A New Mixed Poisson Distribution for Over-dispersed Count Data: Theory and Applications

Ramajeyam Tharshan

•

Postgraduate Institute of Science, University of Peradeniya, Peradeniya, Sri Lanka
Department of Mathematics and Statistics, University of Jaffna, Jaffna, Sri Lanka
tharshan10684@gmail.com

Pushpakanthie Wijekoon

•

Department of Statistics and Computer Science, University of Peradeniya, Peradeniya,Sri Lanka
pushpaw@pdn.ac.lk

Abstract

In this paper, an alternative mixed Poisson distribution is proposed by amalgamating Poisson distribution and a modification of the Quasi Lindley distribution. Some fundamental structural properties of the new distribution, namely the shape of the distribution and moments and related measures, are explored. It was noted that the new distribution to be either unimodal or bimodal, and over-dispersed. Further, it has a tendency to accommodate various right tail behaviors and variance-to-mean ratios. Its unknown parameter estimation by using the maximum likelihood estimation method is examined by a simulation study based on the asymptotic theory. Finally, two real-world data sets are used to illustrate the flexibility and potentiality of the new distribution.

Keywords: over-dispersion, mixed Poisson distribution, Lindley distribution, Quasi Lindley distribution, goodness of fit.

1. Introduction

Most of the real-world applications, especially, reliability, actuarial, biomedical, engineering, ecological sciences, and among others, the variable of interest is in the form of count data. The Poisson distribution is a standard tool to model the count data if the empirical and theoretical properties satisfy the related underline assumptions. A random variable X is said to have a Poisson distribution with parameter λ if both the $E(X)$ and $Var(X)$ of the distribution equal to the parameter λ. This property is commonly known as equidispersion. Even though its probability mass function (pmf) is very flexible to compute its probabilities, in some real-world applications the Poisson distribution fails to match empirical observations. Here the variance of the observed data exceeds the theoretical variance. This phenomenon is explained as over-dispersion or variation inflation (Greenwood and Yule, 1920). The over-dispersion occurs by the failure of the basic assumptions of the Poisson distribution. The reasons might be by phenomena of the clustered structure of the population or population is heterogeneous, and heavy right tail that cannot accommodate by the Poisson distribution (McCullagh et al., 1989; Ridout et al., 1998). The heterogeneity of a population is determined by the Poisson parameter λ which differs individual to individual, and then the $Var(X) = \tau E(X); (\tau > 1)$, where τ is called the index of dispersion.

The mixed Poisson distributions are well-known flexible modeling methods to explain the

heterogeneity of the Poisson parameter as well as heavy right tail behaviors (Feller,1943; Shaked, 1980). The mixed Poisson distribution is a resultant distribution or unconditional distribution by assuming that the Poisson parameter is a random variable that has as a parameterized distribution P. The distribution P and its parameter vector Θ are called prior distribution and hyperparameter, respectively. Then, the resultant distribution of the random variable X can be expressed mathematically in the following form

$$f_X(x) = \int_0^\infty f_{X|\Lambda}(x|\lambda)f_\Lambda^P(\lambda)d\lambda, \tag{1}$$

where $X|\Lambda$ has a Poisson distribution with parameter λ as

$$f_{X|\Lambda}(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \ x = 0,1,2,..., \ \lambda > 0, \tag{2}$$

Λ is the random variable of the Poisson parameter λ, and $f_\Lambda^P(\lambda)$ is the density function of the assumed continuous distribution P to the Poisson parameter λ. Hence, the random variable X has the same support of $X|\Lambda$ with parameter(s) of the prior distribution. Further, Lynch (1988) showed that the form of the mixing distribution has ascendancy over to the form of the resultant mixed distribution.

In literature, researchers assumed the standard lifetime distributions to model the Poisson parameter λ as a classical approach. Greenwood and Yule (1920) used the gamma distribution, and the resultant distribution is negative binomial (NBD). Johnson et al., (1992) assumed the exponential distribution to model the Poisson parameter, and the resultant distribution is geometric distribution (GD). Even though NBD and GD are computationally flexible pmfs, they are not befitting distributions for a higher value of τ and long right tail. In this context, researchers have assumed several modifications of the standard lifetime distributions for the Poisson parameter. They were modified to have more flexibility in their shapes and failure rate criteria than the standard lifetime distributions. Confluent Hypergeometric series, Gamma product ratio, Generalized gamma, Shifted gamma, Inverse gamma, and Modified Bessel of the 3rd kind are used by Bhattacharya (1967), Irwin (1975), Albrecht (1984), Ruohonen (1988), Willmot (1993), and Ong and Muthaloo (1995), respectively to model the Poisson parameter. The pmfs of such distributions are derived through the recursive formulas or Laplace transform technique, or by using the special mathematical functions. Hence, computing the probabilities of such distributions is complicated and they are limited in practice.

The Lindley distribution (LD) is one of the life time distributions introduced by Lindley (1958) having the density function

$$f_\Lambda(\lambda) = \frac{\theta^2}{1+\theta}(1+\lambda)e^{-\theta\lambda}, \quad \lambda > 0, \ \theta > 0, \tag{3}$$

where θ is the shape parameter, and Λ is the respective random variable. Equation (3) presents a two-component mixture of two different continuous distributions namely exponential (θ) and gamma $(2,\theta)$ distributions with the mixing proportion, $p = \dfrac{\theta}{1+\theta}$. Sankaran (1970) introduced the one-parameter discrete Poisson-Lindley distribution (PLD) by combining the Poisson and LD. Its pmf is given as

$$f_X(x) = \frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}}, \ x = 0,1,2,..., \ \theta > 0. \tag{4}$$

Note that the pmf of the PLD is an explicit form. Then, obtaining its probabilities is computationally flexible. However, the PLD flexibility is limited to fit various types of the over-dispersed count data sets since it has only one parameter. Then, as an alternative to PLD, Bhati et al. (2015) have obtained the Generalized Poisson-Lindley distribution (GPLD), where the Poisson parameter is distributed to Two-parameter Lindley distribution (Shanker et al., 2013b); Wongrin

and Bodhisuwan (2016) introduced Poisson-generalized Lindley distribution (PGLD), which was obtained by mixing the Poisson distribution with the generalized Lindley distribution (Elbatal et al., 2013); Grine et al. (2017) have obtained the Quasi-Poisson distribution (PQLD) by modeling Poisson parameter to Quasi Lindley distribution (Shanker et al., 2013a). Table 1 summarizes the mixing proportions, mixing components, and parameters of the above continuous distributions that have been used to model the Poisson parameter. We can see that the mixing proportions of the Two-parameter Lindley and generalized Lindley distributions are incorporated with the scale parameter, θ of the mixing components. Further, the shape parameter of the mixing component gamma $(2, \theta)$ is fixed with value 2 for the Two-parameter Lindley and the Quasi Lindley distributions. These settings of such mixing distributions may limit the flexibility of the above-mentioned Poisson mixtures to fit well for the various types of the right tail heaviness and τ for an over-dispersed count data (Tharshan and Wijekoon, 2020 a b).

Table 1: *Mixing proportions, mixing components, and parameters of some modified-Lindley distributions.*

Distribution	Mixing proportion	Mixing components	Parameters	
			shape	scale
Two-parameter Lindley	$\dfrac{\theta}{\theta + \alpha}$	exponential (θ) , gamma $(2, \theta)$	θ, α	
Generalized Lindley	$\dfrac{\theta}{1 + \theta}$	gamma (α, θ) , gamma (β, θ)	θ, α, β	
Quasi Lindley	$\dfrac{\alpha}{\alpha + 1}$	exponential (θ) , gamma $(2, \theta)$	θ	α

The main contribution of this paper is to propose an alternative mixed Poisson distribution for over-dispersed count data to address the above issues. It is obtained by mixing the Poisson distribution and the Modification of the Quasi Lindley distribution (MQLD) (Tharshan and Wijekoon, 2021). The density function of the MQLD(θ, α, δ) is given as

$$f_\Lambda(\lambda; \theta, \alpha, \delta) = \frac{\theta e^{-\theta \lambda}}{(\alpha^3 + 1)\Gamma(\delta)} \left(\Gamma(\delta)\alpha^3 + (\theta\lambda)^{\delta-1} \right); \lambda > 0, \theta > 0, \alpha^3 > -1, \delta > 0, \qquad (5)$$

where α, and δ are shape parameters, θ is a scale parameter, and Λ is the respective random variable. Equation (5) presents the mixture of two non-identical distributions, exponential (θ) , and gamma (δ, θ) with the mixing proportion, $p = \dfrac{\alpha^3}{\alpha^3 + 1}$. We can clearly observe that its mixing proportion p does not incorporate with scale parameter, θ of the mixing components. Further, the shape parameter of the mixing component gamma distribution, δ is not fixed with a value. The authors have shown that these settings of the MQLD provide the capability to capture the various ranges of right tail heaviness measured by excess kurtosis (kurtosis -3), horizontal symmetry measured by skewness, and heterogeneity measured by Fano factor (variance-to-mean ratio) by setting its parameter values. Further, its density function can be either unimodal or bimodal.

