Statistics and Applications {ISSN 2454-7395 (online)}
Volume 24, No. 1, 2026 (New Series), pp 187–220
http://www.ssca.org.in/journal

Poisson–Transmuted Geometric Convolution for Overdispersed Count Data

Anupama Nandi[1], Partha Jyoti Hazarika[1], Aniket Biswas[2], Morad Alizadeh[3], Hadi Saboori[4] and Mohamed S. Eliwa[5,6]

[1]*Department of Statistics, Dibrugarh University, India*
[2]*Department of Statistics, St. Xavier's University, Kolkata, India*
[3]*Department of Statistics, Persian Gulf university, Bushehr, 75169, Iran*
[4]*Department of Statistics, Faculty of Sciences, University of Zabol, Zabol, Iran*
[5]*Department of Statistics and Operations Research, College of Science, Qassim University, Saudi Arabia*
[6]*Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt*

Received: 10 May 2025; Revised: 01 September 2025; Accepted: 05 September 2025

Abstract

A novel overdispersed count distribution is obtained by convolving independently distributed Poisson and transmuted geometric random variables. This distribution extends the Poisson-geometric, geometric, and transmuted geometric distributions. Essential statistical properties are analyzed. Maximum likelihood estimators of the unknown constants are derived using numerical optimization techniques and the EM algorithm. Extensive simulation studies evaluate the estimators performance under different conditions. Additionally, a flexible regression model built on the proposed distribution is formulated. Real-world applications in modeling overdispersed count data, both with and without covariates, demonstrate the model's relevance. The proposed model demonstrates desirable statistical properties and surpasses its closest competitors in empirical applications.

Key words: Transmuted geometric distribution; EM algorithm; Computer simulation; Prediction model; Failure analysis; Statistics and numerical data.

MSC 2010: 60E05, 62E15

1. Introduction

Overdispersion is a common phenomenon in modeling applications, occurring more frequently than underdispersion or equidispersion. The literature presents numerous count models specifically designed to handle overdispersed data. Given the sustained research interest in this area, the development of a simple yet effective model remains essential. Various distributions and regression models for overdispersed count data have been extensively studied [Moghimbeigi, *et al.* (2008), Rodrigues-Motta, *et al.* (2013), Hassanzadeh

Corresponding Author: Anupama Nandi
Email: anupamanandi0702@gmail.com

and Kazemi (2016), Wongrin and Bodhisuwan (2017), Wang, *et al.* (2017), Sarvi, *et al.* (2019), Moqaddasi Amiri, *et al.* (2019), Tapak, *et al.* (2020), Ghahramani and White (2020), Tüzen, *et al.* (2020), Altun (2020), and Bar-Lev and Ridder (2021)]. Recent developments in overdispersed count regression models include the finite mixtures of mean-parameterized Conway–Maxwell–Poisson (COM-Poisson) regression introduced by Zhan and Young (2024) and the marginalized zero-inflated Bell regression model proposed by Amani, *et al.* (2025).

Although widely used, the Poisson distribution $Poi(\theta)$ is constrained by its equidispersion property. To overcome this limitation, several alternative distributions have been introduced, including the hyper-Poisson by Bardwell and Crow (1964), the double-Poisson by Efron (1986), the weighted Poisson by Del Castillo and Pérez-Casany (1998), the weighted generalized Poisson by Chakraborty (2010), the widely adopted COM-Poisson by Sellers and Shmueli (2010), the Mittag-Leffler function distribution by Chakraborty and Ong (2017), and the Tilted Beta-Binomial Distribution by Hahn (2022).

In addition to these distributions, various modifications of the geometric $Geo(\theta)$ distribution have been studied for modeling overdispersed count data. Notable contributions include the works of Chakraborty and Gupta (2015), Jain and Consul (1971), Philippou, *et al.* (1983), Tripathi, *et al.* (1987), Makcutek (2008), Gómez-Déniz (2010), and Nekoukhou, *et al.* (2012), among others. A key improvement of the geometric distribution is the transmuted geometric distribution (Chakraborty and Bhati, 2016) $TGD(q, \alpha)$ which is developed using the quadratic rank transmutation technique. For a random variable Y following the $TGD(q, \alpha)$ distribution, the probability mass function (pmf) is

$$P(Y = y) = \alpha q^{2y}(1 - q^2) + (1 - \alpha)q^y(1 - q), \;\; y \geq 0. \tag{1}$$

Here, $q \in (0, 1)$ and $\alpha \in (-1, 1)$. It exhibits underdispersion for $\alpha \in (-1, 0)$ and overdispersion for blue $\alpha \in (0, 1)$.

Despite the existence of numerous discrete distributions, there remains a need for new models that can effectively capture overdispersed count data while retaining simplicity and interpretability. The convolution method serves as a powerful yet straightforward technique for constructing new probability distributions. By convolving two known distributions, a new distribution with distinct characteristics can be derived, which is applicable to diverse data-modeling scenarios. Bourguignon and Weiß (2017) introduced a count distribution, termed BerG, by convolving a Bernoulli and a geometric variable. The Bernoulli distribution is underdispersed, with a variance lower than its mean, whereas the geometric distribution is overdispersed, with a variance greater than its mean. The BerG model effectively accommodates underdispersed, overdispersed, and equidispersed data. The Bernoulli-Poisson (BerPoi) distribution of Bourguignon, *et al.* (2022), and the Poisson-geometric (PoiG) distribution of Nandi, *et al.* (2024b) are some recent advancements in flexible count data models based on a similar approach. Later, Nandi, *et al.* (2024a) developed an overdispersed regression model, termed the PoiG regression model, based on the mean parameterization of the PoiG distribution.

The PoiG distribution is derived by convolving the Poisson and geometric distributions. Let $X_1 \sim Poi(\theta)$, where $\theta > 0$ is the parameter of the Poisson distribution and $X_2 \sim G(q)$, where $0 < q < 1$ is the parameter of the geometric distribution. If X_1 and X_2

are independent, then their sum, $X = X_1 + X_2$, follows the pmf

$$P(X = x) = \frac{q^x(1-q)}{\Gamma(x+1)} \exp\left(\frac{\theta(1-q)}{q}\right)\Gamma\left(x+1, \frac{\theta}{q}\right), \qquad x \geq 0. \tag{2}$$

The PoiG regression model of Nandi, *et al.* (2024a) is an overdispersed regression model developed by considering the mean parameterization of the PoiG distribution.

This study introduces the Poisson transmuted geometric (PoiTG) distribution using the above mentioned technique. Compared to the COM-Poisson distribution, the PoiTG distribution is more interpretable as it does not require a complex normalizing constant in its pmf. Additionally, its compact expressions for the mean and variance make it suitable for regression modeling. This distribution extends the Poisson, transmuted geometric, and PoiG distributions, offering a flexible approach to count data modeling.

The paper is structured as follows: Section 2 introduces the PoiTG distribution. Section 3 delves into its key statistical properties, such as recurrence relations, generating functions, moments, skewness, kurtosis, dispersion index, reliability function, hazard rate function, mean residual life function and Shannon entropy. Section 4 covers parameter estimation methods. Section 5 evaluates the performance of maximum likelihood estimators through simulation studies. Section 6 demonstrates the practical applicability of the PoiTG distribution by analyzing real-world data sets. Section 7 proposes a regression model based on the PoiTG distribution and presents an empirical application. Finally, the concluding remarks and potential future research directions are provided in the last section.

2. The PoiTG distribution

Let Y_1 follows the Poisson distribution with parameter $\theta > 0$, that is, $Poi(\theta)$ and Y_2 follows the transmuted geometric distribution with parameters $q \in (0,1)$ and $\alpha \in (0,1)$ (Chakraborty and Bhati, 2016) independent of Y_1, that is, $TGD(q,\alpha)$. Chakraborty and Bhati (2016) originally considered $-1 < \alpha < 1$. However, we restrict α to $0 < \alpha < 1$ to better illustrate the overdispersion in the proposed convolution-based model. This restriction also proves useful in Section 4.2 for implementing the EM algorithm. Since both Y_1 and Y_2 can take values from the set, denoted as N_0, the convolution $Y = Y_1 + Y_2$ also has support on N_0 and

$$\Pr(Y = y) = (1-\alpha)(1-q)q^y e^{-\theta} \sum_{i=0}^{y} \frac{1}{i!}\left(\frac{\theta}{q}\right)^i + \alpha(1-q^2)q^{2y}e^{-\theta}\sum_{i=0}^{y}\frac{1}{i!}\left(\frac{\theta}{q^2}\right)^i. \tag{3}$$

The distribution in (3) is referred to as the PoiTG distribution and we denote it as $Y \sim PoiTG(\theta, q, \alpha)$. Note that the pmf of Y can also be written as

$$p_Y(y) = (1-\alpha)w_1(y,\theta,q) + \alpha w_2(y,\theta,q), \qquad y \geq 0\ , \tag{4}$$

where,

$$\begin{aligned} w_k(y,\theta,q) &= (1-q^k)q^{ky}e^{-\theta}\sum_{i=0}^{y}\frac{1}{i!}\left(\frac{\theta}{q^k}\right)^i \\ &= \frac{(1-q^k)q^{ky}}{\Gamma(y+1)}\exp\left(\frac{\theta}{q^k}-\theta\right)\Gamma\left(y+1,\frac{\theta}{q^k}\right), \qquad k = 1,2\ . \end{aligned}$$

The incomplete gamma function (Abramowitz and Stegun, 1964) in (4) is defined as $\Gamma(k,x) = \int_x^\infty t^{(k-1)}e^{-t}dt$ and it can also be defined as

$$\Gamma(k,x) = (k-1)!\sum_{n=0}^{k-1}\frac{e^{-x}x^n}{n!},$$

which holds for non-negative values of k and any real value of x. The incomplete gamma function in $w_k(y,\theta,q)$ can be rewritten as

$$\Gamma\left(y+1,\frac{\theta}{q^k}\right) = y!\sum_{i=0}^{y}\frac{1}{i!}\exp\left(-\frac{\theta}{q^k}\right)\left(\frac{\theta}{q^k}\right)^i,$$

The nature of the pmf in (3) of $Y \sim PoiTG(\theta,q,\alpha)$ is demonstrated by Figure 1 for different combinations of θ, q, and α. These plots clearly depict that the distribution of $PoiTG(\theta,q,\alpha)$ is unimodal and has, at most, one exponential tail. The cumulative distribution function (cdf) of the proposed distribution is

$$F_Y(y) = \frac{\Gamma(y+1,\theta)}{\Gamma(y+1)} - (1-\alpha)W_1(y,\theta,q) - \alpha W_2(y,\theta,q), \qquad y \geq 0, \tag{5}$$

where,

$$W_k(y,\theta,q) = \frac{q^{k(y+1)}}{\Gamma(y+1)}\exp\left(\frac{\theta(1-q^k)}{q^k}\right)\Gamma\left(y+1,\frac{\theta}{q^k}\right), \qquad k = 1,2\ .$$

The nature of the cdf in (5) of $PoiTG(\theta,q,\alpha)$ is demonstrated by Figure 2 for different combinations of θ, q, and α.

Remark 1: As $\theta \to 0$, the $PoiTG(\theta,q,\alpha)$ distribution exhibits behavior similar to the $TGD(q,\alpha)$. Similarly, as $\alpha \to 0$, it behaves akin to the $PoiG(\theta,1-q)$ distribution. Furthermore, when both θ and $\alpha \to 0$, it resembles the $G(1-q)$ distribution.

3. Statistical properties

3.1. Recurrence relation

The probability recurrence relation facilitates the determination of a term based on its preceding term(s). This approach is particularly useful for computing probability masses at various support points. Note that,

$$p_Y(y) = \alpha(1-q^2)q^{2y}e^{-\theta}s''_y + (1-\alpha)(1-q)q^y e^{-\theta}s'_y$$

where,

$$s'_y = \sum_{i=0}^{y}\frac{1}{i!}\left(\frac{\theta}{q}\right)^i,$$

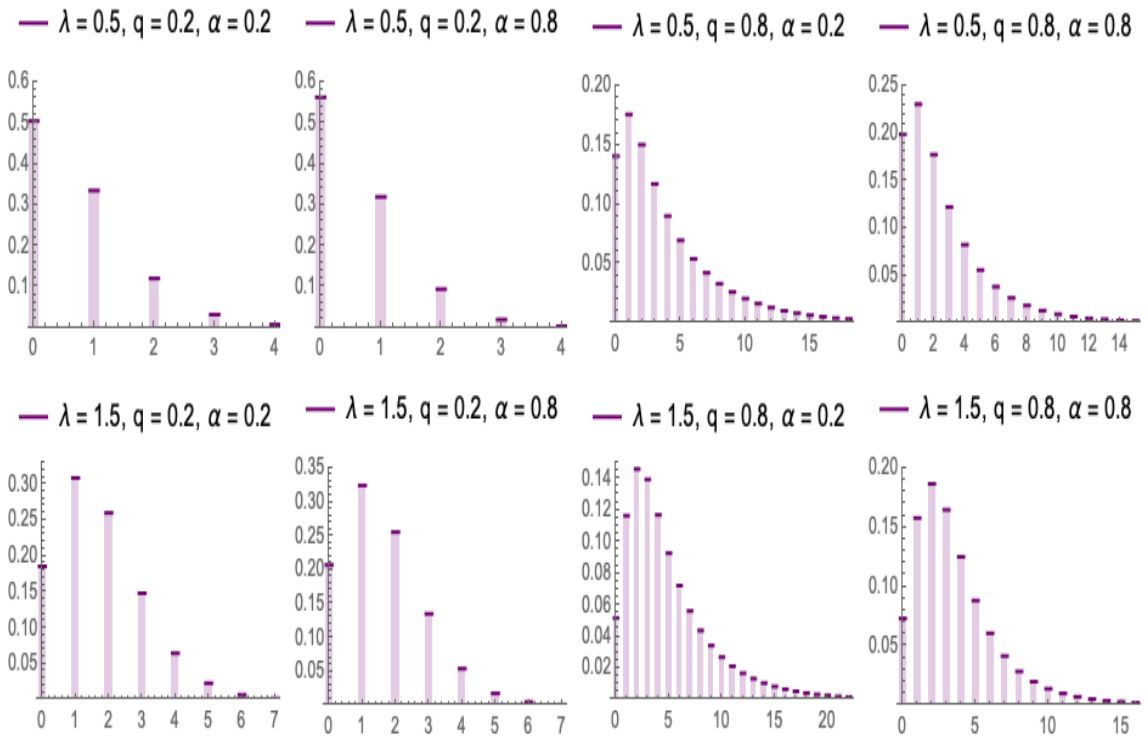

Figure 1: The pmf of $PoiTG(\theta, q, \alpha)$ for multiple choices of θ, q and α

