Journal of Statistical Modelling: Theory and Applications
Vol. 6, No. 1, 2025, pp. 151-166
Yazd University Press 2025

JSMTA

Research Paper

Geometric log-Lindley distribution as an alternative to the Poisson model with application to insurance data

Hossein Zamani*
Deprtment of Statistics, University of Hormozgan, BandarAbbas, Iran

Received: August 14, 2025/ Revised: December 04, 2025/ Accepted: January 22, 2026

Abstract: Mixed discrete distributions are primarily used for modeling over-dispersed count data. The construction of mixed models, such as the mixed Poisson model, is based on the assumption that the distribution parameter of interest is a random variable that follows a specified distribution. In this framework, the marginal distribution of a discrete random variable forms a mixed distribution. This paper introduces a novel discrete distribution derived from the geometric distribution, assuming that the model's parameter follows the log-Lindley distribution. This approach is motivated by situations where the parameter of the geometric distribution is not constant across populations, as is often the case in insurance data, where the probability of a claim varies between different portfolios. This distribution is particularly well-suited for modeling over-dispersed discrete data. The statistical properties of the proposed distribution are examined, and the parameters of the resulting model are estimated. To evaluate the accuracy of the estimates, a simulation study is conducted, and the model's performance is demonstrated using real data.

Keywords: Count data; Geometric; Log-Lindley; Mixed distribution; Over-dispersed.
Mathematics Subject Classification (2010): 62E10.

1 Introduction

Constructing mixed discrete distributions is a useful method for modeling over-dispersed count data. In this regard, several mixed models, such as the Poisson-gamma model known as the negative binomial distribution (Boucher et al., 2007; Denuit et al., 2007; Greene, 2008; Ismail and Zamani, 2013; Klugman et al., 2012) have been proposed and applied in modeling economics, medical, insurance and traffic data. The

*Corresponding author: `zamani.huni@hormozgan.ac.ir`

Poisson-inverse Gaussian (Tremblay, 1992; Willmot, 1987), Poisson-generalized Lindley (Wongrin and Bodhisuwan, 2017), Poisson-log normal (Izsák, 2008) negative binomial-Lindley (Zamani and Ismail, 2010), beta-binomial (Navarro and Perfors, 2005) and beta-geometric (Weinberg and Gladen, 1986) are the other mixed models that have been proposed and applied in various study fields.

The log-Lindley distribution with probability density function (pdf)

$$g(x;\delta,\gamma) = \frac{\delta^2}{1+\delta\gamma}(\gamma - \log x)x^{\delta-1}, \quad 0 < x < 1,\ \ \delta > 0, \quad \gamma \in \mathbb{R}, \tag{1}$$

introduced by Gómez-Déniz et al. (2014). The log-Lindley distribution which is referred to as log-L(δ,γ) was conducted using the transformation $Y = -\log X$ on the generalized Lindley distribution proposed by Zakerzadeh and Dolati (2009) considering one of the parameters equal to one. The r-th moment and inverse moment of log-L(δ,γ) is given by,

$$E(X^r|\delta,\gamma) = \frac{\delta^2}{1+\delta\gamma}\frac{1+\gamma(\delta+r)}{(\delta+r)^2}, \quad r = \ldots, -2, -1, 1, 2, \ldots.$$

This paper introduces a two-parameter geometric log-Lindley (GLL) distribution by mixing the zero-based geometric distribution and the log-Lindley distribution. We indicate that the proposed model is shown to be effective for modeling diverse types of count data. Various properties of the GLL distribution are derived and discussed. The paper is organized as follows.

Section 2 introduces the two-parameter geometric log-Lindley distribution and examines its basic properties including: the behavior of the probability mass function, the expressions for the moments, the survival and hazard rate functions of the random variable, a recursive formula for calculating probabilities of the model, and the zero-truncated and zero-inflated versions of the GLL model. Section 3 discusses the various methods of parameter estimation. In Section 4, the algorithm for simulating random data from the GLL distribution is presented. In this section, a simulation study is performed to investigate the bias, root mean square error, and the coverage probability of the simulated estimates. An application of the GLL distribution, by fitting this distribution to two insurance datasets and comparing it with alternatives, is given in Section 5.

2 Definition and properties

2.1 Geometric log-Lindley distribution

This section details the construction of the GLL distribution. Suppose the conditional random variable $Y|p$, follow a geometric distribution (denoted $G(p)$) with probability mass function (pmf)

$$P(Y = y|p) = p(1-p)^y, \quad y \in \mathbb{N}_0,\ p \in (0,1), \tag{2}$$

and further assume that p is a random variable following a log-Lindley $p \sim \log \mathrm{L}(\delta,\gamma)$ distribution with pdf given in (1). Therefore, the marginal pmf of the random variable Y which we refer to as the geometric log-Lindley distribution, $Y \sim GLL(\delta,\gamma)$ is therefore given by