The remaining part of this paper is organized as follows: In section 2, we introduce the PMQLD with its explicit forms of the probability mass and distribution functions. Its fundamental structural properties are discussed in section 3. The simulation of its random variables and parameter estimations are discussed in section 4. Finally, a simulation study is done to examine the performance of parameter estimation by using the maximum likelihood estimation method, and some real-world examples are taken to show the applicability of the proposed model by comparing it with some other existing Poisson mixtures, NBD, GD, PLD, GPLD, PQLD, PGLD.

2. Formulation of the new mixed Poisson distribution

In this section, we introduce the new mixed Poisson distribution with its pmf and cumulative distribution function (cdf).

Let the random variable X represent the total counts of a specific experiment with mean λ. Then, the traditional distribution to calculate probabilities of such outcomes is the Poisson distribution. The PMQLD is obtained by mixing the Poisson and MQLD (Tharshan and Wijekoon, 2021) for over-dispersed count data. The following theorem gives the pmf of the PMQLD.

Theorem 1. Let $X|\Lambda$ is a random variable that follow the Poisson distribution with parameter λ, abbreviated as $X|\Lambda \sim$ Poisson (λ) and the Poisson parameter $\Lambda \sim$ MQLD (θ, α, δ). Then, the pmf of the PMQLD is defined as

$$f_X(x) = \frac{\theta\left(\Gamma(\delta)\Gamma(x+1)\alpha^3(1+\theta)^{\delta-1} + \theta^{\delta-1}\Gamma(x+\delta)\right)}{x!(\alpha^3+1)(1+\theta)^{x+\delta}\Gamma(\delta)}, x = 0,1,2,...,\theta > 0, \delta > 0, \alpha^3 > -1. \quad (6)$$

Proof. Since $X|\Lambda \sim$ Poisson (λ) and $\Lambda \sim$ MQLD (θ, α, δ), the unconditional distribution of X can be obtained by substituting equations (2) and (5) in equation (1) as below

$$f(X) = \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!} \frac{\theta e^{-\theta\lambda}}{(\alpha^3+1)\Gamma(\delta)}\left(\Gamma(\delta)\alpha^3 + (\theta\lambda)^{\delta-1}\right)d\lambda$$

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

$$= \frac{\theta}{x!(\alpha^3+1)\Gamma(\delta)}\left(\frac{\Gamma(\delta)\alpha^3\Gamma(x+1)}{(1+\theta)^{x+1}} + \frac{\theta^{\delta-1}\Gamma(x+\delta)}{(1+\theta)^{x+\delta}}\right)$$

$$= \frac{\theta}{x!(\alpha^3+1)\Gamma(\delta)(1+\theta)^{x+\delta}}\left((1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x+1) + \theta^{\delta-1}\Gamma(x+\delta)\right).$$

∎

Remarks:

1. Equation (6) presents a two-component mixture of GD$(\frac{\theta}{1+\theta})$ and NBD$(\delta, \frac{1}{1+\theta})$ with the mixing proportion $p = \frac{\alpha^3}{\alpha^3+1}$.

2. For $\alpha \to 0$, the PMQLD reduces to the NBD$(\delta, \frac{1}{1+\theta})$.

3. For $\alpha \to \infty$, the PMQLD reduces to the GD$(\frac{\theta}{1+\theta})$.

The right tail behaviors of the PMQLD for different values of θ, α, and δ are illustrated in Figure 1. For fixed α and δ, it is clear that the distribution's right tail approaches to zero at a faster rate when θ increases. For fixed θ and δ, and when α is increasing, the distribution's right tail approaches to zero at a slower rate when compared with the changes of θ. Further, for fixed θ and α, and when δ is increasing, the distribution captures more right tail. From Figure 2, we may note that the PMQLD may be a bimodal distribution when parameter value δ is very different (higher value) from the parameter values of θ and α.

The corresponding cumulative distribution function of PMQLD is given as

$$F_X(x) = \sum_{t=0}^x f(x)$$

$$= \frac{\delta(1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x+1)((1+\theta)^{x+1}-1)+\theta^\delta\Gamma(x+\delta+1)_2F_1(1,x+\delta+1;\delta+1;\frac{\theta}{1+\theta})}{(\alpha^3+1)\Gamma(\delta)x!\delta(1+\theta)^{x+\delta+1}} \quad (7)$$

$$, x = 0,1,2,...,\theta > 0, \delta > 0, \alpha^3 > -1,$$

where $_2F_1(c,d;r;w)$ is the Gaussian hypergeometric function defined as

$$_2F_1(c,d;r;w) = \sum_{i=0}^{\infty} \frac{(c)_i(d)_i w^i}{(r)_i i!},$$

which is a special case of the generalized hypergeometric function given by the expression

$$_aF_b(p_1,p_2,...p_a;q_1,q_2,...q_b;w) = \sum_{i=0}^{\infty} \frac{(p_1)_i...(p_a)_i w^i}{(q_1)_i...(q_b)_i i!},$$

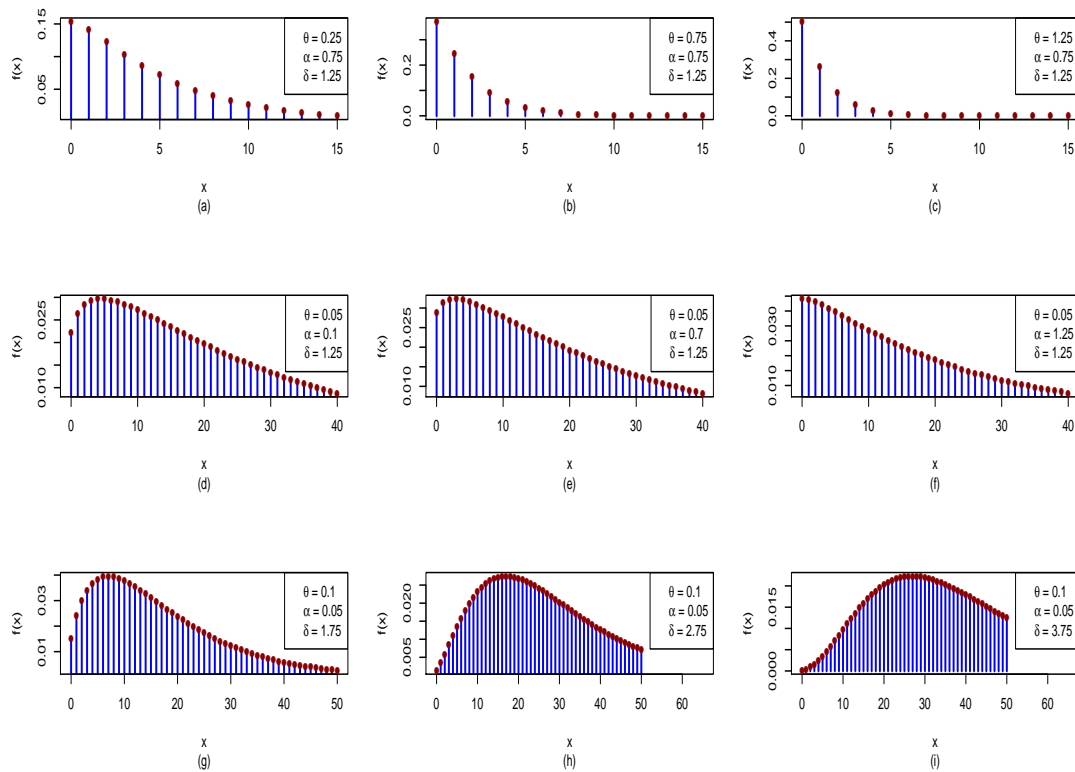and $(p)_i = \frac{\Gamma(p+i)}{\Gamma(p)} = p(p+1)...(p+i+1)$ is the Pochhammer symbol.

Figure 1: *The probability mass function of the PMQLD at different parameter values of $\theta, \alpha,$ and $,\delta$*

3. Statistical properties of PMQLD

In this section, we present some important statistical properties of the PMQLD such as the shape of the distribution, moments and related measures, probability and moment generating functions, and quantile function.
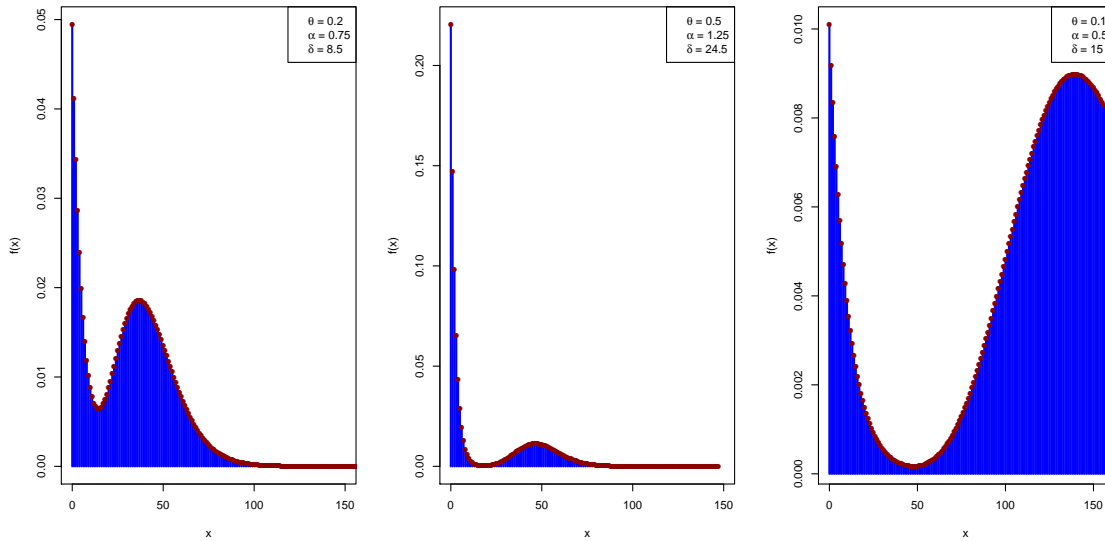
Figure 2: *Some bimodal distributions of the PMQLD*

3.1. Shape of the distribution

From (6) we can easily derive $f(0) = \frac{\theta((1+\theta)^{\delta-1}\alpha^3 + \theta^{\delta-1})}{(\alpha^3+1)(1+\theta)^\delta}$, and $\lim_{x\to\infty} f(x) = 0$.