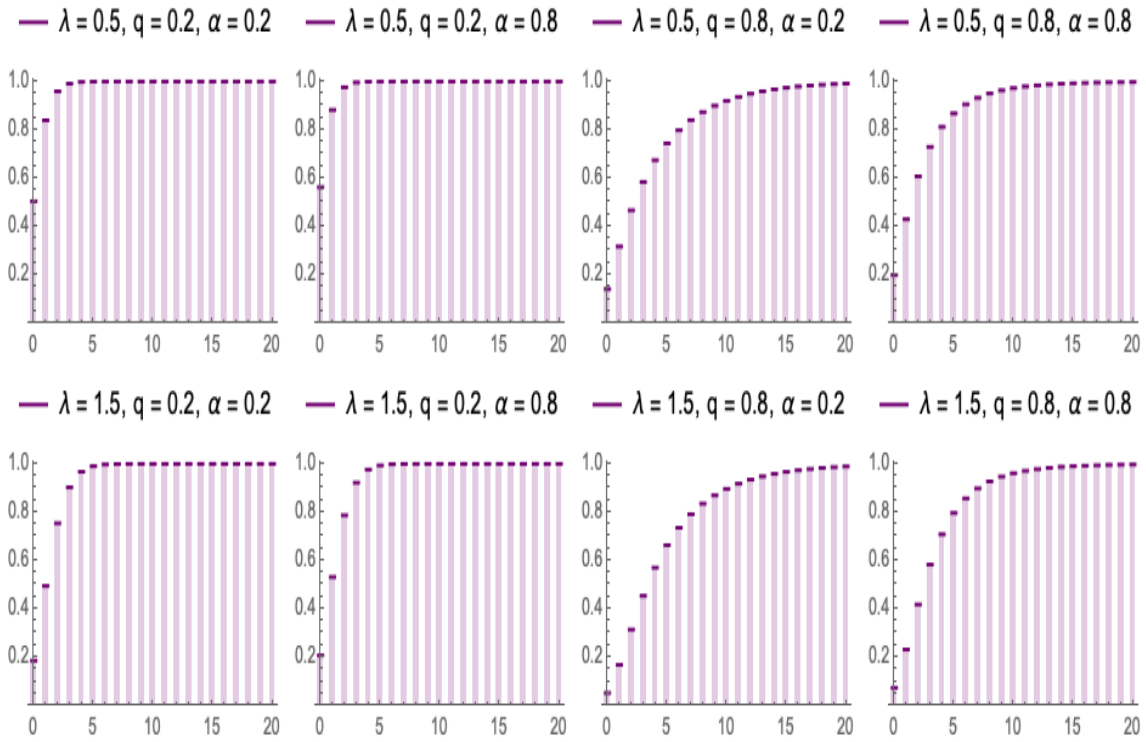

Figure 2: The cdf of $PoiTG(\theta, q, \alpha)$ for multiple choices of θ, q and α

$$s''_y = \sum_{i=0}^{y} \frac{1}{i!} \left(\frac{\theta}{q^2}\right)^i,$$

$$s'_{y+1} = s'_y + \frac{1}{\Gamma(y+2)} \left(\frac{\theta}{q}\right)^{y+1},$$

and

$$s''_{y+1} = s''_y + \frac{1}{\Gamma(y+2)} \left(\frac{\theta}{q^2}\right)^{y+1}.$$

Thus,

$$p_Y(y+1) = \alpha(1-q^2)q^{2(y+1)}e^{-\theta}s''_{y+1} + (1-\alpha)(1-q)q^{y+1}e^{-\theta}s'_{y+1}. \tag{6}$$

Equation 6 is the recurrence relation of the PoiTG distribution. It is easy to verify that

$$\begin{aligned}\frac{p_Y(y+1)}{p_Y(y)} &= q\frac{s'_{y+1} + \dfrac{\alpha(1+q)}{1-\alpha}q^{y+1}s''_{y+1}}{s'_y + \dfrac{\alpha(1+q)}{1-\alpha}q^y s''_y} \\ &= q\left(1 + \frac{\alpha(1+q)q^y s''_y}{(1-\alpha)s'_y}\right)^{-1}. \end{aligned} \tag{7}$$

3.2. Generating functions

We use G to denote the probability generating function (pgf) and to be specific in its use, the corresponding random variable is mentioned in the subscript of the function G. For $Y_1 \sim Poi(\theta)$ and $Y_2 \sim TGD(\alpha, q)$,

$$G_{Y_1}(s) = e^{\theta(s-1)} \quad \text{and} \quad G_{Y_2}(s) = \frac{(1-q)(1+\alpha q(1-s)-q^2 s)}{(1-qs)(1-q^2 s)}, \ |q^2 s| < 1.$$

Since Y_1 and Y_2 are independent, $G_Y(s) = G_{Y_1}(s)G_{Y_2}(s)$. Thus the pgf of $PoiTG(\theta, q, \alpha)$ is

$$G_Y(s) = \frac{e^{\theta(s-1)}(1-q)(1+\alpha q(1-s)-q^2 s)}{(1-qs)(1-q^2 s)}. \tag{8}$$

In similar way, we derive the other generating functions including moment generating function ($M_Y(t)$), characteristic function ($\phi_Y(t)$) and cumulant generating function ($K_Y(t)$). These are provided below as

$$M_Y(t) = \frac{e^{\theta(e^t-1)}(1-q)(1+\alpha q(1-e^t)-q^2 e^t)}{(1-qe^t)(1-q^2 e^t)}, \tag{9}$$

$$\phi_Y(t) = \frac{e^{\theta(e^{it}-1)}(1-q)(1+\alpha q(1-e^{it})-q^2 e^{it})}{(1-qe^{it})(1-q^2 e^{it})}, \tag{10}$$

and

$$K_Y(t) = \log\left[\frac{(1-q)(1+\alpha q(1-e^t)-q^2 e^t)}{(1-qe^t)(1-q^2 e^t)}\right] + \theta(e^t - 1). \tag{11}$$

3.3. Moments, skewness and kurtosis

The r^{th} order raw moment can be computed by taking the r^{th} derivative of the moment generating function specified in (9) with respect to t and evaluating it at $t = 0$.

$$E(Y^r) = M_Y^{(r)}(0) = \frac{d^r}{dt^r}[M_Y(t)]_{t=0}.$$

Let μ'_r denote the r^{th} order raw moment, that is $\mu'_r = E(Y^r)$. The closed form expressions of the first, second, third and fourth moments are

$$\mu'_1 = \theta + \frac{q(1-\alpha)+q^2}{1-q^2}, \tag{12}$$

$$\begin{aligned}\mu'_2 = \frac{1}{(1-q^2)^2}[&\theta(1+\theta)+q(1-\alpha)(1+2\theta)+q^2(3-2\alpha-2\theta^2)\\ &+q^3(1-\alpha)(3-2\theta)+q^4(1-\theta+\theta^2)],\end{aligned} \tag{13}$$

$$\begin{aligned}\mu'_3 = \frac{1}{(1-q^2)^3}[&\theta(\theta^2+3\theta+1)+q(1-\alpha)(3\theta^2+6\theta+1)\\ &-q^2(3\theta^3+6\theta^2-9\theta+6\alpha(\theta+1)-7)-2q^3(1-\alpha)(3\theta^2-8)\\ &+q^4(3\theta^3+3\theta^2-9\theta+6\alpha(\theta-2)+16)+q^5(1-\alpha(3\theta^2-6\theta+7)\\ &-q^6(\theta^3+\theta-1)],\end{aligned} \tag{14}$$

and

$$\begin{aligned}\mu'_4 = \frac{1}{(1-q^2)^4}[&\theta(\theta^3+6\theta^2+7\theta+1)+q(1-\alpha)(4\theta^3+18\theta^2+14\theta+1)\\ &-q^2(4\theta^4+20\theta^3-2\theta^2-46\theta+2\alpha(7+18\theta+6\theta^2)-15)\\ &-q^3(1-\alpha)(12\theta^3+30\theta^2-54\theta-61)\\ &+q^4(6\theta^4+24\theta^3-24\theta^2+8\alpha(3\theta^2-13)+115)\\ &+q^5(1-\alpha)(12\theta^3+30\theta^2-54\theta+115)\\ &-q^6(4\theta^4+12\theta^3-14\theta^2+46\theta+2\alpha(25-18\theta+6\theta^2)-61)\\ &-q^7(1-\alpha)(-15+14\theta-6\theta^2+4\theta^3)+q^8(\theta^4+2\theta^3+\theta^2-\theta+1)].\end{aligned} \tag{15}$$

By utilizing the above formulas for the raw moments, we can derive the explicit formulae for the first, second, four central moments of the random variable Y. The r^{th} order central moment, denoted by μ_r, is defined as $\mu_r = E(Y-\mu'_1)^r$. The explicit expressions for the central moments of the PoiTG distribution are

$$\mu_2 = \frac{\theta(1-q^2)^2+q(1-\alpha+q(2+q(1-\alpha)-\alpha^2))}{(1-q^2)^2}, \tag{16}$$

$$\begin{aligned}\mu_3 = \frac{1}{(1-q^2)^3}(&\theta+q(1-\alpha)-q^2(3\alpha^2+3\theta-4)-2q^3(\alpha^3+2\alpha-3)\\ &-q^4(3\alpha^2-3\theta-4)+q^5(1-\alpha)-q^6\theta),\end{aligned} \tag{17}$$

and

$$\begin{aligned}\mu_4 = \frac{1}{(1-q^2)^4}((&1+q^8)\theta(3\theta+1)+(q+q^7)(1-\alpha)(6\theta+1)-(q^2+q^6)\\ &(\alpha^2(6\theta+4)+6\alpha+12\theta^2-8\theta-11)+(q^3+q^5)(1-\alpha)\\ &(6\alpha^2+12\alpha-6\theta+35)+q^4(3\alpha^4-4\alpha^2(3\theta-7)+12\alpha-2(9\theta^2-9\theta+25))).\end{aligned} \tag{18}$$

Here, explicit expressions for skewness and kurtosis are not provided because of the computational complexity and the involvement of large equations. Instead, 3-D surface plots for these measures are presented in Figure 3 and 4, respectively.

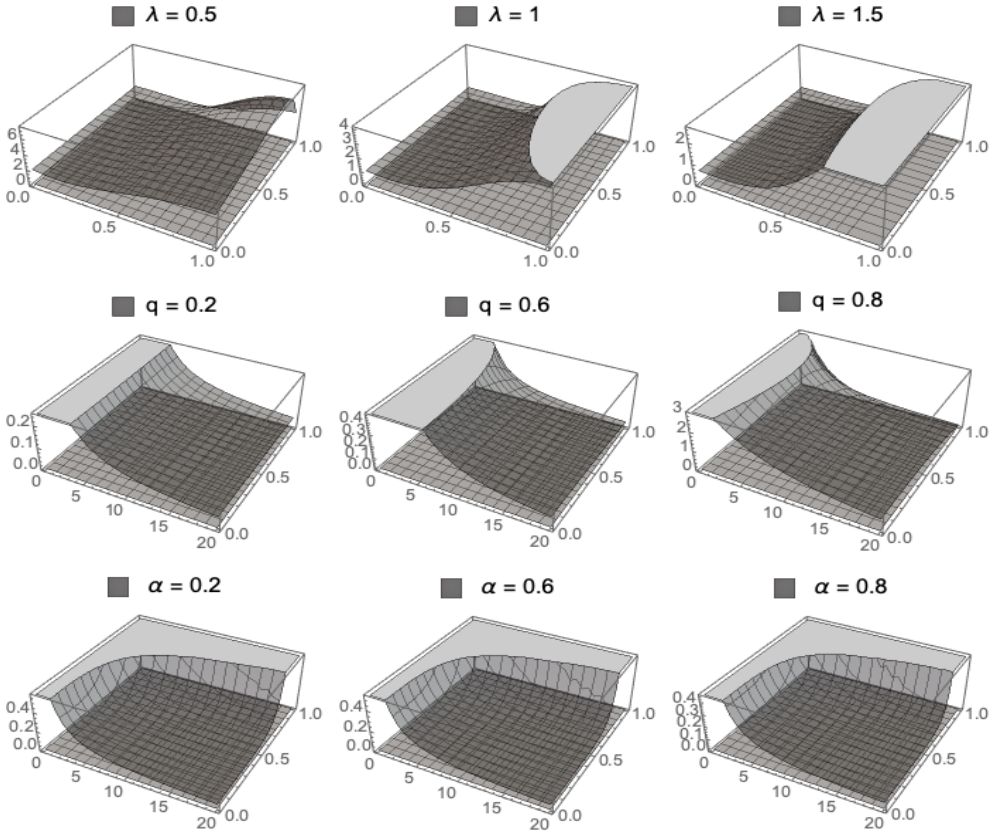

Figure 3: Skewness of $PoiTG(\theta, q, \alpha)$ for different combinations of θ, q and α

Figure 3 represents three rows of surface plots. Each row corresponds to a different combination of fixed parameters: θ, q, and α. The first row shows the surface plot of skewness with respect to q and α, while keeping θ constant. The second row depicts the surface plot of skewness in relation to θ and α, with q held constant. The third row exhibits the surface plot of skewness with respect to θ and q, while maintaining α at a fixed value. Notably, all the surface plots indicate positive skewness for any combination of θ, q, and α.

Similarly, in Figure 4, the same parameter combinations are used for each row as in Figure 3. A horizontal surface is drawn at a value of 3 in each graph, which never intersects the kurtosis surface, indicates the leptokurtic nature of the PoiTG distribution. The second row shows that the parameter q significantly influences kurtosis, with higher q values leading to greater leptokurticity. In contrast, the first and third rows suggest that for large θ and α, the distribution exhibits no substantial variation in peakedness.

3.4. Index of dispersion

The index of dispersion (Hoel, 1943), denoted by ID_Y, provides a measure of the level of dispersion of a distribution. It assesses whether a distribution is appropriate for

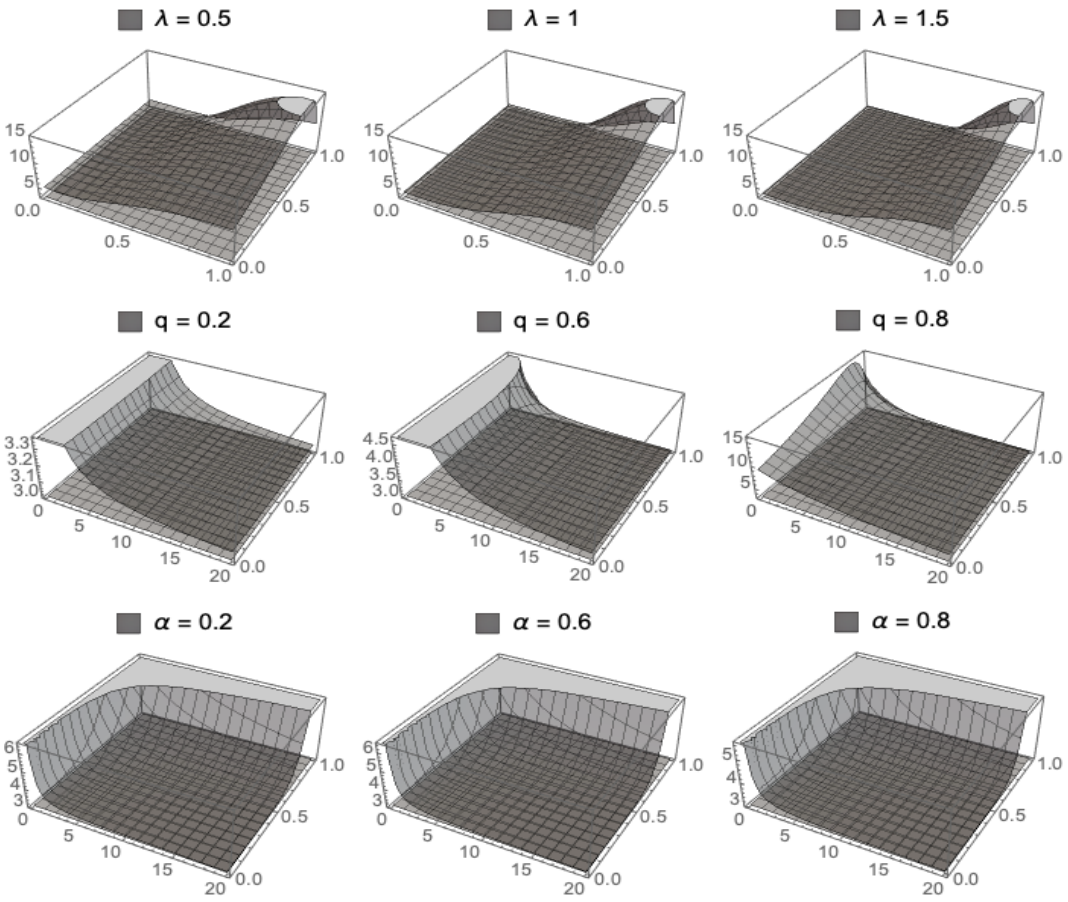

Figure 4: Kurtosis of $PoiTG(\theta, q, \alpha)$ for different combinations of θ, q and α

modeling an overdispersed, underdispersed, or equidispersed dataset. If ID_Y is greater than one, the distribution of Y can accommodate overdispersion, and if ID_Y is less than one, the distribution can accommodate underdispersion. The distribution is said to be equidispersed when $ID_Y = 1$. The dispersion index of $PoiTG(\theta, q, \alpha)$ is given by

$$ID_Y = \frac{\theta(1-q^2)^2 + q(1-\alpha+q(2+q(1-\alpha)-\alpha^2))}{(1-q^2)(\theta(1-q^2)+q(1-\alpha)+q^2)}.$$

We investigate the nature of ID_Y in Figure 5 for different combinations of θ, q, and α. The plots in the first row exhibits the fact that ID_Y is slightly influenced by θ. As the value of θ increases, the corresponding distribution becomes less overdispersed. The values of ID_Y for $\theta = 1.5$ suggest a lower level of overdispersion compared to the case where $\theta = 0.5$ and $\theta = 1$. The plots in the second row shows that ID_Y is greatly influenced by the choice of q values. As the value of q increases, the corresponding distribution becomes more overdispersed. The values of ID_Y for $q = 0.6$ and $q = 0.8$ suggest a higher level of overdispersion compared to the case where $q = 0.2$. The plots in the third row shows that ID_Y is also substantially influenced by the choice of α values. As the value of α increases, the corresponding distribution becomes less overdispersed.

