X).

X	Observed	Expected values of fitting distributions							
		Maximum likelihood estimation				Method of moments			
		Pois	**PL**	**PJ**	**PTJ**	**Pois**	**PL**	**PJ**	**PTJ**
0	33	26.45	31.52	31.98	32.19	25.16	21.61	28.21	30.22
1	12	19.84	14.15	13.72	13.56	20.13	14.12	14.22	13.53
2	6	7.44	6.08	5.88	5.80	8.05	8.65	7.03	6.25
3	3	1.86	2.53	2.52	2.51	2.15	5.10	3.42	2.98
4	1	0.35	1.03	1.08	1.09	0.43	2.92	1.65	1.46
5	1	0.05	0.41	0.46	0.48	0.07	1.64	0.78	0.74
$-\log L$		71.58	66.98	66.92	66.91	-	-	-	-
AIC		145.16	135.96	137.84	139.82	-	-	-	-
BIC		147.19	137.99	141.89	145.90	-	-	-	-
$\hat{\lambda}$		0.7500	-	-	-	0.8000	-	-	-
$\hat{\alpha}$		-		0.0714	0.2886	-	-	0.4437	0.0831
$\hat{\theta}$		-	1.8115	0.0998	0.4087	-	1.0355	0.5788	0.0711
$\hat{\beta}$		-	-	-	0.1864	-	-	-	0.5887
KS test		0.1169	0.0265	0.0182	0.0145	0.1400	0.2034	0.0856	0.0496
p-value		0.4282	1.0000	1.0000	1.0000	0.2227	0.0195	0.8068	0.9991

Table 9 Fitting distributions of the number of hospitalizations per family member (X).

X	Observed	Expected value of fitting distributions							
		Maximum likelihood estimation				Method of moments			
		Pois	PL	PJ	PTJ	Pois	PL	PJ	PTJ
0	2,659	2,649.72	2,546.5	2,661.75	2,659.05	2,649.72	2,561.15	2,562.07	2,570.33
1	244	261.00	329.19	238.74	243.65	261.00	318.2	317.13	308.84
2	19	12.85	42.17	21.41	19.67	12.85	39.2	39.25	38.98
3+	2	0.43	6.14	2.10	1.63	0.43	5.45	5.55	5.85
	$-\log L$	972.26	993.62	969.25	969.07	-	-	-	-
	AIC	1,946.52	1,989.24	1,942.50	1,944.14	-	-	-	-
	BIC	1,952.50	1,995.22	1,954.46	1,962.08	-	-	-	-
	$\hat{\lambda}$	0.0985	-	-	-	0.0985	-	-	-
	$\hat{\alpha}$	-	-	0.1430	0.0785	-	-	0.0014	0.1333
	$\hat{\theta}$	-	7.5459	1.4713	1.0022	-	7.8649	0.0101	0.8667
	$\hat{\beta}$	-	-	-	−0.4998	-	-	-	0.2333
	KS test	0.0032	0.0385	0.0009	0.0001	0.0032	0.0335	0.0311	0.0303
	p-value	1.0000	0.0003	1.0000	1.0000	1.0000	0.0029	0.0069	0.0092

Table 10 Fitting distributions of the number of outbreak of strikes (X).

X	Observed	Expected value of fitting distributions							
		MLEs				MEs			
		Pois	PL	PJ	PTJ	Pois	PL	PJ	PTJ
0	46	57.76	75.24	56.57	70.00	57.76	77.80	79.14	69.72
1	76	57.39	40.55	45.00	42.74	57.39	40.48	39.60	44.18
2	24	28.51	20.73	26.86	22.60	28.51	20.01	19.43	23.04
3	9	9.44	10.23	14.25	11.12	9.44	9.56	9.39	10.84
4	4	2.35	4.93	7.09	5.24	2.35	4.46	4.48	4.79
5+	0	0.55	4.32	6.23	4.30	0.55	3.69	3.96	3.43
	$-\log L$	191.94	212.36	208.00	208.79	-	-	-	-
	AIC	385.88	426.72	420.00	423.58	-	-	-	-
	BIC	388.93	429.77	426.10	432.73	-	-	-	-
	$\hat{\lambda}$	0.9936	-	-	-	0.9936	-	-	-
	$\hat{\alpha}$	-	-	4,878.70	3.7682	-	-	0.4512	0.9111
	$\hat{\theta}$	-	1.4010	7,382.48	5.8566	-	1.4751	0.5972	1.6556
	$\hat{\beta}$	-	-	-	0.0586	-	-	-	−0.6296
	KS test	0.0754	0.1874	0.1310	0.1538	0.0754	0.2038	0.2124	0.1782
	p-value	0.3382	< 0.0001	0.0095	0.0012	0.3382	< 0.0001	< 0.0001	0.0001