The recurrence relation for probabilities is given by

$$\frac{f(x+1)}{f(x)} = \frac{A}{(x+1)(1+\theta)B}; \quad x = 0,1,2,...,\tag{8}$$

where $A = (1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x+2) + \theta^{\delta-1}(x+\delta)\Gamma(x+\delta)$, and

$B = (1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x+1) + \theta^{\delta-1}\Gamma(x+\delta))$.

The PMQLD(θ,α,δ) has a log-concave probability mass function when

$$\Delta\eta(x) = \frac{f(x+1)}{f(x)} - \frac{f(x+2)}{f(x+1)} > 0, \forall x \text{ (Gupta et al., 1997)}$$

$$\Rightarrow \frac{(x+2)A^2}{(x+1)\left((1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x+3) + \theta^{\delta-1}\Gamma(x+\delta+2)\right)B} > 1.$$

Under this condition, the distribution represents a unimodal distribution. Further, by using (8), it can be shown that

(i) For $(1+\theta)^{\delta-1}\alpha^3 + \theta^{\delta-1}\delta < (1+\theta)((1+\theta)^{\delta-1}\alpha^3 + \theta^{\delta-1})$, equation (6) has unique mode at $X = 0$.

(ii) Equation (6) has a unique mode at $X = x_0$, for

$$\frac{(1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x_0+2) + \theta^{\delta-1}(x_0+\delta)\Gamma(x_0+\delta)}{(x_0+1)(1+\theta)\left((1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x_0+1) + \theta^{\delta-1}\Gamma(x_0+\delta)\right)} < 1,$$

and

$$\frac{(1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x_0) + \theta^{\delta-1}\Gamma(x_0+\delta-1)}{x_0(1+\theta)\left((1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x_0+1) + \theta^{\delta-1}\Gamma(x_0+\delta)\right)} < 1.$$

(iii) Equation (6) has two modes at $X = x_0$ and $X = x_0 + 1$, for

$$(1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x_0+2) + \theta^{\delta-1}(x_0+\delta)\Gamma(x_0+\delta)$$

$$= (x_0+1)(1+\theta)((1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x_0+1) + \theta^{\delta-1}\Gamma(x_0+\delta)).$$

The above facts are also shown in Figures 1 and 2 at different parameter settings.

Further, the PMQLD(θ, α, δ) has a log-convex probability mass function when $\Delta\eta(x) \leq 0$:

$$\Rightarrow \frac{(x+2)A^2}{(x+1)\left((1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x+3) + \theta^{\delta-1}\Gamma(x+\delta+2)\right)B} \leq 1.$$

3.2. Survival and hazard rate functions

The survival/reliability function is associated with the probability of a system that will survive beyond a specified time. The survival function of the PMQLD is defined as

$$S(x) = 1 - F(x) = 1 - \frac{\beta_1 + \theta^\delta\Gamma(x+\delta+1)_2F_1(1, x+\delta+1; \delta+1; \frac{\theta}{1+\theta})}{\beta_2} \tag{9}$$

where $\beta_1 = \delta(1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x+1)((1+\theta)^{x+1}-1)$ and $\beta_2 = (\alpha^3+1)\Gamma(\delta)x!\delta(1+\theta)^{x+\delta+1}$. The hazard rate function (hrf) is the instantaneous failure rate. The hrf of the PMQLD is defined as

$$h(x) = \lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x | X > x)}{\Delta x}$$

$$= \frac{f(x)}{S(x)} = \frac{\delta\theta(1+\theta)\left((1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x+1) + \theta^{\delta-1}\Gamma(x+\delta)\right)}{\beta_1 - (\beta_2 + \theta^\delta\Gamma(x+\delta+1)_2F_1(1, x+\delta+1; \delta+1; \frac{\theta}{1+\theta}))}. \tag{10}$$

Figure 3 provides an illustration of the possible shapes of the PMQLD's hazard rate function at different shape parameter values. According to these illustrations, it is clear that the proposed model has the capability to model the bathtub, monotonic increasing, and decreasing failure rate shapes.

3.3. Moments and related measures

The central tendency, horizontal symmetry, tail heaviness, and dispersion are important characteristics of a distribution. These characteristics can be studied by using the moments. The following theorem provides the rth factorial moment of the PMQLD.

Theorem 2. Let $X \sim$PMQLD(θ, α, δ), then the rth factorial moment of X is given as

$$\mu'_{(r)} = \frac{\Gamma(\delta)\Gamma(r+1)\alpha^3 + \Gamma(\delta+r)}{(\alpha^3+1)\Gamma(\delta)\theta^r}. \tag{11}$$

Proof.

$$\mu'_{(r)} = E(\Pi_{i=0}^{r-1}(X-i))$$
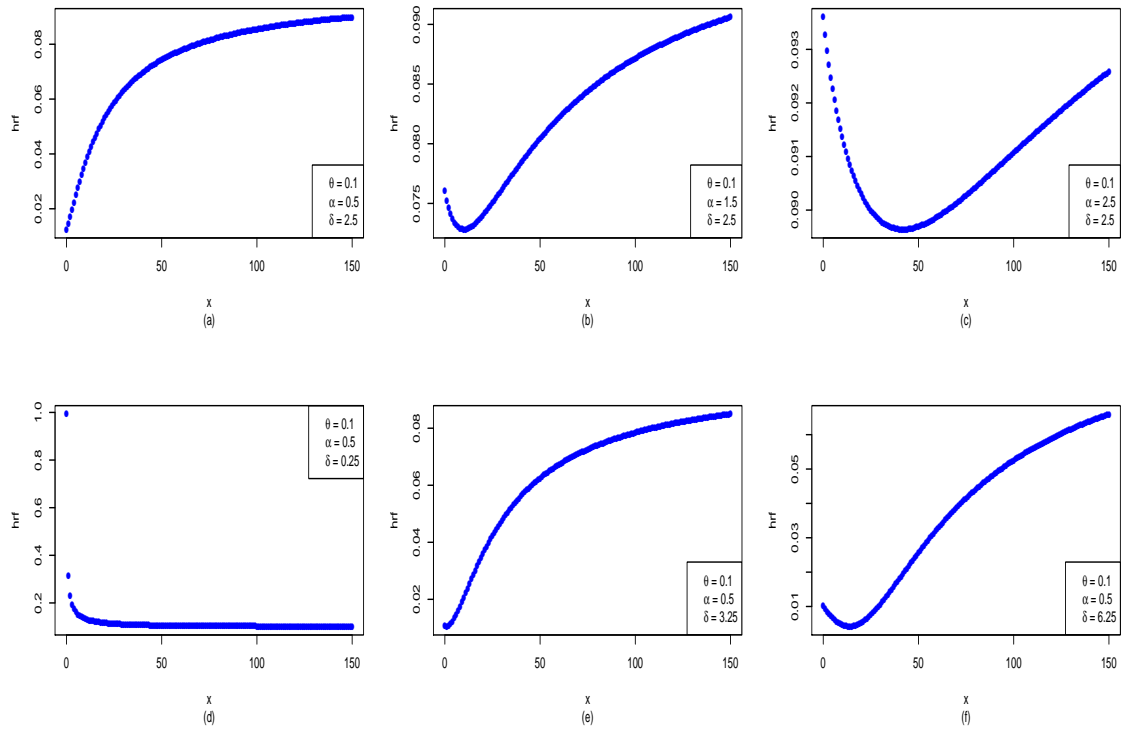
Figure 3: *The hazard rate function at different values of α, and δ*

$$= \sum_{x=0}^{\infty} x^{(r)} \int_0^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \frac{\theta e^{-\theta \lambda}}{(\alpha^3 + 1)\Gamma(\delta)} \left(\Gamma(\delta)\alpha^3 + (\theta\lambda)^{\delta-1} \right) d\lambda$$

$$= \int_0^{\infty} \lambda^r \frac{\theta e^{-\theta \lambda}}{(\alpha^3 + 1)\Gamma(\delta)} \left(\Gamma(\delta)\alpha^3 + (\theta\lambda)^{\delta-1} \right) d\lambda$$

$$= \frac{\theta}{(\alpha^3 + 1)\Gamma(\delta)} \left(\frac{\Gamma(\delta)\alpha^3 \Gamma(r+1)}{\theta^{r+1}} + \frac{\Gamma(\delta+r)}{\theta^{r+1}} \right)$$

$$= \frac{\Gamma(\delta)\Gamma(r+1)\alpha^3 + \Gamma(\delta+r)}{(\alpha^3 + 1)\Gamma(\delta)\theta^r}.$$