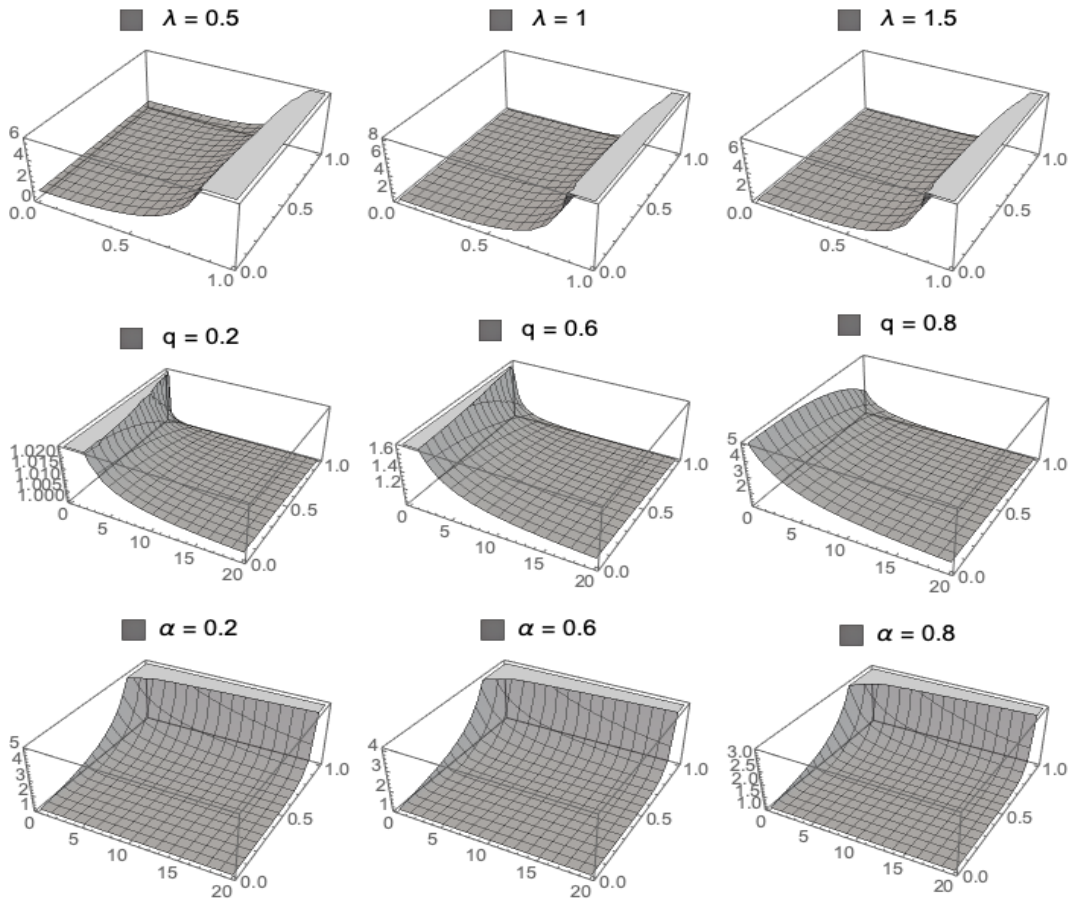

Figure 5: Index of dispersion of $PoiTG(\theta, q, \alpha)$ for multiple combinations of θ, q and α

3.5. Reliability properties

The survival function(sf) is important in reliability analysis and is used to quantify the probability that a system will survive beyond a certain number of events. It is also known as the reliability function because it provides information about the reliability of a system. The sf of $Y \sim PoiTG(\theta, q, \alpha)$ is

$$S_Y(y) = 1 - \frac{\Gamma(y, \theta)}{\Gamma y} + (1 - \alpha)W_1(y - 1, \theta, q) + \alpha W_2(y - 1, \theta, q). \tag{19}$$

In the above expression,

$$W_k(y - 1, \theta, q) = \frac{q^{ky}}{\Gamma y} \exp\left(\frac{\theta(1 - q^k)}{q^k}\right)\Gamma\left(y, \frac{\theta}{q^k}\right), \qquad k = 1, 2\,.$$

The hazard rate function (hrf) of $Y \sim PoiG(\theta, q, \alpha)$ is

$$h_Y(y) = \frac{(1 - q)q^y\left(\alpha(1 + q)e^{\frac{\theta}{q^2}}q^y\Gamma\left(y + 1, \frac{\theta}{q^2}\right) + (1 - \alpha)e^{\frac{\theta}{q}}\Gamma\left(y + 1, \frac{\theta}{q}\right)\right)}{ye^{\theta}(\Gamma y - \Gamma(y, \theta)) + q^y y\left(\alpha e^{\frac{\theta}{q^2}}q^y\Gamma\left(y, \frac{\theta}{q^2}\right) + (1 - \alpha)e^{\frac{\theta}{q}}\Gamma\left(y, \frac{\theta}{q}\right)\right)}. \tag{20}$$

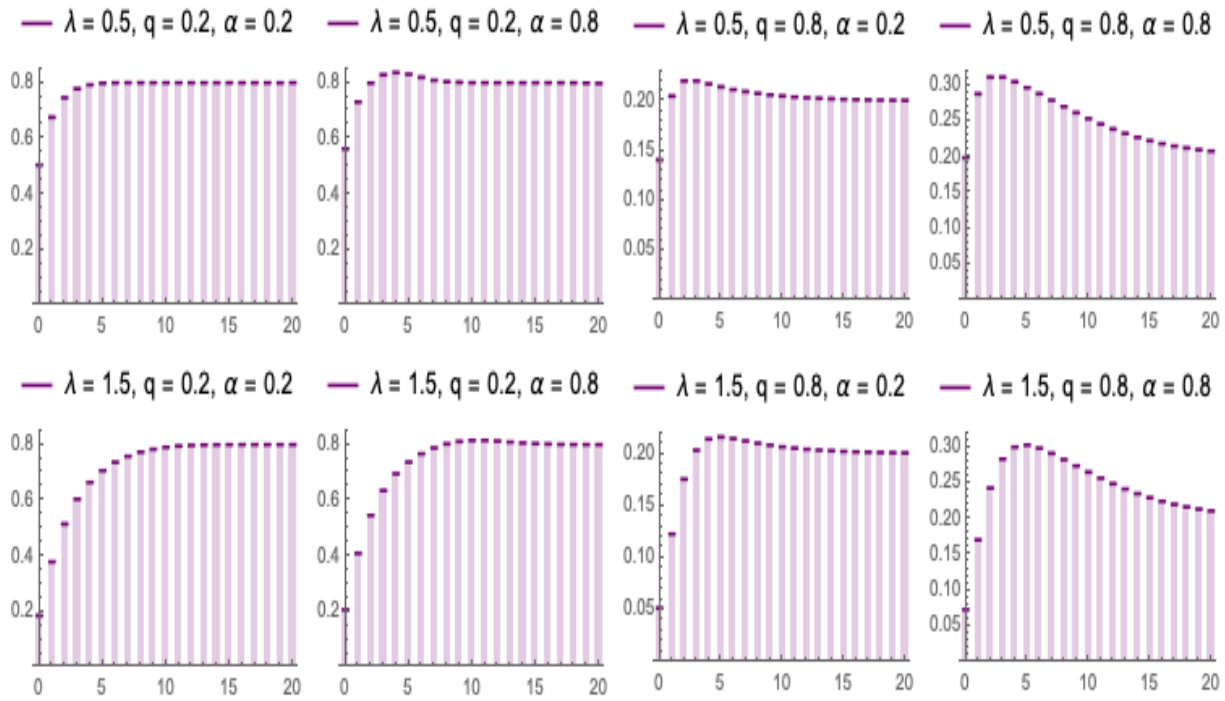

Figure 6: The hrf of $PoiTG(\theta, q, \alpha)$ for multiple choices of θ, q and α

Figure 6 presents the hrf for different parameter settings. In the PoiTG distribution, the behavior of the hrf varies based on parameter values. For small θ, the distribution exhibits a constant failure rate. As θ increases, the failure rate initially rises to a certain point before stabilizing at a constant rate. Conversely, as α increases, the failure rate decreases after a certain time and then remains constant. The mean residual life of Y is

$$\mu_Y(y) = \frac{\sum_{y=t}^{\infty}\left(1 - \dfrac{\Gamma(y,\theta)}{\Gamma y} + (1-\alpha)W_1(y-1,\theta,q) + \alpha W_2(y-1,\theta,q)\right)}{1 - \dfrac{\Gamma(t-1,\theta)}{(\Gamma t - 1)} + (1-\alpha)W_1(t-2,\theta,q) + \alpha W_2(t-2,\theta,q)}. \tag{21}$$

In the derivation of the expression in (21), $\bar{F}(y)$ denotes $1-F(y-1)$, and similarly $\bar{F}(t-1) = 1-F(t-2)$. The functions $W_k(y-1,\theta,q)$ and $W_k(t-2,\theta,q)$ hold their meaning as in (19).

3.6. Shannon entropy

Shannon entropy measures the randomness in a distribution. A higher entropy implies the distribution is more spread out across many possible values, whereas a lower entropy implies the distribution is more concentrated at a few values. Let $p_Y(y)$ be the pmf of random variable $Y \sim PoiTG(\theta, q, \alpha)$ given in (4). The Shannon entropy (Shannon, 1948) of Y can be written as

$$S(\theta, q, \alpha) = -\sum_{y=0}^{\infty}((1-\alpha)w_1(y,\theta,q) + \alpha w_2(y,\theta,q))\log((1-\alpha)w_1(y,\theta,q) + \alpha w_2(y,\theta,q)). \tag{22}$$

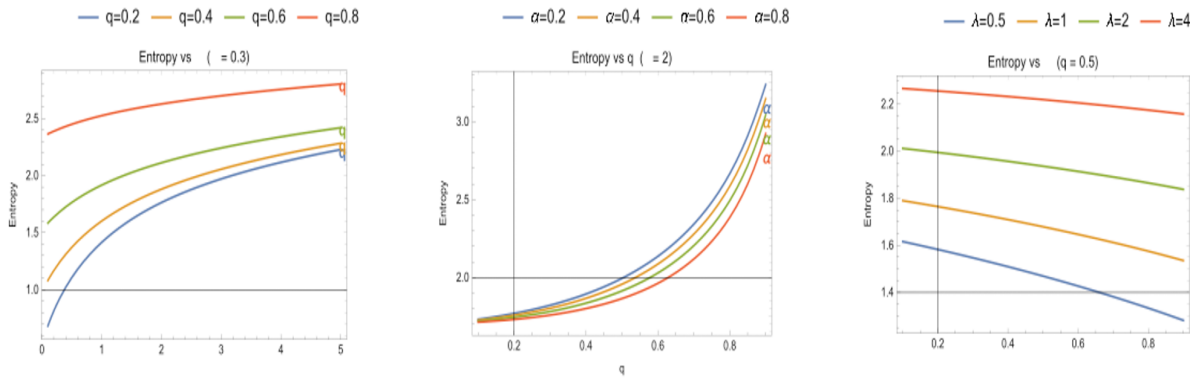

Figure 7: Shannon entropy of PoiTG(λ, q, α)

Figure (7) depicts the Shannon entropy for different parameter choices with the maximum value of y set to 50. The results show that entropy increases with larger values of λ and q, whereas it decreases as α increases. The rising entropy curves for λ and q show the effect of overdispersion increases randomness whereas the falling entropy curves for α illustrate how underdispersion reduces randomness.

4. Parameter estimation

Let $\mathbf{Y}$ be a sample consisting of n observations selected from the $PoiTG(\theta, q, \alpha)$, and a realization on $\mathbf{Y}$ is $(y_1, y_2, \ldots, y_n)$. Here we obtain the maximum likelihood estimates (MLE) of the parameters by two procedures, by direct numerical optimization and by the expectation-maximization (EM) algorithm. Additionally, we derive asymptotic confidence limits for the parameters using the large sample distribution of the MLE.