Conclusions

A new 3-parameter discrete distribution for modeling count data, so-called the Poisson transmuted Janardan (PTJ) distribution, is proposed, which is obtained by mixing the Poisson distribution and the transmuted Janardan distribution. The PTJ has 2 special models like the Poisson Janardan distribution by [8] and Poisson Lindley distribution by [4]. Some mathematical characteristics of the proposed distribution, including the moments, moment generating function and probability generating function, are introduced. We discuss the parameter estimation by using the moments of method and the maximum likelihood estimation and testing of hypotheses for the model parameters with the KS test. The simulation results indicate that the maximum likelihood estimation gives the estimated parameters close to the true parameters more than the method of moments. The results of the application study found that the PTJ distribution is appropriate to fit count data with over-dispersion. The practical relevance and applicability of the new model are demonstrated in 4 real data sets. We verify that some sub-models could provide similar fits while the PTJ distribution being more parsimonious models for count data with over-dispersion.

Acknowledgements

The authors acknowledge profound thanks to anonymous referee for giving comments which have immensely improved the presentation of the paper.

References

[1] D Bhati, P Kumawat and E Gómez-Déniz. A new count model generated from mixed Poisson transmuted exponential family with an application to health care data. *Comm. Stat. Theor. Meth*. 2017; **46**, 11060-76.
[2] D Karlis and E Xekalaki. Mixed Poisson distributions. *Int. Stat. Rev.* 2005; **73**, 35-58.
[3] M Greenwood and GU Yule. 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. *J. Roy. Stat. Soc.* 1920; **83**, 255-79.
[4] M Sankaran. The discrete Poisson-lindley distribution. *Biometrics* 1970; **26**, 145-9.
[5] HS Sichel. On a distribution law for word frequencies. *J. Am. Stat. Assoc.* 1975; **70**, 542-7.
[6] GE Willmot. On recursive evaluation of mixed Poisson probabilities and related quantities. *Scand. Actuarial J.* 1993; **18**, 114-33.
[7] SA Al-Awadhi and ME Ghitany. Statistical properties of Poisson Lomax distribution and its application to repeated accident data. *J. Appl. Stat. Sci.* 2001; **10**, 365-72.
[8] E Mahmoudi and H Zakerzadeh. Generalized Poisson-Lindley distribution. *Comm. Stat. Theor. Meth.* 2010; **39**, 1785-98.
[9] H Zamani, N Ismail and P Faroughi. Poisson-weighted exponential univariate version and regression model with applications. *J. Math. Stat.* 2014; **10**, 148-54.
[10] R Shanker, S Sharma, U Shanker, R Shanker and TA Leonida. The discrete Poisson-Janardan distribution with applications. *Int. J. Soft Comput. Eng.* 2014; **4**, 31-3.
[11] E Gómez-Déniz and E Calderín-Ojeda. Parameters estimation for a new generalized geometric distribution. *Comm. Stat. Simulat. Comput.* 2015; **44**, 2023-39.
[12] W Wongrin and W Bodhisuwan. The Poisson-generalised Lindley distribution and its applications. *Songklanakarin J. Sci. Tech.* 2016; **38**, 645-56.
[13] AI Al-Omari, AM Al-khazaleh and LM Alzoubi. Transmuted Janardan distribution: A generalization of the Janardan distribution. *J. Stat. Appl. Probab.* 2017; **5**, 1-11.
[14] R Shanker, S Sharma, U Shanker and R Shanker. Janardan distribution and its application to waiting times data. *Indian J. Appl. Res.* 2013; **3**, 500-2.
[15] D Veberič. Lambert W function for applications in physics. *Comput. Phys. Comm.* 2012; **183**, 2622-8.
[16] RM Corless, GH Gonnet, DE Hare, DJ Jeffrey and DE Knuth. On the Lambert W function. *Adv. Comput. Math.* 1996; **5**, 329-59.
[17] WT Shaw and IR Buckley. The alchemy of probability distributions: Beyond Gram Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map, Available at: https://arxiv.org/abs/0901.043, accessed January 2020.
[18] R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018.

[19] P Chaussé. Computing generalized method of moments and generalized empirical likelihood with R. *J. Stat. Software* 2010; **34**, 1-35.
[20] CI Bliss and RA Fisher. Fitting the negative binomial distribution to biological data. *Biometrics* 1953; **9**, 176-200.
[21] G Beall. The fit and significance of contagious distributions when applied to observations on larval insects. *Ecology* 1940; **21**, 460-74.
[22] SA Klugman, HH Panjer and GE Willmot. *Loss models: From data to decisions*. John Wiley & Sons, New York, 1998.
[23] MH Alamatsaz, S Dey, T Dey and SS Harandi. Discrete generalized Rayleigh distribution. *Pak. J. Stat.* 2016; **32**, 1-20.