∎

Then, the first four raw moments of X can be derived by the following relationship

$$\mu'_r = E(x^r) = \sum_{i=0}^{r} S(r,i)\mu'_{(i)} \; ; \; r = 1, 2, ...,$$

where $S(r,i)$ is the Stirling numbers of the second kind, and it is defined as

$$S(r,i) = \frac{1}{i!} \sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} j^r \; , \; 0 < i < r.$$

Let

$\kappa_1 = \alpha^3 + \delta, \; \kappa_2 = 2\alpha^3 + \delta(\delta+1), \; \kappa_3 = 6\alpha^3 + \delta(\delta+1)(\delta+2), \; \kappa_4 = 24\alpha^3 + \delta(\delta+1)(\delta+2)(\delta+3).$

Then,

$$\mu'_1 = \frac{\kappa_1}{(\alpha^3+1)\theta} = \mu, \mu'_2 = \frac{\theta\kappa_1+\kappa_2}{(\alpha^3+1)\theta^2}, \ \mu'_3 = \frac{\theta^2\kappa_1+3\theta\kappa_2+\kappa_3}{(\alpha^3+1)\theta^3}, \ \mu'_4 = \frac{\theta^3\kappa_1+7\theta^2\kappa_2+6\theta\kappa_3+\kappa_4}{(\alpha^3+1)\theta^4}.$$

Further, the rth-order moments about the mean can be obtained by using the relationship between moments about the mean and moments about the origin, i.e.

$$\mu_r = E\left[(Y-\mu)^r\right] = \sum_{i=0}^{r}\binom{r}{i}(-1)^{r-i}\mu'_i\mu^{r-i}; \ \ r = 1, 2, \dots .$$

Therefore, the variance of X, σ^2 and index of dispersion, γ_1 are derived as

$$\sigma^2 = \mu_2 = \frac{\alpha^6(1+\theta)+\alpha^3(2+\theta+\delta(\theta+\delta-1))+\delta(\theta+1)}{(\alpha^3+1)^2\theta^2} =$$
$$\mu + \mu^2\left(\frac{\alpha^3(\alpha^3+2+\delta(\delta-1))+\delta}{(\alpha^3+\delta)^2}\right),$$

and

$$\gamma_1 = \frac{\mu_2}{\mu'_1} = 1 + \frac{\alpha^3(\alpha^3+2+\delta(\delta-1))+\delta}{(\alpha^3+1)(\alpha^3+\delta)\theta},$$

respectively. It is clear that the $\gamma_1 > 1$. Then, the PMQLD is an over-dispersed distribution. Since the mathematical expressions of μ_3 and μ_4 are very long, we present the graphical presentations of the skewness $(\gamma_2) = \frac{\mu_3}{(\mu_2)^{3/2}}$ and kurtosis $(\gamma_3) = \frac{\mu_4}{\mu_2^2}$ of the PMQLD in Figure 5. The surface plots in Figures 4, and 5 show some possible values of the index of dispersion, skewness, and kurtosis that can be accommodated by the PMQLD at different settings of the parameters. Hence, these plots indicate that the PMQLD (θ, α, δ) has the capability to accommodate various ranges of the index of dispersion, skewness, and kurtosis at different sets of parameters for over-dispersed count data.
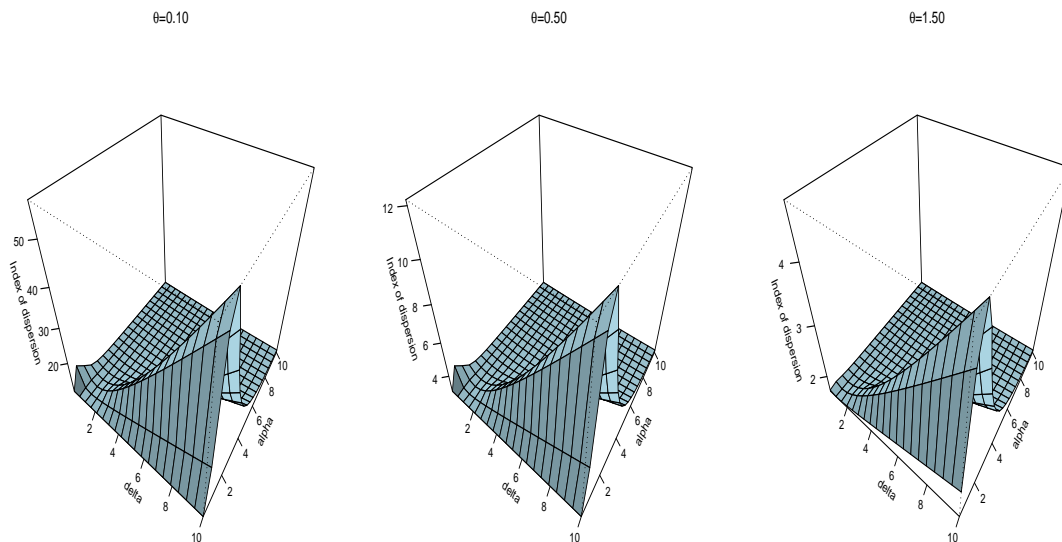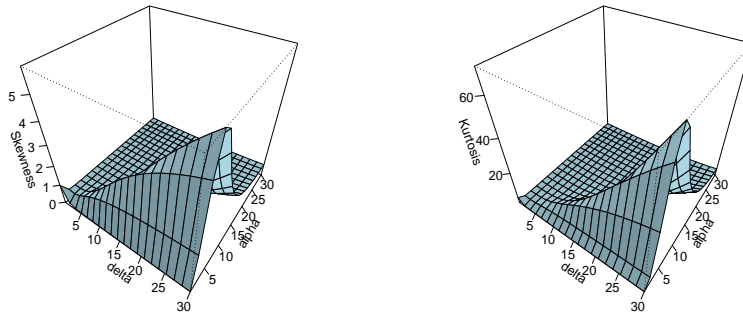

Figure 4: *Surface plots for the index of dispersion at different values of θ, α, and δ*

Figure 5: *Surface plots for the skewness and kurtosis functions at different values of θ, and α*

3.4. Probability and moment generating functions

The characteristics of a probability distribution are directly associated with its probability generating function (pgf) and the moment generating function (mgf). The following theorem provides the pgf of the PMQLD.

Theorem 3. The pgf, $G(t) = E(t^X)$, $X \sim \text{PMQLD}(\theta, \alpha, \delta)$ is given as

$$G(t) = \frac{\theta(\alpha^3(1 - t + \theta)^{\delta-1} + \theta^{\delta-1})}{(\alpha^3 + 1)(1 - t + \theta)^{\delta}}, t \in \mathbb{R}. \tag{12}$$

Proof.

$$G(t) = E(t^X)$$

$$= \sum_{x=0}^{\infty} t^X \int_0^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} \frac{\theta e^{-\theta\lambda}}{(\alpha^3 + 1)\Gamma(\delta)} \left(\Gamma(\delta)\alpha^3 + (\theta\lambda)^{\delta-1}\right) d\lambda$$

$$= \int_0^{\infty} e^{-\lambda(1-t)} \frac{\theta e^{-\theta\lambda}}{(\alpha^3 + 1)\Gamma(\delta)} \left(\Gamma(\delta)\alpha^3 + (\theta\lambda)^{\delta-1}\right) d\lambda$$

$$= \frac{\theta}{(\alpha^3 + 1)\Gamma(\delta)} \left(\Gamma(\delta)\alpha^3 \int_0^{\infty} e^{-\lambda(1-t+\theta)}d\lambda + \theta^{\delta-1} \int_0^{\infty} e^{-\lambda(1-t+\theta)}\lambda^{\delta-1}d\lambda\right)$$

$$= \frac{\theta(\alpha^3(1 - t + \theta)^{\delta-1} + \theta^{\delta-1})}{(\alpha^3 + 1)(1 - t + \theta)^{\delta}}.$$

∎

The mgf can be obtained effortlessly from pgf by using the relationship $G(e^t) = E(e^{tX}) = M_X(t)$, and given as

$$M_X(t) = \frac{\theta(\alpha^3(1 - e^t + \theta)^{\delta-1} + \theta^{\delta-1})}{(\alpha^3 + 1)(1 - e^t + \theta)^{\delta}}, t \in \mathbb{R}. \tag{13}$$

3.5. Quantile function

The quantile function is a useful function to estimate the quantiles. Let us define the quantiles for random variable $X \sim \text{PMQLD}(\theta, \alpha, \delta)$. The uth quantile can be derived by solving $F(x_u) = u$ for $x_u, 0 < u < 1$. Then, the uth quantile function of the PMQLD is given as

$$\beta_1(x_u) + \theta^{\delta}\Gamma(x_u + \delta + 1)_2F_1(1, x_u + \delta + 1; \delta + 1; \frac{\theta}{1 + \theta}) - u\beta_2(x_u) = 0, \ 0 < u < 1, \tag{14}$$

where $\beta_1(x_u) = \delta(1+\theta)^{\delta-1}\Gamma(\delta)\alpha^3\Gamma(x_u+1)((1+\theta)^{x_u+1}-1)$ and $\beta_2(x_u) = (\alpha^3+1)\Gamma(\delta)x_u!\delta(1+\theta)^{x_u+\delta+1}$.