4.1. Maximum likelihood estimator

Using the pmf of Y in (4), the log-likelihood function of the PoiTG distribution can easily be written as

$$l(\theta, q, \alpha; \boldsymbol{y}) = \sum_{i=1}^{n} \log((1-\alpha)w_{i1} + \alpha w_{i2}). \tag{23}$$

In the above expression,

$$w_{i1} = (1-q)e^{\theta(1-q)/q}\frac{q^{y_i}}{y_i!}\Gamma\left(y_i+1, \frac{\theta}{q}\right), \text{ and}$$
$$w_{i2} = (1-q^2)e^{\theta(1-q^2)/q^2}\frac{q^{2y_i}}{y_i!}\Gamma\left(y_i+1, \frac{\theta}{q^2}\right),$$

for each $i = 1, 2, ..., n$. We list some useful notations below to be used in writing the score functions. For each $i = 1, 2, ..., n$,

$$u_{ij} = \frac{e^{-\theta} q^{j(y_i-1)}}{y_i!} \left(\frac{\theta}{q^j}\right)^{y_i} (1-q^j),$$
$$v_{ij} = \frac{e^{\theta(1-q^j)/q^j} q^{jy_i}}{y_i!} \Gamma\left(y_i+1, \frac{\theta}{q^j}\right), \text{ and}$$
$$q_{jk} = \frac{1-q^j}{q^k}, \qquad j = 1,2 \text{ and } k = 0,1,2,3.$$

The score functions are obtained by differentiating (23) with respect to the parameters, these are given as

$$s_1(\theta, q, \alpha; \boldsymbol{y}) = \sum_{i=1}^{n} \left(\frac{\alpha(u_{i1} - u_{i2} + q^2 q_{22}^2 v_{i2} - q q_{11}^2 v_{i1}) + u_{i1} + q q_{11}^2 v_{i1}}{\alpha q_{20} v_{i2} + (1-\alpha) q_{10} v_{i1}} \right), \tag{24}$$

$$s_2(\theta, q, \alpha; \boldsymbol{y}) \tag{25}$$
$$= \sum_{i=1}^{n} \left(\frac{2\alpha(\frac{\theta u_{i2}}{q} - q v_{i2} - \theta q_{23} v_{i2} + q_{21} v_{i2} y_i) + (1-\alpha)(\frac{\theta u_{i1}}{q} - v_{i1} - \theta q_{12} v_{i1} + q_{11} v_{i1} y_i)}{\alpha q_{20} v_{i2} + (1-\alpha) q_{10} v_{i1}} \right), \tag{26}$$

and

$$s_3(\theta, q, \alpha; \boldsymbol{y}) = \sum_{i=1}^{n} \left(\frac{q_{20} v_{i2} - q_{10} v_{i1}}{\alpha q_{20} v_{i2} + (1-\alpha) q_{10} v_{i1}} \right). \tag{27}$$

Due to the structural complexity of equations (24), (26), and (27), obtaining explicit expressions for the MLE is challenging. Therefore, we employ numerical optimization methods to maximize the log-likelihood function with respect to the parameters. Specifically, we use the *constrOptim* function in *R*, which adequately implements the Nelder-Mead method. Let $\hat{\theta}_{ML}$, $\hat{q}_{ML}$, and $\hat{\alpha}_{ML}$ denote the MLEs of θ, q, and α, respectively.

To obtain the information matrix, the second order partial derivatives of the function given in (23) with respect to θ, q, and α must be computed. However, deriving exact expressions for all second-order partial derivatives is analytically intractable.

The recently introduced $PoiG(\theta, q)$ distribution (Nandi, *et al.*, 2024b) is a special case of the $PoiTG(\theta, q, \alpha)$ distribution when the parameter α is set to zero. Now, we express the pmf of $PoiTG(\theta, q, \alpha)$ as a mixture of the pmf of $PoiG(\theta, q)$ provided in (2). Note that, $Y \sim PoiTG(\theta, q, \alpha)$ can be expressed as

$$Y = \begin{cases} X_1 \text{ with probability } (1-\alpha), \\ X_2 \text{ with probability } \alpha, \end{cases}$$

where $X_1 \sim PoiG(\theta, q)$ and $X_2 \sim PoiG(\theta, q^2)$ are independent. Thus, the pmf is $p_Y(y) = (1-\alpha)p_{X_1}(y) + \alpha p_{X_2}(y)$.Using this formulation, we have derived the Fisher's information matrix and constructed confidence intervals for all parameters of $PoiTG(\theta, q, \alpha)$ (see Appendix I).

4.2. Expectation maximization algorithm

The expectation-maximization (EM) algorithm is a useful iterative procedure to compute MLEs of parameters of a mixture distribution. First, the mixture distribution is represented in such a way that the mixing coefficient is the parameter of a latent Bernoulli random variable. As pointed out earlier, $Y = (1-\alpha)X_1 + \alpha X_2$, where $X_1 \sim PoiG(\theta, q)$ and $X_2 \sim PoiG(\theta, q^2)$. We consider Z to be another random variable such that, Z is independent of X_1 and X_2, and Y is expressed as the following linear combination of X_1 and X_2.

$$Y = (1-Z)X_1 + ZX_2. \tag{28}$$

Clearly, a reasonable assumption is $Z \sim Bernoulli(\alpha), 0 < \alpha < 1$. The estimates obtained by this method are consistent and unique (Dempster*, et al.* (1977), Redner and Walker (1984)). The EM algorithm iterates through two steps: the Expectation (E-step) and the Maximization (M-step). We have n iid copies of Y. However, in the incomplete-data framework of EM algorithm, we find that observations of Z are not available. Thus, the hypothetical complete dataset is $(Y_i, Z_i) : i = 1, 2, ..., n$. Under the formulation, E-step of each EM iteration requires the expectation of $(Z|Y; \Psi^{(k)})$, where $\Psi^{(k)} = (\theta^{(k)}, q^{(k)}, \alpha^{(k)})$ is the estimate of $\Psi = (\theta, q, \alpha)$ in the k^{th} iteration. Since,

$$Z_i|Y_i, \Psi^{(k)} \sim Bernoulli(\alpha_i^{(k)}),$$

$$\alpha_i^{(k+1)} = \frac{\alpha^{(k)}(1-q^{(k)^2})2^{2y_i}\, e^{\frac{\theta^{(k)}(1-q^{(k)^2})}{q^{(k)^2}}}\, \Gamma\left(y_i+1, \frac{\theta^{(k)}}{q^{(k)^2}}\right)}{\alpha^{(k)}(1-q^{(k)^2})2^{2y_i}\, e^{\frac{\theta^{(k)}(1-q^{(k)^2})}{q^{(k)^2}}}\, \Gamma\left(y_i+1, \frac{\theta^{(k)}}{q^{(k)^2}}\right) + (1-\alpha^{(k)})(1-q^k)q^{y_i}\, e^{\frac{\theta^{(k)}(1-q^{(k)})}{q^{(k)}}}\, \Gamma\left(y_i+1, \frac{\theta^{(k)}}{q^{(k)}}\right)}. \tag{29}$$

Now, it is clear that ,

$$E[Z_i|Y_i, \Psi^{(k)}] = \alpha_i^{k+1} \text{and } V[Z_i|Y_i, \Psi^{(k)}] = \alpha_i^{k+1}(1-\alpha_i^{k+1}). \tag{30}$$

We estimate α for the $(k+1)^{th}$ iteration by

$$\alpha^{k+1} = \frac{1}{n}\sum_{i=1}^{n} \alpha_i^{k+1}. \tag{31}$$

Then, we proceed to M-step, where we maximize the log-likelihood of observed $(y_1, y_2, ..., y_n)$ on $(Y_1, Y_2, ..., Y_n)$ with respect to θ and q for given $\alpha^{(k+1)}$ and obtain $\theta^{(k+1)}$ and $q^{(k+1)}$. That is,

$$(\theta^{(k+1)}, q^{(k+1)}) = \arg \max_{\theta,q} l(\theta, q, \alpha^{(k+1)}; \vec{y}).$$

We continue iterating the successive E-step and M-step until $|\theta^{(k+1)} - \theta^{(k)}| < \epsilon$, $|q^{(k+1)} - q^{(k)}| < \epsilon$, and $|\alpha^{(k+1)} - \alpha^{(k)}| < \epsilon$, simultaneously. We set $\alpha_0 = 0.5$ and $\epsilon = 0.0001$ in the implementation of EM algorithm.

5. Simulation study

The initial aim is to extract a sample of size n from $Y \sim PoiTG(\theta, q, \alpha)$. We obtain two samples, each of size n, namely $\mathbf{y_1}$ from $Y_1 \sim Poi(\theta)$ and $\mathbf{y_2}$ from $Y_2 \sim TGD(q, \alpha)$. Since Y is the sum of Y_1 and Y_2, a sample from the PoiTG distribution can be obtained as $\mathbf{y} = \mathbf{y_1} + \mathbf{y_2}$.

The main objective is to assess the performance of the MLEs obtained using a general optimization approach and the estimates obtained using the EM algorithm. We generate a sample of size n from the PoiTG distribution for 8 different combinations of the parameters θ, q, and α. Additionally, we consider six different values of n, such as 30, 50, 100, 250, 500 and 1000. The MLEs obtained through the EM algorithm are denoted as $\hat{\theta}_{EM}$, $\hat{q}_{EM}$, and $\hat{\alpha}_{EM}$. For each sample of size n, we obtain the MLEs using both approaches mentioned in Section 4.1 and Section 4.2. Subsequently, we calculate the bias and squared error of these estimates.

To evaluate the performance of the estimators, we replicate the experiment 10,000 times for each value of n and each fixed combination of parameter values. We then compute the average bias and average mean squared error (MSE). Table 5 (see Appendix I) displays the values of the corresponding average biases for $\hat{\theta}_{ML}$, $\hat{q}_{ML}$, $\hat{\alpha}_{ML}$, $\hat{\theta}_{EM}$, $\hat{q}_{EM}$, and $\hat{\alpha}_{EM}$ in the non-shaded regions. The shaded regions represent the average MSEs of these estimates.

From Table 5, it can be observed that the average biases and the average MSEs of $\hat{\theta}_{ML}$, $\hat{\theta}_{EM}$, $\hat{q}_{ML}$, and $\hat{q}_{EM}$ decreases with the increase in sample size n. In fact, as the sample size increases from 30 to 1000, a consistent pattern is observed, where both the average bias and the MSE decrease with larger n. It is also observed that for particular combination of n, θ, and greater value of q the average biases and mean squared errors of $\hat{\theta}_{ML}$ and $\hat{\theta}_{EM}$ decrease as α increases. However, for the same combinations, the average mean squared error of $\hat{q}_{ML}$ and $\hat{q}_{EM}$ exhibits an opposite trend. When keeping n, q, and α fixed, the average mean squared errors of both the estimators $\hat{\theta}_{ML}(\hat{\theta}_{EM})$ and $\hat{q}_{ML}$ $(\hat{q}_{EM})$ increase with θ. Similarly, for fixed n, θ, and α, the average mean squared errors of $\hat{\theta}_{ML}$ and $\hat{\theta}_{EM}$ increase with q, while the same for $\hat{q}_{ML}$ and $\hat{q}_{EM}$ decrease with increasing q. On the other hand, the biases and MSEs of $\hat{\alpha}_{ML}$ and $\hat{\alpha}_{EM}$ do not exhibit any clear trend. Additionally, it is noteworthy that for any combination of parameters, the average biases and average mean squared errors of the maximum likelihood estimators obtained using the EM algorithm ($\hat{\theta}_{EM}$, $\hat{q}_{EM}$, and $\hat{\alpha}_{EM}$) are smaller compared to their counterparts for the optimization approach ($\hat{\theta}_{ML}$, $\hat{q}_{ML}$, and $\hat{\alpha}_{ML}$).

6. Goodness-of-Fit evaluation in data analysis

This section highlights the flexibility of the PoiTG distribution in effectively modeling diverse datasets across various fields. The goodness of fit is assessed using the Chi-square (χ^2) test along with its associated P-value. Additionally, the fit of the PoiTG distribution is compared with that of other competing distributions listed below.

Table 1: Competing models

Distribution	Abbreviation	Author(s)
Discrete log-logistic	DsLogL	Para and Jan (2016a)
Discrete inverse Weibull	DsIW	Jazi, *et al.* (2010)
Discrete Inverted Nadarajah-Haghighi	DsINH	Singh, *et al.* (2022)
Discrete Burr type II	DsBX-II	Para and Jan (2016b)
Discrete Belal	DsBL	Altun, *et al.* (2022)
Geometric	Geo	-
Discrete Rayleigh	DsR	Roy (2004)
Discrete inverse Rayleigh	DsIR	Hussain and Ahmad (2014)
Poisson	Poi	-
Discrete Burr-Hatke	DsBH	El-Morshedy, *et al.* (2020)
Discrete Pareto	DsPa	Krishna and Pundir (2009)
Poisson Geometric	PoiG	Nandi, *et al.* (2024b)
Transmuted Geometric	TGD	Chakraborty and Bhati (2016)

6.1. Dataset I

The data set is sourced from (https://www.worldometers.info/coronavirus/country/south-korea/) and consists of daily COVID-19 fatalities in South Korea from February 15 to December 12, 2020. Figure 8 provides a primary graphical representation of these data using nonparametric methods. In Figure 8, it is observed that the frequency distribution is unimodal, has a mode at zero, exhibits a heavier tail compared to the Poisson distribution, and contains a small proportion of outliers. Furthermore, Table 2 presents the observed frequency (OF), expected frequency (EF), maximum likelihood estimators (MLEs) of the parameters, log-likelihood ($-L$), and Chi-square test values along with their corresponding p values for each competing distribution applied to Data set I. In particular, Table 2 shows that the χ^2 test fails to reject the hypothesis of a good fit only for the PoiTG distribution.

Among all models tested, the PoiTG distribution exhibits superior performance as the optimal fit for this data set. This conclusion is supported by its lowest χ^2 value compared to other distributions. The estimated pmfs for Data set I are illustrated in Figure 9.

6.2. Dataset II

The data set in question captures the number of computer malfunctions over 128 successive weeks of operation, as documented by Hand, *et al.* (1995). Non-parametric graphs are recommended for the first visual analysis of this data, with the findings illustrated in Figure 10. From Figure 10, we note that the frequency distribution is multimodal, has maximum frequency at 2, has a heavy tail and a moderate proportion of outliers. Table 3 outlines the OF, EF, MLEs for the parameters, $-L$, and Chi-square test results, along with their respective P-values, for each competing distribution within dataset II.

From Table 3 we note that the χ^2 test does not reject the hypothesis of a good fit for the PoiTG distribution. Among all the tested models, the PoiTG distribution and others distributions such as PoiG and TGD perform in similar manner for this dataset. The estimated pmfs for Dataset II are illustrated in Figure 11.