Since equation (14) is not a closed-form in x_u, the estimates of the quantiles can be evaluated by using any numerical method. Further, the first three quartiles can be calculated by substituting $u = 0.25, 0.50$, and 0.75 in equation (14) and solving the respective equations.

4. SIMULATION AND PARAMETER ESTIMATION

4.1. Simulation of the random variables

Here, we provide two different algorithms to simulate the random variables $x_1, x_2, ..., x_n$ from the PMQLD(θ, α, δ) with size n based on the inverse transform method.

The first algorithm is obtained by considering the mixing of the PMQLD. Since $X|\Lambda \sim$Poisson (λ) and $\Lambda \sim$MQLD(θ, α, δ), the first algorithm is obtained as follows

Algorithm I:

i Simulate the random variables, $u_i \sim$uniform$(0,1)$; $i = 1, 2, ..., n$.

ii Solve the non-linear equation for λ_i: $\Gamma(\delta)(1+\alpha^3(1-e^{-\theta\lambda_i})) - \Gamma(\delta, \theta\lambda_i) - u_i(\alpha^3+1)\Gamma(\delta) = 0$ to simulate the random variables, $\lambda_i \sim$MQLD(θ, α, δ); $i = 1, 2, ..., n$.

iii Simulate x_i from Poisson (λ_i); $i = 1, 2, ..., n$.

The second algorithm is obtained from the quantile function of PMQLD discussed in subsection 3.5, and the steps are as follows

Algorithm II:

i Simulate the random variables, $u_i \sim$uniform$(0,1)$; $i = 1, 2, ..., n$.

ii Solve the non-linear equation for $[x_{u_i}]$;
$\beta_1(x_{u_i}) + \theta^\delta\Gamma(x_{u_i}+\delta+1)_2F_1(1, x_{u_i}+\delta+1; \delta+1; \frac{\theta}{1+\theta}) - u_i\beta_2(x_{u_i}) = 0$, where $\beta_1(x_{u_i})$ and $\beta_2(x_{u_i})$ are defined as in section 3.6. $[.]$ denotes the integer part.

4.2. Parameter estimation of PMQLD

In this subsection, we discuss the parameter estimation of the PMQLD by using the method of moment estimation and the maximum likelihood estimation method.

4.2.1 Method of moment estimation (MME)

Given a random samples $x_1, x_2...x_n$ with size n from the PMQLD(θ, α, δ), the method of moment estimators of θ, α, and δ, abbreviated as $\hat{\theta}_{MME}, \hat{\alpha}_{MME}$, and $\hat{\delta}_{MME}$, can be evaluated by equating the raw-moments, say μ'_r, to the sample moments, say $\dfrac{\sum_{i=1}^{n} x_i^r}{n}, r = 1, 2, 3$, i.e. we need to find the solutions of the following system of non-linear equations:

$$n\kappa_1 - (\alpha^3+1)\theta\sum_{i=1}^{n}x_i = 0; \quad n(\theta\kappa_1+\kappa_2) - (\alpha^3+1)\theta^2\sum_{i=1}^{n}x_i^2 = 0;$$

$$n(\theta^2\kappa_1+3\kappa_2+\kappa_3) - (\alpha^3+1)\theta^3\sum_{i=1}^{n}x_i^3 = 0,$$

where κ_1, κ_2, and κ_3 are defined in subsection 3.3. It is clear that these equations are not a closed form. However, the solutions can be derived by using a numerical method.

4.2.2 Maximum likelihood estimation (MLE)

Given a random samples $x_1, x_2 ... x_n$ with size n from the PMQLD(θ, α, δ), the likelihood function of the ith sample value x_i is given as

$$L(\theta, \alpha, \delta | x) = \frac{\theta}{x_i!(\alpha^3 + 1)(1+\theta)^{x_i+\delta}\Gamma(\delta)} \left(\Gamma(\delta)\Gamma(x_i + 1)\alpha^3(1+\theta)^{\delta-1} + \theta^{\delta-1}\Gamma(x_i + \delta) \right).$$

Then, the log-likelihood function is given as

$$log(L(\theta, \alpha, \delta | x_i)) = l(\theta, \alpha, \delta | x)$$

$$= n \left(log(\theta) - log(\alpha^3 + 1) - log(\Gamma(\delta)) \right) + \sum_{i=1}^{n} log \left(\Gamma(\delta)\Gamma(x_i + 1)\alpha^3(1+\theta)^{\delta-1} + \theta^{\delta-1}\Gamma(x_i + \delta) \right)$$

$$- log(x_i!) - (x_i + \delta)log(1 + \theta). \tag{15}$$

The score functions are

$$\frac{\partial l(\theta, \alpha, \delta | x)}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \frac{\Gamma(\delta)\Gamma(x_i + 1)\alpha^3(\delta - 1)(1+\theta)^{\delta-2} + \Gamma(x_i + \delta)(\delta - 1)\theta^{\delta-2}}{\Gamma(\delta)\Gamma(x_i + 1)\alpha^3(1+\theta)^{\delta-1} + \theta^{\delta-1}\Gamma(x_i + \delta)} - \sum_{i=1}^{n} \frac{x_i + \delta}{1 + \theta},$$

$$\frac{\partial l(\theta, \alpha, \delta | x)}{\partial \alpha} = \sum_{i=1}^{n} \frac{3\alpha^2 \Gamma(\delta)\Gamma(x_i + 1)(1+\theta)^{\delta-1}}{\Gamma(\delta)\Gamma(x_i + 1)\alpha^3(1+\theta)^{\delta-1} + \theta^{\delta-1}\Gamma(x_i + \delta)} - \frac{3n\alpha^2}{\alpha^3 + 1},$$

and

$$\frac{\partial l(\theta, \alpha, \delta | x)}{\partial \delta} =$$

$$\sum_{i=1}^{n} \frac{\Gamma(x_i + 1)\alpha^3(\Gamma(\delta)(1+\theta)^{\delta-1}log(1+\theta) + (1+\theta)^{\delta-1}\Gamma(\delta)\psi(\delta)) + \Gamma(x_i + \delta)\theta^{\delta-1}(\log(\theta) + \psi(x_i + \delta))}{\Gamma(\delta)\Gamma(x_i + 1)\alpha^3(1+\theta)^{\delta-1} + \theta^{\delta-1}\Gamma(x_i + \delta)}$$

$$- n(log(1 + \theta) + \psi(\delta)),$$

where $\psi(a) = \frac{\partial}{\partial a}log\Gamma(a) = \frac{\Gamma'(a)}{\Gamma(a)}$. By setting the score functions equal to zero, the maximum likelihood estimators of θ, α, and δ, abbreviated as $\hat{\theta}_{MLE}, \hat{\alpha}_{MLE}$, and $\hat{\delta}_{MLE}$ can be derived. These systems of non-linear equations can be solved by a numerical method. Here, the solutions of the parameter estimates will be obtained by using the *optim* function in the R package *stats*.

The asymptotic confidence intervals for the parameters θ, α, and δ are derived by the asymptotic theory. The estimators are asymptotic three-variate normal with mean (θ, α, δ) and the observed information matrix

$$I(\theta, \alpha, \delta) = \begin{pmatrix} -\dfrac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \theta^2} & -\dfrac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \theta \partial \alpha} & -\dfrac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \theta \partial \delta} \\ -\dfrac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \alpha \partial \theta} & -\dfrac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \alpha^2} & -\dfrac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \alpha \partial \delta} \\ -\dfrac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \delta \partial \theta} & -\dfrac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \delta \partial \alpha} & -\dfrac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \delta^2} \end{pmatrix}$$

at $\theta = \hat{\theta}_{MLE}, \alpha = \hat{\alpha}_{MLE}$, and $\delta = \hat{\delta}_{MLE}$, i.e. $(\hat{\theta}_{MLE}, \hat{\alpha}_{MLE}, \hat{\delta}_{MLE}) \sim N_3((\theta, \alpha, \delta), I^{-1}(\theta, \alpha, \delta))$. The second order partial derivatives of the log-likelihood function are given in Appendix.

Therefore, a $(1 - a)100\%$ confidence interval for the parameters θ, α, and δ are given by

$$\hat{\theta_{MLE}} \pm z_{a/2}\sqrt{Var(\hat{\theta_{MLE}})}, \quad \hat{\alpha}_{MLE} \pm z_{a/2}\sqrt{Var(\hat{\alpha}_{MLE})}, \quad \hat{\delta}_{MLE} \pm z_{a/2}\sqrt{Var(\hat{\delta}_{MLE})},$$

wherein, the $Var(\hat{\theta}_{MLE}), Var(\hat{\alpha}_{MLE})$, and $Var(\hat{\delta}_{MLE})$ are the variance of $\hat{\theta}_{MLE}, \hat{\alpha}_{MLE}$, and $\hat{\delta}_{MLE}$, respectively, and can be derived by diagonal elements of $I^{-1}(\theta, \alpha, \delta)$ and $z_{a/2}$ is the critical value at a level of significance.

5. MONTE CARLO SIMULATION STUDY AND REAL-WORLD APPLICATION

This section is devoted to discuss the simulation study and the applicability of PMQLD.

5.1. Monte Carlo simulation study

Here, we examine the accuracy of the MLE method in the unknown parameter estimation of the PMQLD with respect to sample size n. The second algorithm given in subsection 4.1 is used to simulate the random variables from the PMQLD. The sample sizes are taken as $60, 100, 200$, and 300, and the simulation study is repeated 1000 times. The study is designed as follows

(i) Simulate 1000 samples of size n.

(ii) Compute the maximum likelihood estimates for the 1000 samples, say $(\hat{\theta}_i, \hat{\alpha}_i, \hat{\delta}_i), i = 1, 2, ...1000$.

(iii) Compute the average MLEs, biases, and mean square errors (MSEs) by using the following equations

$$\hat{s}(n) = \frac{1}{1000} \sum_{i=1}^{1000} \hat{s}_i, \quad bias_s(n) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{s}_i - s), \text{ and } MSE_s(n) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{s}_i - s)^2,$$

for $s = \theta, \alpha, \delta$, and $n = 60, 100, 200, 300$.

Tables 2 to 5 represent the average MLEs, biases, and MSEs (in parentheses) of θ, α, and δ for different values of θ, α, and δ which are $\theta = 0.1, 0.3$; $\alpha = 0.25, 0.50, 0.75$; and $\delta = 2.50, 3.50, 4.50$. Note that the biases and MSEs decrease as n increases for all parameters. Then, MLE method verifies the asymptotic property for all parameter estimates, and the parameters θ, α, and δ are over estimated. Further, while the estimation of θ is good for small value of θ, the estimation of α doses not show a good estimation for small value of α based on average biases and MSEs. However, there is no particular pattern for estimation of δ.

Table 2: *Performance of MLE method for the PMQLD($\theta = 0.10, \alpha = 0.50, \delta$)*

	$n = 60$		$n = 100$		$n = 200$		$n = 300$	
	MLE	Bias(MSE)	MLE	Bias(MSE)	MLE	Bias(MSE)	MLE	Bias(MSE)
$\delta = 2.50$								
θ	0.1151	0.0152(0.0014)	0.1086	0.0086(0.0007)	0.1067	0.0067(0.0002)	0.1009	0.0009(0.0001)
α	0.7519	0.2519(0.0691)	0.7482	0.2482(0.0648)	0.7290	0.2290(0.0557)	0.7195	0.2195(0.0494)
δ	3.0685	0.5685(1.0107)	3.0545	0.5545 (0.9155)	2.8466	0.3466(0.3078)	2.8358	0.3358(0.2186)
$\delta = 3.50$								
θ	0.1101	0.0101(0.0010)	0.1051	0.0051(0.0006)	0.1044	0.0044(0.0002)	0.1029	0.0029(0.0001)
α	0.6733	0.1733(0.0348)	0.6662	0.1662(0.0317)	0.6684	0.1684(0.0300)	0.6509	0.1509(0.0259)
δ	3.8991	0.3991(0.9110)	3.8494	0.3494(0.8928)	3.6791	0.1791(0.2898)	3.6010	0.1010(0.0887)
$\delta = 4.50$								
θ	0.1087	0.0087(0.0009)	0.1054	0.0054(0.0006)	0.1037	0.0037(0.0002)	0.1029	0.0021(0.0001)
α	0.6299	0.1299(0.0213)	0.6268	0.1268(0.0191)	0.6217	0.1217(0.0161)	0.6192	0.1192(0.0156)
δ	4.8664	0.3664(1.3188)	4.7506	0.2506(1.1957)	4.6354	0.1354(0.3871)	4.5805	0.0805(0.0122)

Table 3: *Performance of MLE method for PMQLD($\theta = 0.30, \alpha = 0.50, \delta$)*

	$n = 60$		$n = 100$		$n = 200$		$n = 300$	
	MLE	Bias(MSE)	MLE	Bias(MSE)	MLE	Bias(MSE)	MLE	Bias(MSE)
$\delta = 2.50$								
θ	0.3791	0.0790 (0.0242)	0.3635	0.0635(0.0215)	0.3378	0.0378 (0.0075)	0.3367	0.0367(0.0039)
α	0.8392	0.3392 (0.1199)	0.8243	0.3243(0.1109)	0.7956	0.2956 (0.0933)	0.7503	0.2503(0.0853)
δ	3.6229	1.1229(3.5560)	3.4672	0.9672(2.9115)	3.1876	0.6876(1.2228)	3.1509	0.6509 (0.7488)
$\delta = 3.50$								
θ	0.3839	0.0839 (0.0245)	0.3710	0.0710(0.0216)	0.3275	0.0275(0.0042)	0.3207	0.0207(0.0018)
α	0.7493	0.2493 (0.0698)	0.7215	0.2215(0.0543)	0.7170	0.2170(0.0487)	0.6983	0.1983(0.0457)
δ	4.3284	0.8284(2.5256)	4.1190	0.6190(1.9316)	3.7680	0.2680(0.4910)	3.7393	0.2393(0.2547)
$\delta = 4.50$								
θ	0.3638	0.0638 (0.0210)	0.3484	0.0484(0.0118)	0.3143	0.0143(0.0033)	0.3117	0.0117(0.0016)
α	0.6784	0.1784 (0.0373)	0.6703	0.1703(0.0330)	0.6658	0.1658(0.0299)	0.6578	0.1578(0.0277)
δ	5.1958	0.6958(3.0589)	4.9240	0.4240(2.0855)	4.6154	0.1154(0.6430)	4.6713	0.1713(0.2667)

Table 4: *Performance of MLE method for PMQLD($\theta = 0.10, \alpha, \delta = 2.50$)*

	$n = 60$		$n = 100$		$n = 200$		$n = 300$	
	MLE	Bias(MSE)	MLE	Bias(MSE)	MLE	Bias(MSE)	MLE	Bias(MSE)
$\alpha = 0.25$								
θ	0.1128	0.0128(0.0014)	0.1079	0.0079(0.0007)	0.1080	0.0080(0.0002)	0.1027	0.0027 (0.0001)
α	0.7109	0.4609(0.2184)	0.7034	0.4534(0.2088)	0.6957	0.4457(0.2000)	0.6804	0.4304 (0.1888)
δ	3.0685	0.5685(1.0447)	3.0085	0.5085(0.7195)	2.9490	0.4490(0.3138)	2.8473	0.3473 (0.1865)
$\alpha = 0.75$								
θ	0.1351	0.0351(0.0065)	0.1217	0.0217(0.0017)	0.1203	0.0203(0.0023)	0.1156	0.0156(0.0004)
α	0.8487	0.0987(0.0167)	0.8414	0.0914(0.0135)	0.8318	0.0818(0.0095)	0.8190	0.0690(0.0081)
δ	3.1905	0.6905(2.0102)	3.1765	0.6765(1.4733)	3.0104	0.5104(0.7818)	2.8781	0.3781(0.3592)

Table 5: *Performance of MLE method for PMQLD($\theta = 0.30, \alpha, \delta = 2.50$)*

|	$n = 60$		$n = 100$		$n = 200$		$n = 300$	
	MLE	Bias(MSE)	MLE	Bias(MSE)	MLE	Bias(MSE)	MLE	Bias(MSE)
$\alpha = 0.25$								
θ	0.3998	0.0998(0.0329)	0.3805	0.0805(0.0240)	0.3482	0.0482(0.0060)	0.3336	0.0336(0.0030)
α	0.7823	0.5323(0.2928)	0.7795	0.5295(0.2873)	0.7611	0.5111(0.2636)	0.7437	0.4937(0.2508)
δ	3.5545	1.0545(2.7164)	3.4774	0.9774(2.2678)	3.1584	0.6584(0.7994)	3.0583	0.5583(0.5396)
$\alpha = 0.75$								
θ	0.3773	0.0773(0.0621)	0.3566	0.0566(0.0506)	0.3275	0.0275(0.0267)	0.3097	0.0097(0.0225)
α	0.8962	0.1462(0.0280)	0.8889	0.1389(0.0221)	0.8693	0.1193(0.0186)	0.8475	0.0975(0.0175)
δ	2.9598	0.4598(4.7991)	2.8278	0.3278(4.4320)	2.6089	0.1089(2.7618)	2.5723	0.0723(2.4811)

5.2. Real-world applications

In this subsection, we discuss the real-world applications of the proposed mixed Poisson distribution. Two data sets are considered to illustrate whether the proposed distribution is well fitted compared to some other existing competing Poisson mixtures. The best-fitted distribution was selected based on the negative log-likelihood ($-2logL$), Akaike Information Criterion (AIC), and chi-square goodness of fit statistic. The unknown parameters of the models are estimated by using the MLE method. Tables 6 and 7 summarize all these statistical measures for each data set, and the standard errors of the parameter estimates are reported in parentheses.

The first data set contains the epileptic seizure counts (Chakraborty, 2010). The sample index dispersion (τ) of this data set is 1.867. Since τ value is greater than one, the distribution of

this data set is clearly over-dispersed. Also, the skewness and excess kurtosis for this example are 1.239 and 1.680, respectively, which show that the distribution of the data set is positively skewed and leptokurtic. This data set was used to fit the PMQLD, GD, NBD, PLD, GPLD, PQLD, and PGLD. Table 6 presents the estimates of the parameters of distributions and the goodness of fit test. Of all eight distributions, the PMQLD performs well based on the smallest AIC value of 1191.83 and the smallest chi-square value (χ^2) of 2.93 (p-value=0.71).