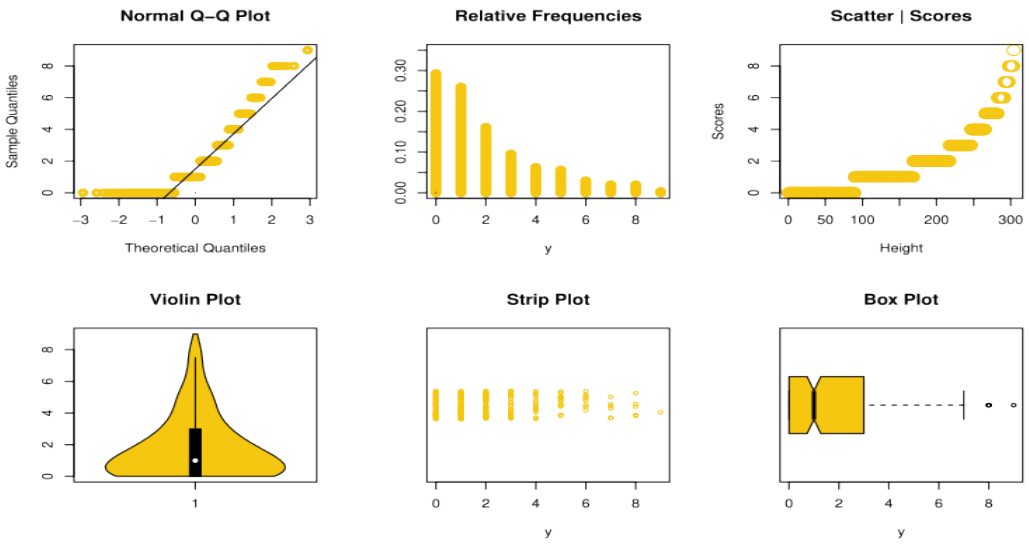

Figure 8: Plots based on non-parametric methods for dataset I

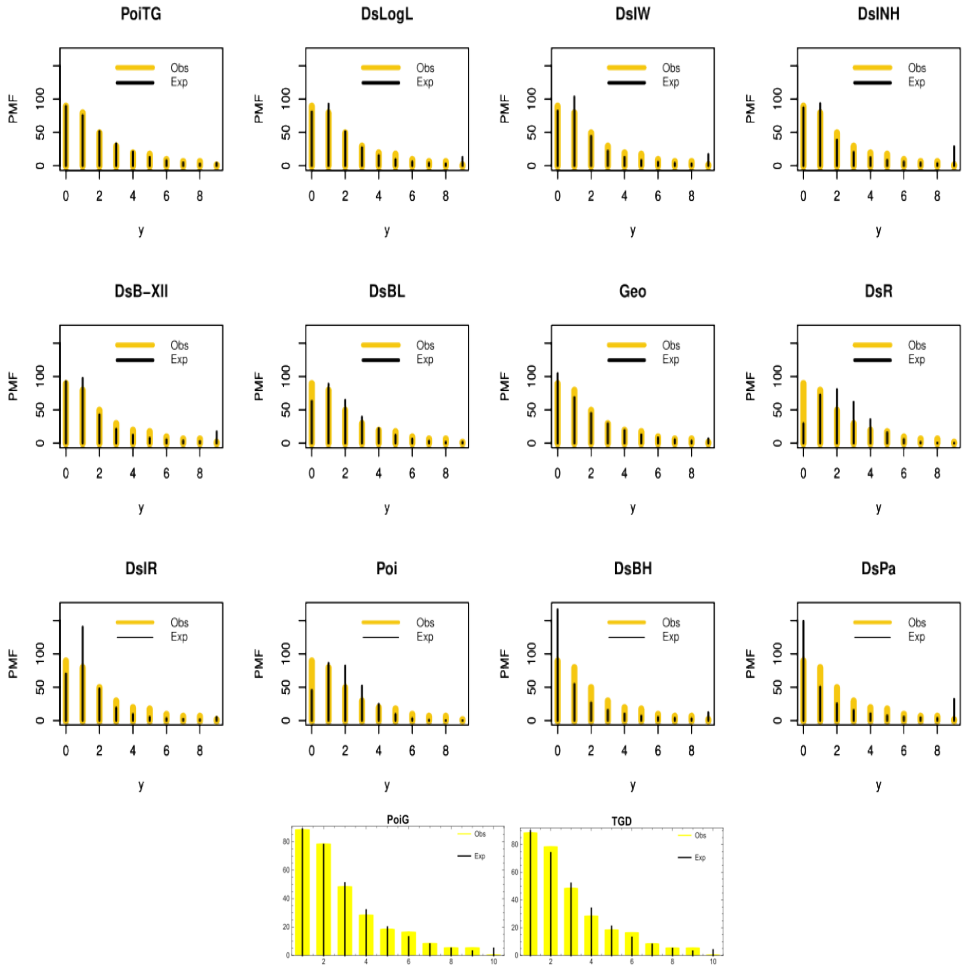

Figure 9: The estimated probability mass functions for dataset I

Table 2: The MLEs and goodness-of fit measures of all competing distributions for dataset I

No. ECB	OF	Expected frequency (EF)												
		PoiTG	TGD	PoiG	DsLogL	DsIW	DsINH	DsB-XII	DsBL	Geo	DsR	DsIR	Poi	DsBH	DsPa
0	89	89.423	89.772	89.129	80.931	82.351	87.190	92.887	63.089	104.780	29.499	69.890	45.408	166.600	149.356
1	79	75.642	74.420	78.023	92.775	103.702	93.636	97.788	89.133	68.665	72.407	140.613	86.336	54.598	50.500
2	49	51.726	51.965	51.414	51.431	44.390	38.648	42.676	64.868	44.998	80.783	47.685	82.074	26.667	25.478
3	29	33.332	33.767	32.241	27.336	22.317	20.378	21.174	39.578	29.489	61.938	19.125	52.014	15.540	15.379
4	19	20.893	21.168	20.082	15.559	12.926	12.508	12.172	22.307	19.325	35.679	9.333	24.725	10.017	10.312
5	17	12.898	13.015	12.500	9.518	8.248	8.442	7.754	12.037	12.664	15.984	5.198	9.402	6.885	7.387
6	9	7.893	7.914	7.780	6.193	5.636	6.076	5.308	6.328	8.299	5.664	3.162	2.979	4.953	5.533
7	6	4.805	4.781	4.842	4.247	4.052	4.581	3.831	3.273	5.439	1.603	2.098	0.809	3.682	4.408
8	6	2.917	2.878	3.014	3.061	3.031	3.576	2.879	1.674	3.564	0.364	1.429	0.192	2.808	3.466
9	1	4.471	4.316	5.008	12.949	17.347	28.965	17.531	1.713	6.777	0.079	5.467	0.061	12.25	32.181
Total	304	304	304	304	304	304	304	304	304	304	304	304	304	304	304
$-L$		565.047	565.087	565.258	577.011	586.855	596.877	587.652	575.338	568.083	638.905	606.870	621.098	620.466	633.531
MLE_α		0.275	−0.441	–	1.716	0.271	6.241	0.591	0.707	0.655	0.903	0.229	1.901	0.904	0.377
MLE_q		0.604	0.598	0.377	1.878	1.411	0.139	2.466	–	–	–	–	–	–	–
MLE_θ		0.115	–	0.2529	–	–	–	–	–	–	–	–	–	–	–
χ^2		2.771	2.9967	3.59	25.019	41.868	51.008	44.784	28.203	11.767	184.989	92.204	115.896	109.333	128.631
d.f		4	5	6	6	6	6	6	6	7	5	6	4	6	7
P-value		0.597	0.7005	0.7319	≤ 0.001	≤ 0.001	≤ 0.001	≤ 0.001	≤ 0.001	0.109	≤ 0.001	≤ 0.001	≤ 0.001	≤ 0.001	≤ 0.001

Table 3: The MLEs and goodness-of fit measures for all competing distributions in dataset II

X	OF	PoiTG	TGD	PoiG	DsLogL	DsIW	DsINH	DsB-XII	DsBL	Geo	DsR	DsIR	Poi	DsBH	DsPa
0	15	13.754	14.871	13.758	11.012	9.689	12.401	19.886	9.7251	25.527	3.636	3.237	2.310	65.853	47.806
1	19	21.773	20.205	21.772	23.309	33.087	30.161	36.588	19.441	20.434	10.302	47.809	9.272	21.915	19.190
2	23	21.140	19.678	21.137	23.010	23.104	19.935	19.052	20.731	16.360	15.306	34.021	18.613	10.931	10.760
3	14	17.242	16.958	17.238	17.906	14.481	12.782	10.690	18.511	13.096	18.041	16.649	24.912	6.538	7.021
4	15	13.267	13.740	13.265	12.868	9.539	8.721	6.832	15.191	10.491	18.440	8.773	25.014	4.342	5.001
5	10	10.045	10.741	10.044	9.105	6.637	6.291	4.762	11.873	8.385	16.918	5.079	20.085	3.088	3.774
6	8	7.577	8.211	7.577	6.506	4.840	4.732	3.520	8.993	6.720	14.172	3.174	13.440	2.304	2.967
7	4	5.712	6.184	5.712	4.738	3.653	3.693	2.724	6.666	5.381	10.944	2.106	7.713	1.782	2.404
8	6	4.305	4.612	4.307	3.525	2.849	2.950	2.171	4.865	4.312	7.839	1.466	3.870	1.417	1.994
9	2	3.245	3.415	3.246	2.677	2.273	2.412	1.774	3.510	3.456	5.228	1.059	1.736	1.151	1.686
10	3	2.446	2.517	2.447	2.073	1.848	2.013	1.482	2.510	2.753	3.256	0.789	0.690	0.953	1.447
11	3	1.843	1.848	1.844	1.633	1.533	1.691	1.251	1.783	2.205	1.897	0.604	0.253	0.800	1.258
12	2	1.389	1.354	1.390	1.307	1.284	1.460	1.080	1.260	1.766	1.036	0.472	0.080	0.680	1.106
+13	4	4.254	3.660	4.257	8.331	13.183	18.758	16.188	2.941	7.114	0.985	2.762	0.0062	6.246	21.586
Total	128	128	128	128	128	128	128	128	128	128	128	128	128	128	128
$-l$		311.638	311.038	311.663	319.854	330.446	331.931	342.581	318.545	320.703	347.148	356.525	384.974	379.346	369.766
MLE_α		< 0.001	–	-0.787	3.337	0.076	1.244	0.785	0.831	0.801	0.972	0.025	4.016	0.971	0.509
MLE_q		0.753	0.728	–	1.960	1.235	1.634	3.309	–	–	–	–	–	–	–
MLE_θ		0.829	0.727	0.829	–	–	–	–	–	–	–	–	–	–	–
χ^2		1.540	1.580	1.539	6.018	18.709	20.199	40.183	6.361	11.651	49.897	50.533	88.995	143.174	94.971
d.f		6	7	7	7	6	6	5	8	9	8	5	6	5	6
P-value		0.9567	0.9793	0.9809	0.537	0.005	0.003	0	0.607	0.234	0	0	0	0	0

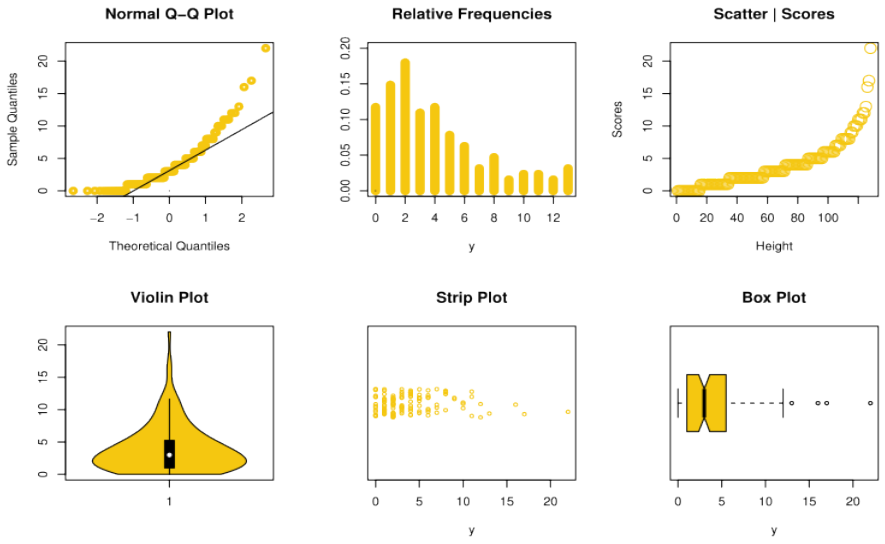

Figure 10: Plots based on non-parametric methods for dataset II

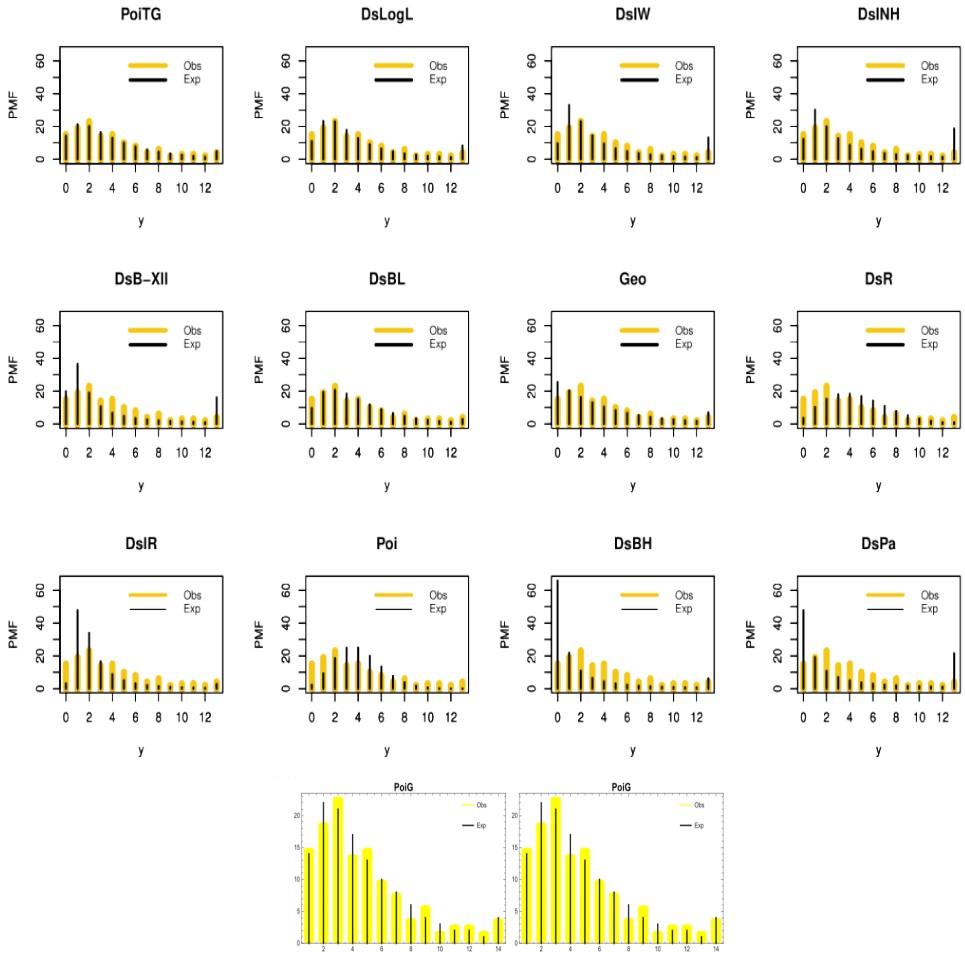

Figure 11: The estimated probability mass functions for dataset II

Remark 2: The PoiTG distribution provides a noticeably better fit compared to most of the competing models considered for Datasets I & II. It is noted, however, that its performance is equivalent to that of PoiG and TGD. Despite this equivalence, the PoiTG distribution remains particularly useful, as it unifies different hazard rate shapes within a single framework and demonstrates greater flexibility when extended to regression settings, as shown in Section 7. This makes it a strong candidate for wider application in future case studies.

7. Regression models: theory and practical application

In this section, we reparameterize the PoiTG model in terms of its mean (μ) and the two parameters q and α for modeling overdispersed count responses. Using this reparametrization, we develop a new count regression model within the framework of generalized linear models (GLM). The $PoiTG(\theta, q, \alpha)$ model is re-parametrized by considering $\theta = \mu - g(q, \alpha)$, where $g(q, \alpha) = [q(1-\alpha) + q^2]/[1-q^2]$ with the restriction that $\mu > g(q, \alpha)$ since $\theta > 0$. The reparameterized PoiTG distribution is is denoted as $PoiTG(\mu, q, \alpha)$ with $\mu > g(q, \alpha)$, $0 < q < 1$ and $0 < \alpha < 1$. Under this reparameterization, the pmf of the PoiTG distribution in (4) can be rewritten as

$$p_Y(y) = (1-\alpha)\, w_1(y, \mu, q, \alpha) + \alpha\, w_2(y, \mu, q, \alpha), \qquad y = 0, 1, 2, \ldots\ . \tag{32}$$

In the above expression,

$$\begin{aligned} w_k(y, \mu, q, \alpha) &= (1-q^k)q^{ky} \exp[-(\mu - g(q,\alpha))] \sum_{i=0}^{y} \frac{1}{i!} \left(\frac{\mu - g(q,\alpha)}{q^k}\right)^i \\ &= \frac{(1-q^k)q^{ky}}{\Gamma(y+1)} \exp\left[(\mu - g(q,\alpha))\left(\frac{1}{q^k} - 1\right)\right] \Gamma\left(y+1, \frac{\mu - g(q,\alpha)}{q^k}\right),\ k = 1, 2\,. \end{aligned}$$

Clearly, the mean is $E(Y) = \mu$ and the variance is

$$V(Y) = \frac{\mu(1-q^2)^2 + q^2[(1+q)^2 - 2q\alpha - \alpha^2]}{(1-q^2)^2}.$$

Let y be a random sample of n from the $PoiTG(\mu, q, \alpha)$ distribution, as defined in (32). The PoiTG_{GLM} is formulated using the log link function *i.e.*, which establishes the relationship between the p-dimensional covariates and the mean response.

$$\eta_i = g(\mu_i) = \log(\mu_i) = \sum_{j=0}^{p} \gamma_j x_{ij} = \boldsymbol{x_i'\gamma}, \quad i = 1, 2, \ldots n, \tag{33}$$

where $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2, \ldots, \gamma_p)'$ and $\boldsymbol{x_i'} = (1, x_{i1}, x_{i2}, \ldots, x_{ip})$ are the vector of unknown regression coefficients and covariates augmented with a 1 to account for the intercept. Let $Y_i|\boldsymbol{x_i'}$ be a response for a given set of covariates, $\boldsymbol{x_i'}$. By substituting $\mu = \mu_i = \exp(\boldsymbol{x_i'\gamma})$ into (32), the pmf of $Y_i|\boldsymbol{x_i'} \sim PoiTG_{GLM}(\mu, q, \alpha)$ can be expressed as

$$p_Y(y_i|\boldsymbol{x_i'}) = (1-\alpha)\, w_1(y_i, \exp(\boldsymbol{x_i'\gamma}), q, \alpha) + \alpha\, w_2(y_i, \exp(\boldsymbol{x_i'\gamma}), q, \alpha), \quad i = 1, 2, \ldots n\ . \tag{34}$$

In the above expression,

$$w_k(y_i, \exp(\boldsymbol{x_i'\gamma}), q, \alpha) = \frac{(1-q^k)q^{ky_i}}{\Gamma(y_i+1)} \exp\left[(\exp(\boldsymbol{x_i'\gamma}) - g(q,\alpha))\left(\frac{1}{q^k}-1\right)\right] \Gamma\left(y_i+1, \frac{\exp(\boldsymbol{x_i'\gamma}) - g(q,\alpha)}{q^k}\right), \ k = 1, 2\ .$$

The log-likelihood function of the PoiTG regression model is given by

$$l(\boldsymbol{\delta}) = \sum_{i=1}^{n} \log\left[(1-\alpha)\, w_1(y_i, \exp(\boldsymbol{x_i'\gamma}), q, \alpha) + \alpha\, w_2(y_i, \exp(\boldsymbol{x_i'\gamma}), q, \alpha)\right], \tag{35}$$

where $\boldsymbol{\delta} = (\boldsymbol{\gamma},\, q,\, \alpha)'$. The unknown parameters, $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2, ..., \gamma_p)'$, q and α are estimated by maximizing (34) using the *constrOptim* function in R, which efficiently implements the Nelder-Mead method. The asymptotic confidence intervals for the regression coefficients $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \gamma_2, ..., \gamma_p)'$, q and α can be derived from asymptotic theory as $n \to \infty$. Based on this theory,

$$\sqrt{n}(\hat{\boldsymbol{\delta}} - \boldsymbol{\delta}) \to N_{p+3}(0,\, I^{-1}(\hat{\boldsymbol{\delta}})) \text{ as } n \to \infty,$$

where $I(\hat{\boldsymbol{\delta}})$ is the observed information matrix of $\hat{\boldsymbol{\delta}}$.

7.1. Bootstrap standard error and confidence interval

Calculating asymptotic confidence intervals for each regression coefficient and the parameters q and α is laborious. Bootstrapping provides an alternative approach to obtain precise standard errors (SE) and confidence intervals (CI) for the estimates. To determine these for the PoiTG regression model, we adopt the following nonparametric bootstrap routine.

First, we randomly select a sample with replacement B times from the set of row indices $\{1, 2, ..., n\}$ of the given dataset. For the j^{th} repetition, let $\{i_1^{(j)}, i_2^{(j)}, ..., i_n^{(j)}\}$ denote the randomly selected indices. The j^{th} bootstrap dataset is

$$D^{(j)} = \begin{pmatrix} y_{i_1^{(j)}} & x_{1i_1^{(j)}} & \dots & x_{pi_1^{(j)}} \\ y_{i_2^{(j)}} & x_{1i_2^{(j)}} & \dots & x_{pi_2^{(j)}} \\ \vdots & \vdots & & \vdots \\ y_{i_n^{(j)}} & x_{1i_n^{(j)}} & \dots & x_{pi_n^{(j)}} \end{pmatrix}, \ j = 1, 2, ..., B.$$

Based on $D^{(j)}$, we compute the MLEs of $\gamma_1, \gamma_2, ..., \gamma_p$ as $\hat{\gamma_1}^{(j)}, \hat{\gamma_2}^{(j)}, ..., \hat{\gamma_p}^{(j)}$ for each $j = 1, 2, ..., B$. Consequently, for any given $k \in 0, 1, ..., p$, we obtain a set of MLEs of γ_k, represented as $\boldsymbol{\gamma_k^*} = (\hat{\gamma_k}^{(1)}, \hat{\gamma_k}^{(2)}, ..., \hat{\gamma_k}^{(B)})$.

To obtain the ordinary bootstrap estimate of the SE of $\hat{\gamma_k}$, we compute the standard deviation of the array $\boldsymbol{\gamma_k^*}$ as

$$SE(\hat{\gamma_k}) = \sqrt{\frac{1}{B}\sum_{j=1}^{B}(\gamma_k^{(j)} - \bar{\gamma_k})^2},$$

where, $\bar{\gamma_k}$ is the mean of the bootstrap estimates. To construct $100(1-\nu)\%$ bootstrap CI for γ_k, first we obtain the empirical cumulative distribution function:

$$\hat{F_B}(t) = \frac{1}{B}\sum_{j=1}^{B} I(\sqrt{n}(\hat{\gamma}_k^{(j)} - \hat{\gamma_k}) \leq t).$$

Let, $t_{\nu/2} = \hat{F_B^{-1}}(\nu/2)$ and $t_{1-\nu/2} = \hat{F_B^{-1}}(1-\nu/2)$. Using these threshold points, the $100(1-\nu)\%$ bootstrap CI for γ_k is given by

$$CI(\hat{\gamma_k}) = \left[\hat{\gamma_k} - \frac{t_{1-\nu/2}}{\sqrt{n}},\, \hat{\gamma_k} + \frac{t_{\nu/2}}{\sqrt{n}}\right].$$

Typically, B is chosen to be large, on the order of thousands. A similar procedure can be applied to estimate the SEs and CIs for q and α. The corresponding formulas for q and α are as follows:

$$SE(\hat{q}) = \sqrt{\frac{1}{B}\sum_{j=1}^{B}(q^{(j)} - \bar{q})^2}\ ,\ SE(\hat{\alpha}) = \sqrt{\frac{1}{B}\sum_{j=1}^{B}(\alpha^{(j)} - \bar{\alpha})^2},$$

$$CI(\hat{q}) = \left[\hat{q} - \frac{t_{1-\nu/2}}{\sqrt{n}},\, \hat{q} + \frac{t_{\nu/2}}{\sqrt{n}}\right] \text{ and } CI(\hat{\alpha}) = \left[\hat{\alpha} - \frac{t_{1-\nu/2}}{\sqrt{n}},\, \hat{\alpha} + \frac{t_{\nu/2}}{\sqrt{n}}\right].$$

7.2. Applied regression fitting

In this section, we evaluate the practicality of the PoiTG regression model (PoiTG_{GLM}) using a real-world dataset. DeHart, *et al.* (2008) conducted a study in which individuals classified as "moderate to heavy drinkers" (defined as consuming at least 12 alcoholic drinks per week for women and 15 for men) were asked to maintain a daily record of their alcohol intake over a 30-day period. Participants also completed various rating scales related to daily life events and self-esteem.

One of the key hypotheses explored in the study was that negative events, particularly those involving romantic relationships, could be associated with the amount of alcohol consumed, especially among people with low self-esteem. The primary focus of our analysis is to model the variable *numall*, which represents the number of alcoholic beverages (or "drinks") consumed in a single day. Furthermore, we examine how *number* is influenced by factors such as daily desire to drink (*desired*), age (*age*), negative romantic relationship events (*nevgent*), and state self-esteem (*state*). The predictive model is formulated as follows:

$$g(\mu_i) = \log(\mu_i) = \gamma_0 + \gamma_1\, negevent_i + \gamma_2\, state_i + \gamma_3\, age_i + \gamma_4\, desired_i$$

for $i = 1, ..., 618$.

We compare the fitted values of the PoiTG regression model (PoiTG_{GLM}) with those of established regression models for equidispersed count data, such as the Poisson regression model (Poi_{GLM}), and for overdispersed count data, including the PoiG regression model

Table 4: Analysis of the Alcohol consumption dataset

	Estimators(SE)				
	Poi_{GLM}	NB_{GLM}	CMP_{GLM}	PoiG_{GLM}	PoiTG_{GLM}
γ_0	-0.02541(0.3198)	-0.2559(0.4662)	-0.0486(0.2225)	0.4054(0.5121)	0.7128(0.16209)
	(-0.6522, 0.6014)	(-1.1696 ,0.6578)	(-0.4847 ,0.3875)	(-0.5983,1.4091)	(0.3951,1.0305)
γ_1	-0.2726(0.0709)	-0.2420(0.1002)	-0.1452(0.0509)	-2.0305(0.12040)	0.6884(0.0347)
	(-0.4116,-0.1336)	(-0.4384 ,-0.0456)	(-0.2450,-0.0454)	(-2.2665,-1.7945)	(0.6204,0.7564)
γ_2	-0.0514(0.0570)	-0.0600(0.0843)	-0.0670(0.0389)	-0.6384(0.0749)	-0.1148(0.0244)
	(-0.1631,0.0603)	(-0.2252,0.1052)	(-0.1432,0.0092)	(-0.7851,-0.4917)	(-0.1627,-0.0669)
γ_3	-0.0018(0.0058)	0.0039(0.0083)	-0.0058(0.0040)	-0.1126(0.0138)	0.2720(0.0029)
	(-0.0132,0.0096)	(-0.0124,0.0202)	(-0.0136,0.0020)	(-0.1396,-0.0856)	(0.2663,0.2777)
γ_4	0.2755(0.01636)	0.2871(0.0232)	0.1342(0.0139)	0.0133(0.0502)	-0.0959(0.0077)
	(0.2434,0.3076)	(0.2416,0.3326)	(0.1070,0.1614)	(-0.0851,0.1117)	(-0.1109,-0.0809)
ϕ	—	2.4720(0.3050)	-1.2003(0.1392)	0.4611(0.0190)	—
	—	(1.8742,3.0698)	(-1.4731,-0.9275)	(0.4239,0.4983)	—
q	—	—	—	—	0.0148(0.0071)
	—	—	—	—	(0.0009,0.0287)
α	—	—	—	—	0.2167(0.0334)
					(0.1512,0.2822)
$-\text{l}$	1317.750	1217.189	1223.979	1227.197	**1197.856**
AIC	2645.500	2446.378	2459.958	2466.394	**2409.712**
BIC	2667.632	2472.937	2486.517	2492.953	**2440.697**

The bootstrap SEs and CIs of the MLEs are presented in parentheses next to and below the estimates.

(PoiG_{GLM}), the COM-Poisson regression model (CMP_{GLM}), and the NB regression model (NB_{GLM}). Model selection is based on the estimated negative log-likelihood (-LL), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). Table 4 presents the estimated parameters, -LL, AIC, and BIC values for the fitted models using the Alcohol consumption dataset. The results demonstrate that the PoiTG regression model offers the highest modeling accuracy, as it attains the lowest -LL, AIC, and BIC values. The estimated parameters of the PoiTG regression model further implies that the predictive model can be written as

$$\mu_i = \exp(0.7128 + 0.6884\, negevent_i - 0.1148\, state_i + 0.2720\, age_i - 0.0959\, desired_i).$$

Figure 12 includes scatter plot, Q-Q plot and histogram of the randomized quantile residuals of PoiTG regression model fitted to Alcohol consumption dataset.

The question arises whether there is a statistically significant difference between the PoiTG regression model and its closest competitor, the NB regression model, in fitting this data set. To answer this question, we use the generalized likelihood ratio test based on LLs, whose test statistic is $\text{LL}(\text{NB}_{GLM})$-$\text{LL}(\text{PoiTG}_{GLM})$=19.333, which results in a significance level of 0.00001. To perform this test, we compare the obtained statistics with $\chi^2(1)$ (Chi-square distribution with one degree of freedom). Therefore, there is a significant difference

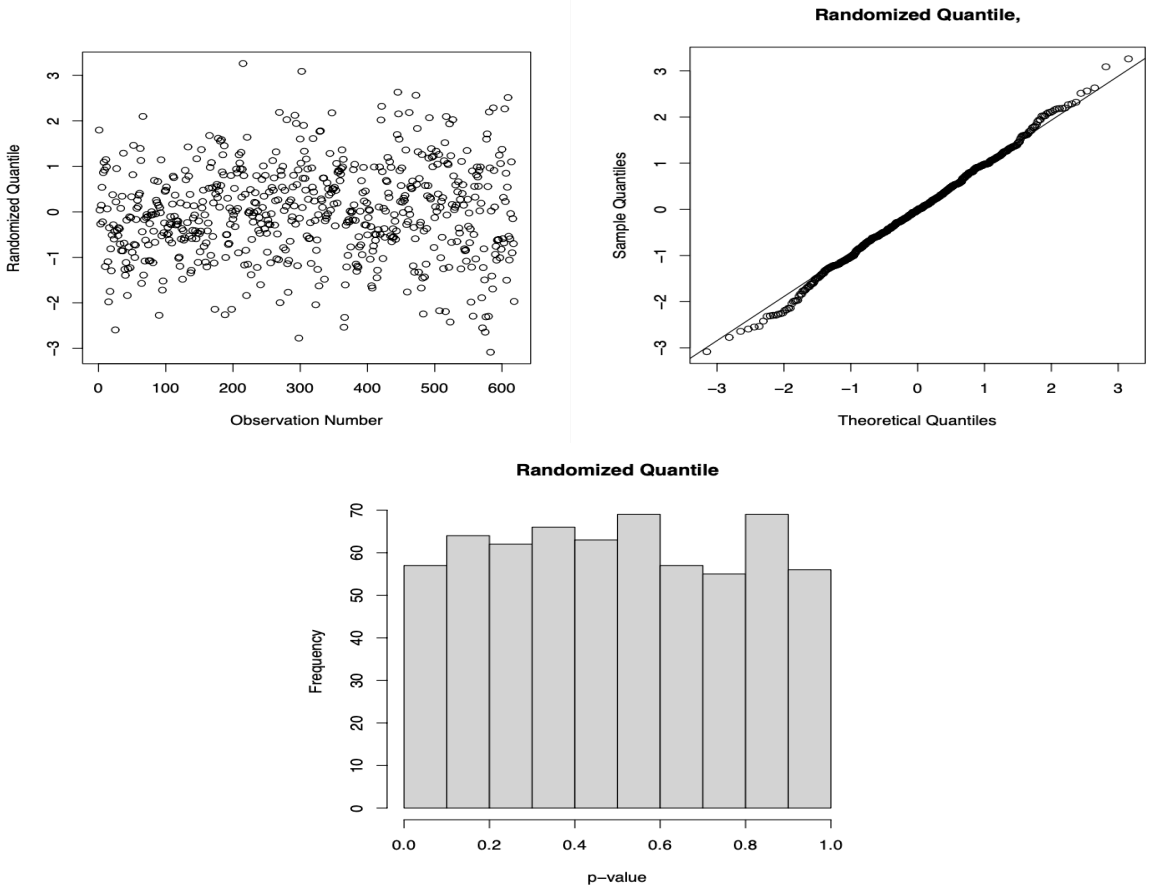

Figure 12: The randomized quantile residuals and corresponding Q-Q plot and histogram

between these two models (for more details, see Self and Liang (1987)). Thus, among the used models, the PoiTG regression model is the best model to fit this data set.