The second data set represents the number of roots produced by 270 micro-propagated shoots of the columnar apple cultivar Trajan (Ridout et al., 1998). This is a bimodal data set for which the sample index dispersion, skewness, and excess kurtosis are 3.077, 0.182, and -1.056, respectively. These values indicate that the distribution of the data set is extremely over-dispersed, mild positively skewed, and platykurtic. This data set was also used to fit the same distributions that we used for the first example. Table 7 summarizes the results of parameter estimations and the goodness of fit test. The results show that the PMQLD having AIC=1350.20, $\chi^2 = 11.75$, p-value=0.47 outperforms clearly than other distributions.

Figure 6 illustrates how the expected values of the proposed distribution adhere with the observed value for the data sets. We can see that the observed values of the first and second data sets are very close to the expected values of the PMQLD, and the observed values of the third data set are very close to the expected values of the ZMPQLD.
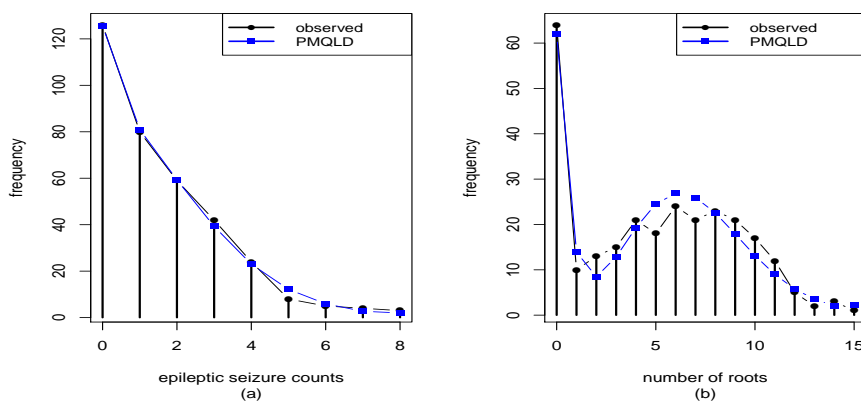
Figure 6: *Performance of PMQLD for the real-data sets*

Table 6. Epileptic seizure counts

Counts	Observed	Expected						
		GD	NBD	PLD	GPLD	PQLD	PGLD	PMQLD
0	126	137.95	120.22	128.72	121.93	121.82	122.85	125.65
1	80	83.73	93.00	87.14	90.92	90.95	90.08	80.98
2	59	50.82	59.18	55.26	58.72	58.76	57.85	59.30
3	42	30.85	34.94	33.63	35.20	35.23	35.20	39.25
4	24	18.72	19.84	19.89	20.16	20.17	20.53	22.99
5	8	11.37	10.99	11.52	11.20	11.20	11.54	12.17
6	5	6.90	5.98	6.57	6.08	6.08	6.27	5.94
7	4	4.19	3.22	3.70	3.25	3.25	3.31	2.72
8	3	6.47	3.63	4.57	3.54	3.54	3.37	2.00
Total	351	351	351	351	351	351	351	351
MLE		$\hat{\theta} = 0.65$ (0.04)	$\hat{\beta} = 1.00$ (0.19) $\hat{\alpha} = 1.55$ (0.28)	$\hat{\theta} = 0.97$ (0.05)	$\hat{\theta} = 1.11$ (0.13) $\hat{\alpha} = 2.76$ (2.76)	$\hat{\theta} = 1.12$ (0.13) $\hat{\alpha} = 0.38$ (0.33)	$\hat{\theta} = 1.57$ (0.66) $\hat{\alpha} = 1.49$ (0.57) $\hat{\beta} = 3.89$ (2.38)	$\hat{\theta} = 2.70$ (1.26) $\hat{\alpha} = 0.82$ (0.07) $\hat{\delta} = 5.89$ (2.83)
χ^2		11.42	5.67	5.84	4.85	4.86	4.66	2.93
p-value		0.12	0.46	0.56	0.56	0.56	0.46	0.71
$-2logL$		1196.79	1189.88	1190.36	1188.96	1188.96	1188.54	1185.83
AIC		1198.79	1193.88	1192.36	1192.96	1192.96	1194.54	1191.83

Table 7. Number of roots

Counts	Observed	Expected						
		GD	NBD	PLD	GPLD	PQLD	PGLD	PMQLD
0	64	44.62	36.87	31.09	35.46	35.45	82.81	61.93
1	10	37.25	36.05	32.94	34.00	33.99	17.81	13.92
2	13	31.09	32.16	31.79	31.19	31.19	15.31	8.47
3	15	25.95	27.77	29.06	27.78	27.77	16.46	12.85
4	21	21.66	23.58	25.63	24.20	24.20	17.53	19.30
5	18	18.08	19.83	22.04	20.74	20.74	17.70	24.53
6	24	15.09	16.56	18.60	17.55	17.55	16.95	26.85
7	21	12.60	13.76	15.48	14.69	14.69	15.53	25.94
8	23	10.52	11.39	12.73	12.20	12.20	13.70	22.55
9	21	8.78	9.40	10.37	10.05	10.05	11.72	17.90
10	17	7.33	7.74	8.39	8.23	8.24	9.76	13.13
11	12	6.12	6.37	6.74	6.71	6.71	7.96	8.99
12	5	5.11	5.23	5.38	5.44	5.45	6.36	5.78
13	2	4.26	4.28	4.27	4.40	4.40	5.01	3.52
14	3	3.56	3.51	3.38	3.54	3.54	3.88	2.03
≥ 15	1	17.98	15.50	12.11	13.82	13.83	11.51	2.31
Total	270	270	270	270	270	270	270	270
MLE		$\hat{\theta} = 0.12$ (0.01)	$\hat{\beta} = 0.24$ (0.03) $\hat{\alpha} = 1.21$ (0.16)	$\hat{\theta} = 0.35$ (0.02)	$\hat{\theta} = 0.37$ (0.03) $\hat{\alpha} = 0.47$ (0.26)	$\hat{\theta} = 0.32$ (0.03) $\hat{\alpha} = 0.68$ (0.27)	$\hat{\theta} = 0.59$ (0.08) $\hat{\alpha} = 0.22$ (0.12) $\hat{\beta} = 4.25$ (0.66)	$\hat{\theta} = 4.23$ (1.36) $\hat{\alpha} = 0.73$ (0.04) $\hat{\delta} = 29.34$ (9.28)
χ^2		121.92	120.76	117.44	110.58	110.58	46.72	11.74
p-value		0.00	0.00	0.00	0.00	0.00	0.00	0.466
$-2logL$		1464.90	1462.63	1454.10	1451.45	1451.45	1384.21	1344.20
AIC		1466.90	1466.63	1456.10	1455.45	1455.45	1390.21	1350.20

6. Conclusion

This paper proposes an alternative mixed Poisson distribution to model the over-dispersed count data. Explicit expressions of the pmf, hazard rate function, moments, mean, variance, skewness, and kurtosis were derived for the proposed distribution. Its pmf possesses to be either unimodal or bimodal, and hazard rate function presents monotonic increasing, decreasing, and bathtub shapes. The kurtosis and the variance-to-mean ratio functions of the new distribution indicate that

the distribution can capture various ranges of right tail weights as well as the index of dispersions. Further, its structural properties show that the new distribution is much more flexible than its predecessors, negative binomial, geometric, and Poisson-Lindley distributions. The maximum likelihood method was employed to estimate the parameters of the distribution, and the observed information matrix has also been derived. The proposed distribution and some other competing Poisson mixtures have been fitted to two real-world data sets. The results show that the proposed distribution could provide a better fit than a set of common Poisson mixtures considered in these applications.

Acknowledgement

References

[1] Albrecht, P. (1984). Laplace Transforms, Mellin Transforms and Mixed Poisson Processes. Scandinavian Actuarial Journal, 11: 58-64. https://doi.org/10.1080/03461238.1984.10413753

[2] Bhati, D., Sastry, D. V. S., and Qadri, P. Z. (2015). A New Generalized Poisson Lindley Distribution, applications and Properties. Austrian Journal of Statistics, 44: 35-51. https://doi.org/10.17713/ajs.v44i4.54

[3] Bhattacharya, S. K. (1967). A Result in Accident Proneness. Biometrika, 54: 324-325. https://doi.org/10.1093/biomet/54.1-2.324

[4] Cameron, A. C. and Trivedi, P. K. Regression Analysis of Count Data, (2nd Ed.) Cambridge University Press, New York, USA, 2013. https://doi.org/10.1017/CBO9781139013567

[5] Chakraborty, S. (2010). On Some Distributional Properties of the Family of Weighted Generalized Poisson Distribution. Communications in Statistics - Theory and Methods, 39: 2767-2788. https://doi.org/10.1080/03610920903129141

[6] Elbatal, I., Merovci, F., and Elgarhy, M. (2013). A new generalized Lindley distribution, Mathematical Theory and Modeling, 3: 30-47.