To check the efficiency of the fitted models, in addition to using the previous criteria, we also use the Pearson chi-square statistic χ^2 Tests.

$$\chi^2 = \sum_{i=1}^{c} \frac{(o_i - e_i)^2}{e_i},$$

where, o_i is observed frequency, e_i is the expected frequency of the ith category, and c is the total number of categories. To perform this test, we sorted the data from 0 to 7 and more than 7. Then, the observed and expected frequency of the data set are calculated. The value of the test statistic is 7.87 and the P-value corresponding to the test is 0.2475, which supports a good fit of the PoiTG model to the data.

8. Discussion

This article introduces a novel discrete distribution and thoroughly studies its important properties. This distribution possesses a closed form of the mean, which eventually facilitates the corresponding generalized linear model. However, the usp of the present work is the real-life applications which should motivate the practitioners to adopt it as a viable alternative to the COM-Poisson and other overdispersed count data models. The interpretability of the proposed model is also worth adoption in practice. The parameters θ, q, and α control the distribution's tail behavior, while q and α account for the overdispersion in the dataset. Their combined influence provides flexibility in shaping the distribution. In the case of large θ, the distribution maintains a bell-shaped mass distribution. In Section 3.5, we observed that the hazard rate function of the PoiTG distribution can capture near-constant, increasing, and decreasing behaviors under different parameter settings. One possible extension of this work is to suitably modify the distribution so as to capture bathtub-shaped hazard behaviors, thereby enriching its applicability to a wider range of reliability scenarios. This work is complete in its objectives. However, further research is warranted on exploring the applicability of this new distribution for count time series data. Another important direction of future research may be to study the inferential aspects of the proposed model under the Bayesian perspective.

Acknowledgements

The authors gratefully acknowledge the editor and the learned reviewers for the insightful comments and constructive suggestions, which significantly improved the clarity and presentation of the manuscript.

Conflict of interest

The authors declare that there is no conflict of interest. Additionally, no funding was received for the preparation of this manuscript.

References

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. US Government printing office.

Altun, E. (2020). A new generalization of geometric distribution with properties and applications. *Communications in Statistics-Simulation and Computation*, **49**, 793–807.

Altun, E., El-Morshedy, M., and Eliwa, M. S. (2022). A study on discrete Bilal distribution with properties and applications on integervalued autoregressive process. *REVSTAT-Statistical Journal*, **20**, 501–528.

Amani, K. M., Kouakou, K. J. G., and Hili, O. (2025). Marginalized zero-inflated bell regression models for overdispersed count data. *Journal of Statistical Theory and Practice*, **19**, 17.

Bar-Lev, S. K. and Ridder, A. (2021). Exponential dispersion models for overdispersed zero-inflated count data. *Communications in Statistics-Simulation and Computation*, **52**, 3286–3304.

Bardwell, G. E. and Crow, E. L. (1964). A two parameter family of hyper-Poisson distributions. *Journal of the American Statistical Association*, **59**, 133–141.

Bourguignon, M., Gallardo, D. I., and Medeiros, R. M. R. (2022). A simple and useful regression model for underdispersed count data based on Bernoulli–Poisson convolution. *Statistical Papers*, **63**, 821–848.

Bourguignon, M. and Weiß, C. H. (2017). An INAR (1) process for modeling count time series with equidispersion, underdispersion and overdispersion. *Test*, **26**, 847–868.

Chakraborty, S. (2010). On some distributional properties of the family of weighted generalized poisson distribution. *Communications in Statistics—Theory and Methods*, **39**, 2767–2788.

Chakraborty, S. and Bhati, D. (2016). Transmuted geometric distribution with applications in modeling and regression analysis of count data. *SORT-Statistics and Operations Research Transactions*, **40**, 153–176.

Chakraborty, S. and Gupta, R. D. (2015). Exponentiated geometric distribution: another generalization of geometric distribution. *Communications in Statistics-Theory and Methods*, **44**, 1143–1157.

Chakraborty, S. and Ong, S. H. (2017). Mittag-leffler function distribution-a new generalization of hyper-Poisson distribution. *Journal of Statistical Distributions and Applications*, **4**, 1–17.

DeHart, T., Tennen, H., Armeli, S., Todd, M., and Affleck, G. (2008). Drinking to regulate negative romantic relationship interactions: The moderating role of self-esteem. *Journal of Experimental Social Psychology*, **44**, 527–538.

Del Castillo, J. and Pérez-Casany, M. (1998). Weighted poisson distributions for overdispersion and underdispersion situations. *Annals of the Institute of Statistical Mathematics*, **50**, 567–585.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, **39**, 1–22.

Efron, B. (1986). Double exponential-families and their use in generalized linear-regression. *Journal of the American Statistical Association*, **81**, 709–721.

El-Morshedy, M., Eliwa, M. S., and Altun, E. (2020). Discrete Burr-Hatke distribution with properties, estimation methods and regression model. *IEEE Access*, **8**, 74359–74370.

Ghahramani, M. and White, S. S. (2020). Time series regression for zero-inflated and overdispersed count data: A functional response model approach. *Journal of Statistical Theory and Practice*, **14**, 29.

Gómez-Déniz, E. (2010). Another generalization of the geometric distribution. *Test*, **19**, 399–415.

Hahn, E. D. (2022). The tilted beta-binomial distribution in overdispersed data: Maximum likelihood and Bayesian estimation. *Journal of Statistical Theory and Practice*, **16**, 43.

Hand, D. J., Daly, F., Lunn, A. D., McConway, K., and Ostrowski, E. (1995). A handbook of small data sets. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **158**, 339.

Hassanzadeh, F. and Kazemi, I. (2016). Analysis of over-dispersed count data with extra zeros using the Poisson log-skew-normal distribution. *Journal of Statistical Computation and Simulation*, **86**, 2644–2662.

Hoel, P. G. (1943). On indices of dispersion. *The Annals of Mathematical Statistics*, **14**, 155–162.

Hussain, T. and Ahmad, M. (2014). Discrete inverse Rayleigh distribution. *Pakistan Journal of Statistics*, **30**, 203–2022.

Jain, G. C. and Consul, P. C. (1971). A generalized negative binomial distribution. *SIAM Journal on Applied Mathematics*, **21**, 501–513.

Jazi, M. A., Lai, C. D., and Alamatsaz, M. H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. *Statistical Methodology*, **7**, 121–132.

Krishna, H. and Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. *Statistical Methodology*, **6**, 177–188.

Makcutek, J. (2008). A generalization of the geometric distribution and its application in quantitative linguistics. *Romanian Reports in Physics*, **60**, 501–509.

Moghimbeigi, A., Eshraghian, M. R., Mohammad, K., and Mcardle, B. (2008). Multilevel zero-inflated negative binomial regression modeling for over-dispersed count data with extra zeros. *Journal of Applied Statistics*, **35**, 1193–1202.

Moqaddasi Amiri, M., Tapak, L., and Faradmal, J. (2019). A mixed-effects least square support vector regression model for three-level count data. *Journal of Statistical Computation and Simulation*, **89**, 2801–2812.

Nandi, A., Biswas, A., Hazarika, P. J., and Das, J. (2024a). A new regression model for over-dispersed count responses based on Poisson and geometric convolution. *Journal of the Indian Society for Probability and Statistics*, **Online first**, 1–18.

Nandi, A., Chakraborty, S., and Biswas, A. (2024b). A new over-dispersed count model based on poisson-geometric convolution. *Communications in Statistics - Simulation and Computation*, **Online first**, 1–26.

Nekoukhou, V., Alamatsaz, M. H., and Bidram, H. (2012). A discrete analogue of the generalized exponential distribution. *Communications in Statistics - Theory and Methods*, **41**, 2000–2013.

Para, B. A. and Jan, T. R. (2016a). Discrete version of log-logistic distribution and its applications in genetics. *International Journal of Modern Mathematical Sciences*, **14**, 407–422.

Para, B. A. and Jan, T. R. (2016b). On discrete three-parameter Burr type XII and discrete Lomax distributions and their applications to model count data from medical science. *Biometrics and Biostatistics International Journal*, **4**, 1–15.

Philippou, A. N., Georghiou, C., and Philippou, G. N. (1983). A generalized geometric distribution and some of its properties. *Statistics and Probability Letters*, **1**, 171–175.

Redner, R. A. and Walker, H. F. (1984). Mixture densities, maximum likelihood and the em algorithm. *SIAM review*, **26**, 195–239.

Rodrigues-Motta, M., Pinheiro, H. P., Martins, E. G., Araújo, M. S., and dos Reis, S. F. (2013). Multivariate models for correlated count data. *Journal of Applied Statistics*, **40**, 1586–1596.

Roy, D. (2004). Discrete rayleigh distribution. *IEEE Transactions on Reliability*, **53**, 255–260.

Sarvi, F., Moghimbeigi, A., and Mahjub, H. (2019). GEE-based zero-inflated generalized Poisson model for clustered over or under-dispersed count data. *Journal of Statistical Computation and Simulation*, **89**, 2711–2732.

Self, S. G. and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, **82**, 605–610.

Sellers, K. F. and Shmueli, G. (2010). A flexible regression model for count data. *The Annals of Applied Statistics*, **4**, 943–961.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, **27**, 379–423.

Singh, B., Singh, R. P., Nayal, A. S., and Tyagi, A. (2022). Discrete inverted Nadarajah-Haghighi distribution: Properties and classical estimation with application to complete and censored data. *Statistics, Optimization & Information Computing*, **10**, 1293–1313.

Tapak, L., Hamidi, O., Amini, P., and Verbeke, G. (2020). Random effect exponentiated-exponential geometric model for clustered/longitudinal zero-inflated count data. *Journal of Applied Statistics*, **47**, 2272–2288.

Tripathi, R. C., Gupta, R. C., and White, T. J. (1987). Some generalizations of the geometric distribution. *Sankhya Ser. B*, **49**, 218–223.

Tüzen, F., Erbaş, S., and Olmuş, H. (2020). A simulation study for count data models under varying degrees of outliers and zeros. *Communications in Statistics-Simulation and Computation*, **49**, 1078–1088.

Wang, S., Cadigan, N. G., and Benoît, H. P. (2017). Inference about regression parameters using highly stratified survey count data with over-dispersion and repeated measurements. *Journal of Applied Statistics*, **44**, 1013–1030.

Wongrin, W. and Bodhisuwan, W. (2017). Generalized Poisson–Lindley linear model for count data. *Journal of Applied Statistics*, **44**, 2659–2671.

Zhan, D. and Young, D. S. (2024). Finite mixtures of mean-parameterized Conway–Maxwell–Poisson regressions. *Journal of Statistical Theory and Practice*, **65**, 1469–1492.

Appendix I

The log-likelihood functions of the $PoiG(\theta, q)$ distribution is

$$l'(\theta, q; \boldsymbol{y}) = n\log(1-q) + n\overline{y}\log q + \frac{n\theta(1-q)}{q} + \sum_{i=0}^{n}\log\left(\frac{\Gamma\left(y_i+1, \dfrac{\theta}{q}\right)}{\Gamma(y_i+1)}\right). \tag{36}$$

The score functions are

$$s_1'(\theta, q; \boldsymbol{y}) = \frac{\partial}{\partial\theta} l'(\theta, q; \boldsymbol{y}) = \frac{n(1-q)}{q} - \sum_{i=1}^{n} \alpha_1(y_i)\beta^{y_i},$$

$$\text{and} \quad s_2'(\theta, q; \boldsymbol{y}) = \frac{\partial}{\partial q} l'(\theta, q; \boldsymbol{y}) = -\frac{n}{1-q} - \frac{n(\theta - \bar{y})}{q} - \frac{n\theta(1-q)}{q^2} + \sum_{i=1}^{n} \theta\alpha_2(y_i)\beta^{y_i}.$$

In the above expressions, $\beta = \dfrac{\theta}{q}$ and

$$\alpha_j(y_i) = \frac{e^{-\beta}}{\Gamma(y_i + 1, \beta)} \frac{1}{q^j} \text{ for } i = 1, 2, ..., n, j = 1, 2.$$

The second order partial derivatives of the log-likelihood function of the $PoiG(\theta, q)$ distribution (Nandi, *et al.* (2024b)) are

$$\frac{\partial^2 l'(\theta, q; \boldsymbol{y})}{\partial\theta^2} = \sum_{i=1}^{n} \left[(\beta^{y_i} - y_i\beta^{y_i - 1})\alpha_2(y_i) - \beta^{2y_i}\alpha_1(y_i)^2 \right],$$

$$\frac{\partial^2 l'(\theta, q; \boldsymbol{y})}{\partial\theta\partial q} = -\frac{n}{q^2} - \sum_{i=1}^{n} \left[\theta(\beta^{y_i} - y_i\beta^{y_i - 1})\alpha_3(y_i) - \beta^{y_i}\alpha_2(y_i) - \frac{\theta}{q}\beta^{2y_i}\alpha_1(y_i)^2 \right],$$

$$\text{and} \quad \frac{\partial^2 l'(\theta, q; \boldsymbol{y})}{\partial q^2} = \frac{2n\theta - n\bar{y}q}{q^3} - \frac{n}{(1-q)^2} + \sum_{i=1}^{n} \left[((\theta^2 - 2\theta q)\beta^{y_i} - \theta^2 y_i\beta^{y_i - 1})\alpha_4(y_i) - \theta^2\beta^{2y_i}\alpha_2(y_i)^2 \right].$$

The above listed second order partial derivatives of the log-likelihood functions for the PoiG distribution (Nandi, *et al.*, 2024b) can be utilized to derive the expression for the 2^{nd} partial derivatives of the log-likelihood function for PoiTG distribution,

$$l(\theta, q, \alpha; \boldsymbol{y}) = \sum_{i=1}^{n} \log L(\theta, q, \alpha; y_i) = \sum_{i=1}^{n} \log((1-\alpha)P(\theta, q; y_i) + \alpha P(\theta, q_*; y_i)).$$

Here, we use q^* to denote q^2. For each $i = 1, 2, ..., n$, let us introduce the following notations for further use.

$$A_{i1} = (1-\alpha)L'(\theta, q; y_i)\frac{\partial l'(\theta, q; y_i)}{\partial\theta} + \alpha L'(\theta, q_*; y_i)\frac{\partial l'(\theta, q_*; y_i)}{\partial\theta},$$

$$A_{i2} = (1-\alpha)L'(\theta, q; y_i)\left(\left(\frac{\partial l'(\theta, q; y_i)}{\partial\theta}\right)^2 + \frac{\partial^2 l'(\theta, q; y_i)}{\partial\theta^2}\right),$$

$$A_{i3} = \alpha L'(\theta, q_*; y_i)\left(\left(\frac{\partial l'(\theta, q_*; y_i)}{\partial\theta}\right)^2 + \frac{\partial^2 l'(\theta, q_*; y_i)}{\partial\theta^2}\right),$$

$$B_{i1} = (1-\alpha)L'(\theta, q; y_i)\frac{\partial l'(\theta, q; y_i)}{\partial q} + 2\alpha q L'(\theta, q_*; y_i)\frac{\partial l'(\theta, q_*; y_i)}{\partial q^*},$$

$$B_{i2} = (1-\alpha)L'(\theta, q; y_i)\left(\left(\frac{\partial l'(\theta, q; y_i)}{\partial q}\right)^2 + \frac{\partial^2 l'(\theta, q; y_i)}{\partial q^2}\right),$$

$$B_{i3} = 2\alpha L'(\theta, q_*; y_i)\left(\frac{\partial l'(\theta, q_*; y_i)}{\partial q_*} + 2q^2\left(\frac{\partial l'(\theta, q_*; y_i)}{\partial q_*}\right)^2 + 2q^2\frac{\partial^2 l'(\theta, q_*; y_i)}{\partial q_*^2}\right),$$

$$C_{i1} = (1-\alpha)L'(\theta, q; y_i)\left(\frac{\partial l'(\theta, q; y_i}{\partial \theta}\frac{\partial l'(\theta, q; y_i)}{\partial q} + \frac{\partial^2 l'(\theta, q; y_i)}{\partial \theta \partial q}\right),$$

$$C_{i2} = 2\alpha q L'(\theta, q_*; y_i)\left(\frac{\partial l'(\theta, q_*; y_i)}{\partial \theta}\frac{\partial l'(\theta, q_*; y_i)}{\partial q_*} + \frac{\partial^2 l'(\theta, q_*; y_i)}{\partial \theta \partial q_*}\right),$$

and $$C_{i3} = 2qL'(\theta, q_*; y_i)\frac{\partial l'(\theta, q_*; y_i)}{\partial q_*} - L'(\theta, q; y_i)\frac{\partial l'(\theta, q; y_i)}{\partial q}.$$

The second order partial derivatives of the log-likelihood function for $PoiTG(\theta, q, \alpha)$ are

$$\frac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \theta^2} = \sum_{i=1}^{n}\frac{L(\theta, q, \alpha; y_i)(A_{i2} + A_{i3}) - A_{i1}^2}{L(\theta, q, \alpha; y_i)^2}, \tag{37}$$

$$\frac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial q^2} = \sum_{i=1}^{n}\frac{L(\theta, q, \alpha; y_i)(B_{i2} + B_{i3}) - B_{i1}^2}{L(\theta, q, \alpha; y_i)^2}, \tag{38}$$

$$\frac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \theta \partial q} = \sum_{i=1}^{n}\frac{L(\theta, q, \alpha y_i)(C_{i1} + C_{i2}) - A_{i1}B_{i1}}{L(\theta, q, \alpha; y_i)^2}, \tag{39}$$

$$\frac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \alpha^2} = -\sum_{i=1}^{n}\frac{\left(L'(\theta, q_*; y_i) - L'(\theta, q; y_i)\right)^2}{L(\theta, q, \alpha;)^2}, \tag{40}$$

$$\frac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \theta \partial \alpha} = \sum_{i=1}^{n}\frac{\left(L'(\theta, q_*; y_i)\dfrac{\partial l'(\theta, q_*; y_i)}{\partial \theta} - L'(\theta, q; y_i)\dfrac{\partial l'(\theta, q; y_i)}{\partial \theta}\right)}{L(\theta, q, \alpha; y_i)^2}, \tag{41}$$

and $$\frac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial q \partial \alpha} = \sum_{i=1}^{n}\frac{L(\theta, q, \alpha; y_i)C_{i3}}{L(\theta, q, \alpha; y_i)^2} - \sum_{i=1}^{n}\frac{\left(L'(\theta, q_*; y_i) - L'(\theta, q; y_i)\right)B_{i1}}{L(\theta, q, \alpha; y_i)^2}. \tag{42}$$

The resulting fisher's information matrix for θ, q, and α is

$$I_Y(\theta, q, \alpha) = \begin{pmatrix} -E\left(\dfrac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \theta^2}\right) & -E\left(\dfrac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \theta \partial q}\right) & -E\left(\dfrac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \theta \partial \alpha}\right) \\ -E\left(\dfrac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \theta \partial q}\right) & -E\left(\dfrac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial q^2}\right) & -E\left(\dfrac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial q \partial \alpha}\right) \\ -E\left(\dfrac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \theta \partial \alpha}\right) & -E\left(\dfrac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial q \partial \alpha}\right) & -E\left(\dfrac{\partial^2 l(\theta, q, \alpha; \boldsymbol{y})}{\partial \alpha^2}\right) \end{pmatrix}.$$

This can be approximated by

$$\widehat{I}_Y(\theta,q,\alpha) \approx \begin{pmatrix} -\frac{\partial^2 l(\theta,q,\alpha;\boldsymbol{y})}{\partial\theta^2} & -\frac{\partial^2 l(\theta,q,\alpha;\boldsymbol{y})}{\partial\theta\partial q} & -\frac{\partial^2 l(\theta,q,\alpha;\boldsymbol{y})}{\partial\theta\partial\alpha} \\ -\frac{\partial^2 l(\theta,q,\alpha;\boldsymbol{y})}{\partial\theta\partial q} & -\frac{\partial^2 l(\theta,q,\alpha;\boldsymbol{y})}{\partial q^2} & -\frac{\partial^2 l(\theta,q,\alpha;\boldsymbol{y})}{\partial q\partial\alpha} \\ -\frac{\partial^2 l(\theta,q,\alpha;\boldsymbol{y})}{\partial\theta\partial\alpha} & -\frac{\partial^2 l(\theta,q,\alpha;\boldsymbol{y})}{\partial q\partial\alpha} & -\frac{\partial^2 l(\theta,q,\alpha;\boldsymbol{y})}{\partial\alpha^2} \end{pmatrix}_{(\theta,q,\alpha)=(\hat{\theta}_{ML},\hat{q}_{ML},\hat{\alpha}_{ML})}.$$

The MLEs $\hat{\theta}_{ML}$, $\hat{q}_{ML}$, and $\hat{\alpha}_{ML}$ are consistent and asymptotically normal with the mean vector (θ, q, α) and the dispersion matrix $\hat{I}_Y^{-1} = [d_{ij}]_{3\times 3}$. Let z_ν denote the $(1-\nu)$-th quantile of the standard normal distribution. The $(1-\nu)\times 100\%$ confidence interval for the parameters θ, q and α are given by

$$\left(\hat{\theta}_{ML} - z_{\nu/2}\sqrt{d_{11}}\ ,\ \hat{\theta}_{ML} + z_{\nu/2}\sqrt{d_{11}}\right), \left(\hat{q}_{ML} - z_{\nu/2}\sqrt{d_{22}}\ ,\ \hat{q}_{ML} + z_{\nu/2}\sqrt{d_{22}}\right)$$

and

$$\left(\hat{\alpha}_{ML} - z_{\nu/2}\sqrt{d_{33}}\ ,\ \hat{\alpha}_{ML} + z_{\nu/2}\sqrt{d_{33}}\right), \text{ respectively.}$$

Table 5: Simulated bias and mean squared error of the estimators

Bias (Unshaded region) and MSE (Shaded region)									
θ	q	α	n	$\hat{\theta}_{ML}$	$\hat{q}_{ML}$	$\hat{\alpha}_{ML}$	$\hat{\theta}_{EM}$	$\hat{q}_{EM}$	$\hat{\alpha}_{EM}$
0.5	0.2	0.2	30	-0.1791	0.1034	-0.0222	-0.1141	0.1105	0.2673
				0.0780	0.0240	0.0726	0.0611	0.0270	0.0737
			50	-0.1324	0.0847	0.0021	-0.0705	0.0881	0.2753
				0.0529	0.0191	0.0893	0.0385	0.0202	0.0772
			100	-0.0808	0.0641	0.0357	-0.0258	0.0624	0.2836
				0.0296	0.0139	0.1092	0.0204	0.0130	0.0808
			250	-0.0359	0.0465	0.0798	0.0103	0.0376	0.2886
				0.0140	0.0103	0.1353	0.0101	0.0077	0.0834
			500	-0.0143	0.0370	0.1037	0.0273	0.0247	0.2905
				0.0085	0.0082	0.1496	0.0069	0.0052	0.0845
			1000	-0.0003	0.0327	0.1266	0.0376	0.0175	0.2916
				0.0059	0.0071	0.1611	0.0054	0.0037	0.0851
0.5	0.2	0.8	30	-0.2459	0.0732	-0.6155	-0.1894	0.0789	-0.3301
				0.0955	0.0176	0.4509	0.0733	0.0198	0.1107
			50	-0.2033	0.0524	-0.6031	-0.1480	0.0548	-0.3232
				0.0697	0.0139	0.4501	0.0487	0.0145	0.1055
			100	-0.1618	0.0350	-0.5610	-0.1131	0.0308	-0.3155

				0.0444	0.0101	0.4265	0.0279	0.0089	0.0999
			250	-0.1082	0.0110	-0.5106	-0.0695	-0.0013	-0.3095
				0.0211	0.0076	0.3947	0.0120	0.0052	0.0959
			500	-0.0828	0.0005	-0.4653	-0.0501	-0.0198	-0.3073
				0.0129	0.0069	0.3675	0.0070	0.0043	0.0945
			1000	-0.0609	-0.0096	-0.4260	-0.0331	-0.0359	-0.3057
				0.0078	0.0064	0.3425	0.0040	0.0044	0.0935
0.5	0.8	0.2	30	0.1161	-0.0112	0.0396	0.1821	0.0055	0.2266
				0.2452	0.0033	0.0735	0.2834	0.0026	0.0616
			50	0.0583	-0.0025	0.0524	0.1268	0.0133	0.232
				0.1462	0.0018	0.0796	0.1685	0.0013	0.0612
			100	0.0134	0.0033	0.0628	0.0863	0.0187	0.2424
				0.0748	0.0012	0.0823	0.0845	0.0004	0.0616
			250	-0.0075	0.0050	0.0513	0.0678	0.0219	0.2479
				0.0313	0.0008	0.0721	0.0346	0.0007	0.0622
			500	-0.0070	0.0049	0.0469	0.0648	0.0226	0.2496
				0.0175	0.0006	0.0609	0.0194	0.0006	0.0624
			1000	-0.0104	0.0032	0.0257	0.0584	0.0231	0.2495
				0.0095	0.0004	0.0411	0.0108	0.0006	0.0623
0.5	0.8	0.8	30	0.0894	-0.0992	-0.5173	0.1509	-0.0840	-0.3562
				0.1982	0.0187	0.3517	0.2322	0.0148	0.1386
			50	0.0374	-0.0809	-0.4789	0.0972	-0.0689	-0.3494
				0.1162	0.0114	0.3228	0.1314	0.0081	0.1302
			100	0.0045	-0.0655	-0.4264	0.0619	-0.0592	-0.3414
				0.0569	0.0072	0.2816	0.0614	0.0048	0.1199
			250	0.0014	-0.0516	-0.3507	0.0499	-0.0548	-0.3389
				0.0250	0.0048	0.2184	0.0258	0.0035	0.1156
			500	0.0009	-0.0435	-0.2985	0.0417	-0.0534	-0.3395
				0.0124	0.0036	0.1733	0.0127	0.0031	0.1155
			1000	0.0081	-0.0358	-0.2422	0.0404	-0.0530	-0.3403
				0.0068	0.0027	0.1255	0.0072	0.0029	0.1159
1.5	0.2	0.2	30	-0.3463	0.1663	-0.0455	-0.2485	0.1748	0.2729
				0.2543	0.0456	0.0686	0.1904	0.0506	0.0772
			50	-0.2760	0.1476	-0.0160	-0.1866	0.1532	0.2824
				0.1621	0.0366	0.0840	0.1144	0.0394	0.0809
			100	-0.1969	0.1227	0.0231	-0.1193	0.1220	0.2891
				0.0901	0.0278	0.1038	0.0589	0.0274	0.0838
			250	-0.1208	0.0932	0.0590	-0.0554	0.0851	0.2929
				0.0412	0.0191	0.1263	0.0252	0.0161	0.0858
			500	-0.0809	0.0762	0.0854	-0.0223	0.0626	0.2944
				0.0233	0.0149	0.1410	0.0138	0.0108	0.0867

			1000	-0.0459	0.0605	0.1089	0.0080	0.0412	0.2955
				0.0134	0.0118	0.1547	0.0087	0.0072	0.0873
1.5	0.2	0.8	30	-0.4406	0.1534	-0.6475	-0.3446	0.1611	-0.3263
				0.3136	0.0407	0.4862	0.2331	0.0450	0.1090
			50	-0.3688	0.1326	-0.6151	-0.2831	0.1368	-0.3171
				0.2141	0.0322	0.4623	0.1518	0.0343	0.1015
			100	-0.2953	0.1057	-0.5853	-0.2207	0.1057	-0.3105
				0.1335	0.0235	0.4427	0.0890	0.0232	0.0966
			250	-0.2164	0.0748	-0.5409	-0.1544	0.0651	-0.3066
				0.0701	0.0157	0.4167	0.0430	0.0127	0.0940
			500	-0.1711	0.0554	-0.5070	-0.1169	0.0386	-0.3050
				0.0438	0.0121	0.3942	0.0253	0.0080	0.0930
			1000	-0.1326	0.0380	-0.4703	-0.0846	0.0128	-0.3038
				0.0268	0.0097	0.3699	0.0143	0.0052	0.0923
1.5	0.8	0.2	30	0.1059	-0.0143	0.0282	0.1997	0.0014	0.2211
				0.4456	0.0044	0.0739	0.4877	0.0035	0.0629
			50	0.0477	-0.0044	0.0398	0.1439	0.0108	0.2279
				0.2499	0.0021	0.0786	0.2696	0.0016	0.0622
			100	0.0121	0.0029	0.0584	0.1106	0.0173	0.2415
				0.1283	0.0013	0.0847	0.1361	0.0009	0.0633
			250	-0.0128	0.0050	0.0500	0.0888	0.0210	0.2511
				0.0529	0.0008	0.0748	0.0559	0.0007	0.0642
			500	-0.0111	0.0046	0.0423	0.0877	0.0217	0.2535
				0.0293	0.0006	0.0623	0.0319	0.0006	0.0645
			1000	-0.0129	0.0034	0.0276	0.0821	0.0224	0.2541
				0.0163	0.0004	0.0453	0.0187	0.0006	0.0646
1.5	0.8	0.8	30	0.0615	-0.1017	-0.5276	0.1475	-0.0875	-0.3514
				0.3552	0.0210	0.3636	0.3868	0.0175	0.1350
			50	0.0292	-0.0844	-0.4853	0.1106	-0.0731	-0.3420
				0.2117	0.0136	0.3322	0.2293	0.0103	0.1255
			100	0.0009	-0.0664	-0.4291	0.0763	-0.0607	-0.3333
				0.1049	0.0077	0.2878	0.1086	0.0054	0.1147
			250	-0.0105	-0.0515	-0.3545	0.0537	-0.0547	-0.3309
				0.0426	0.0049	0.2269	0.0418	0.0036	0.1105
			500	-0.0067	-0.0420	-0.2935	0.0450	-0.0527	-0.3315
				0.0229	0.0035	0.1736	0.0215	0.0031	0.1103
			1000	-0.0005	-0.0340	-0.2342	0.0398	-0.0521	-0.3331
				0.0117	0.0025	0.1242	0.0111	0.0028	0.1111