[7] Feller, W. (1943). On a Generalized Class of Contagious Distributions. Annals of Mathematical Statistics, 14: 389-400. https://www.jstor.org/stable/2235926

[8] Greenwood, M. and Yule, G. U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents, Journal of the Royal Statistical Society, 83: 255-279. https://www.jstor.org/stable/2341080

[9] Grine, R. and Zeghdoudi, H. (2017). On Poisson quasi-Lindley distribution and its applications. J. of Modern Applied Statistical Methods, 16: 403-417.

[10] Irwin, J. (1975). The Generalized Waring Distribution Parts I, II, III. Journal of the Royal Statistical Society A, 138: 18-31 (Part I), 204-227 (Part II), 374-384 (Part III).

[11] Johnson, N. L., Kotz, S., and Kemp, A. W. Univariate Discrete Distributions, (2nd Ed.) John Wiley and Sons, New York, USA, 1992.

[12] Lindley, D. V. (1958). Fiducial Distributions and Bayes' Theorem. Journal of the Royal Statistical Society, 20: 102-107. https://www.jstor.org/stable/2983909

[13] Lynch, J. (1988). Mixtures, Generalized Convexity and Balayages, Scandinavian Journal of Statistics, 15: 203-210.

[14] McCullagh, P. and Nelder, J. Generalized Linear Models, (2nd Ed.) Chapman and Hall, New York, USA, 1989.

[15] Ong, S.H. and Muthaloo, S. (1995). A Class of Discrete Distributions Suited to Fitting Very Long Tailed Data. Communication in Statistics-Simulation and Computation, 24: 929-945. https://doi.org/10.1080/03610919508813285

[16] Ridout, M., Demktrio, C. G. B, and Hinde, J. (1998). Models for count data with many zeros, Invited paper presented at the Nineteenth International Biometric Conference, Capetown, South Africa.

[17] Ruohonen, M. (1988). A Model for the Claim Number Process. ASTIN Bulletin, 18: 57-68. https://doi.org/10.2143/AST.18.1.2014960

[18] Sankaran, M. (1970). The discrete Poisson-Lindley distribution, Biometrics, 26: 145-149. https://www.jstor.org/stable/2529053

[19] Shaked, M. (1980). On Mixtures from Exponential Families, Journal of the Royal Statistical Society B, 42: 192-198. https://doi.org/10.1111/j.2517-6161.1980.tb01118.x

[20] Shanker, R. and Mishra, A. (2013a). A Quasi Lindley Distribution. African Journal of Mathematics and Computer Science Research, 6: 64-71.

[21] Shanker, R., Sharma, S., and Shanker, R. (2013b). A Two-Parameter Lindley Distribution for Modeling Waiting and Survival Times Data. Applied Mathematics, 4: 363-368.

[22] Tharshan, R. and Wijekoon, P. (2020a). Location Based Generalized Akash Distribution: Properties and Applications, Open Journal of Statistics, 10: 163-187.

[23] Tharshan, R. and Wijekoon, P. (2020b). A comparison study on a new five-parameter generalized Lindley distribution with its sub-models. Statistics in Transition new Series, 21: 89-117. https://doi.org/10.21307/stattrans-2020-015

[24] Tharshan, R. and Wijekoon, P. (2021). A modification of the Quasi Lindley distribution. Open Journal of Statistics, 11: 369-392.

[25] Willmot, G. E. (1993). On Recursive Evaluation of Mixed Poisson Probabilities and Related Quantities, Scandinavian Actuarial Journal, 18: 114-133.

[26] Wongrin, W. and Bodhisuwan, W. (2016). The Poisson-generalised Lindley distribution and its applications, Songklankarin J.Sci. Technol, 38: 645-66.

Appendix: Elements of the observed information matrix, $I(\theta, \alpha, \delta)$ defined in subsection 4.2.1:

Let us define $T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8, T_9, T_{10}, T_{11}, T_{12}, T_{13}$ and T_{14} as follows

$$T_1 = \Gamma(\delta)\Gamma(x_i+1)\alpha^3(\delta-1)(1+\theta)^{\delta-2} + \Gamma(x_i+\delta)(\delta-1)\theta^{\delta-2},$$

$$T_2 = \Gamma(\delta)\Gamma(x_i+1)\alpha^3(1+\theta)^{\delta-1} + \theta^{\delta-1}\Gamma(x_i+\delta),$$

$$T_3 = 3\alpha^2\Gamma(\delta)\Gamma(x_i+1)(1+\theta)^{\delta-1},$$

$$T_4 =$$
$$\Gamma(x_i+1)\alpha^3(\Gamma(\delta)(1+\theta)^{\delta-1}log(1+\theta) + (1+\theta)^{\delta-1}\Gamma(\delta)\psi(\delta)) + \Gamma(x_i+\delta)\theta^{\delta-1}(\log(\theta) + \psi(x_i+\delta)),$$

$$T_5 = \Gamma(\delta)\Gamma(x_i+1)\alpha^3(1+\theta)^{\delta-3} + \Gamma(x_i+\delta)\theta^{\delta-3},$$

$$T_6 = \Gamma(\delta)\Gamma(x_i+1)\alpha^3(1+\theta)^{\delta-2} + \Gamma(x_i+\delta)\theta^{\delta-2},$$

$$T_7 = log(1+\theta)(\Gamma(\delta)(1+\theta)^{\delta-1}log(1+\theta) + (1+\theta)^{\delta-1}\Gamma(\delta)\psi(\delta))$$

$$T_8 = (1+\theta)^{\delta-1}(\Gamma(\delta)\psi_1(\delta) + (\psi(\delta))^2\Gamma(\delta)) + \Gamma(\delta)\psi(\delta)(1+\theta)^{\delta-1}log(1+\theta),$$

$$T_9 = \Gamma(x_i+\delta)\theta^{\delta-1}log(\theta) + \theta^{\delta-1}\Gamma(x_i+\delta)\psi(x_i+\delta),$$

$$T_{10} = \Gamma(x_i+\delta)\theta^{\delta-1}\psi_1(x_i+\delta),$$

$$T_{11} = \Gamma(x_i+1)\alpha^3(\Gamma(\delta)(1+\theta)^{\delta-1}log(1+\theta) + (1+\theta)^{\delta-1}\Gamma(\delta)\psi(\delta)),$$

$$T_{12} = \theta^{\delta-1}\Gamma(x_i+\delta)\psi(x_i+\delta) + \Gamma(x_i+\delta)\theta^{\delta-1}log(\theta),$$

$$T_{13} = \Gamma(\delta)((1+\theta)^{\delta-2} + log(1+\theta)(\delta-1)(1+\theta)^{\delta-2}) + (\delta-1)(1+\theta)^{\delta-2}\Gamma(\delta)\psi(\delta),$$

and

$$T_{14} = (log(\theta) + \psi(x_i+\delta))\Gamma(x_i+\delta)(\delta-1)\theta^{\delta-2}.$$

Then, the second order partial derivatives of the log-likelihood function are as follows

$$\frac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \theta^2} = \frac{-n}{\theta^2} + \sum_{i=1}^{n} \frac{x_i + \delta}{(1 + \theta)^2} + \sum_{i=1}^{n} \frac{T_2(\delta - 1)(\delta - 2)T_5 - T_1(\delta - 1)T_6}{T_2^2},$$

$$\frac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \alpha^2} =$$
$$\sum_{i=1}^{n} \frac{T_2(6\alpha\Gamma(\delta)\Gamma(x_i + 1)(1 + \theta)^{\delta-1}) - T_3(3\alpha^2\Gamma(\delta)\Gamma(x_i + 1)(1 + \theta)^{\delta-1})}{T_2^2} - \frac{3n\alpha(2(\alpha^2 + 1) - 3\alpha^3)}{(\alpha^3 + 1)^2},$$

$$\frac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \delta^2} =$$
$$\sum_{i=1}^{n} \frac{T_2(\Gamma(x_i + 1)\alpha^3(T_7 + T_8) + (log(\theta) + \psi(x_i + \delta))T_9 + T_{10}) - T_4(T_{11} + T_{12})}{T_2^2} - n\psi_1(\delta),$$

$$\frac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \theta \partial \alpha} = \sum_{i=1}^{n} \frac{T_2(3\alpha^2\Gamma(\delta)\Gamma(x_i + 1)(\delta - 1)(1 + \theta)^{\delta-1}) - T_1 T_3}{T_1^2},$$

$$\frac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \delta \partial \alpha} = \sum_{i=1}^{n} \frac{T_2(3\alpha^2\Gamma(x_i + 1)(\Gamma(\delta)(1 + \theta)^{\delta-1}log(1 + \theta) + (1 + \theta)^{\delta-1}\Gamma(\delta)\psi(\delta))) - T_4 T_3}{T_2^2},$$

and

$$\frac{\partial^2 l(\theta, \alpha, \delta | x)}{\partial \delta \partial \theta} = \sum_{i=1}^{n} \frac{T_2(\Gamma(x_i + 1)\alpha^3 T_{13} + \Gamma(x_i + \delta)\theta^{\delta-2} + T_{14}) - T_4 T_1}{T_2^2} - \frac{n}{1 + \theta},$$

where $\psi_1(s)$ is the trigamma function and defined as $\psi_1(s) = \frac{d^2}{ds^2}log(\Gamma(s)) = \sum_{i=1}^{n} \frac{1}{(s + k)^2}$.