$$\begin{aligned} P(y) &= \int_0^1 P(y|p)g(p)dp \\ &= \frac{\delta^2}{1+\delta\gamma}\int_0^1 (\gamma - \log p)p^{\delta}(1-p)^y dp \\ &= \frac{\delta^2}{1+\delta\gamma}\left\{\gamma\int_0^1 p^{\delta}(1-p)^y dp - \frac{\partial}{\partial\delta}\int_0^1 p^{\delta}(1-p)^y dp\right\} \\ &= \frac{\delta^2}{1+\delta\gamma}\left[\gamma \mathrm{B}(\delta{+}1, y+1) - \frac{\partial}{\partial\delta}\mathrm{B}(\delta{+}1, y+1)\right], \quad y \in \mathbb{N}_0 \end{aligned} \tag{3}$$

where $\delta > 0$, $\gamma \in \mathbb{R}$ and $1+\delta\gamma > 0$. Moreover, $\mathrm{B}(r,k) = \frac{\Gamma(r)\Gamma(k)}{\Gamma(r+k)}$ and $\Gamma(k) = \int_0^\infty t^{k-1}e^{-t}dt$. After simplification, the resulting pmf is

$$\frac{\partial}{\partial\delta}\mathrm{B}(\delta{+}1, y+1) = \mathrm{B}(\delta{+}1, y+1)\left\{\psi(\delta+1) - \psi(\delta+y+2)\right\},$$

where $\psi(.)$ is digamma function defined as $\psi(k) = \frac{\partial}{\partial k}\log\Gamma(k) = \frac{\Gamma'(k)}{\Gamma(k)}$. Therefore, the pmf in (3) can be expressed as,

$$P(y;\delta,\gamma) = \frac{\delta^2}{1+\delta\gamma}\mathrm{B}(\delta{+}1, y+1)\left\{\gamma + \psi(\delta+y+2) - \psi(\delta+1)\right\}, \quad y \in \mathbb{N}_0, \tag{4}$$

where $\delta > 0$, $\gamma \in \mathbb{R}$ and $1+\delta\gamma > 0$.

Proposition 2.1. *The expression* (4) *is a proper pmf.*

Proof. Here, we have to show that $\sum_{y=0}^{\infty} P(y;\delta,\gamma) = 1$.

$$\begin{aligned} \sum_{y=0}^{\infty} P(y;\delta,\gamma) &= \sum_{y=0}^{\infty}\frac{\delta^2}{1+\delta\gamma}\left\{\gamma\int_0^1 p^{\delta}(1-p)^y dp - \frac{\partial}{\partial\delta}\int_0^1 p^{\delta}(1-p)^y dp\right\} \\ &= \frac{\delta^2}{1+\delta\gamma}\left\{\gamma\int_0^1 p^{\delta}\sum_{y=0}^{\infty}(1-p)^y dp - \frac{\partial}{\partial\delta}\int_0^1 p^{\delta}\sum_{y=0}^{\infty}(1-p)^y dp\right\} \\ &= \frac{\delta^2}{1+\delta\gamma}\left\{\gamma\int_0^1 p^{\delta-1}dp - \frac{\partial}{\partial\delta}\int_0^1 p^{\delta-1}dp\right\} \\ &= \frac{\delta^2}{1+\delta\gamma}\left\{\frac{\gamma}{\delta} + \frac{1}{\delta^2}\right\} = 1. \end{aligned}$$

□

Proposition 2.2. *Let $Y \sim GLL(\delta,\gamma)$. Then, the r-th factorial moment of the random variable Y is given by*

$$E(Y^{(r)}) = r!\frac{\delta^2}{1+\delta\gamma}\frac{1+\gamma(\delta-r)}{(\delta-r)^2}, \quad r = 1,2,3,\ldots. \tag{5}$$

Proof. Considering the r-th factorial moment of the geometric distribution, the r-th factorial moment of GLL distribution is given by

$$\begin{aligned} E(Y^{(r)}) &= E(E(Y^{(r)}|p)) = E(\frac{r!}{P^r}) \\ &= r!\frac{\delta^2}{1+\delta\gamma}\int_0^1 (\gamma - \log p)p^{\delta-r-1}dp \\ &= r!\frac{\delta^2}{1+\delta\gamma}\left\{\gamma\int_0^1 p^{\delta-r-1}dp - \frac{\partial}{\partial\delta}\int_0^1 p^{\delta-r-1}dp\right\} \\ &= r!\frac{\delta^2}{1+\delta\gamma}\frac{1+\gamma(\delta-r)}{(\delta-r)^2}. \end{aligned}$$

□

The mean and the second factorial moment of the random variable Y are given by using $r = 1, 2$ in (5) respectively as

$$\begin{aligned} E(Y) &= \frac{\delta^2}{1+\delta\gamma}\frac{1+\gamma(\delta-1)}{(\delta-1)^2}, \\ E(Y(Y-1)) &= 2\frac{\delta^2}{1+\delta\gamma}\frac{1+\gamma(\delta-2)}{(\delta-2)^2}. \end{aligned}$$

Using these expressions, the variance of the model is given by

$$Var(Y) = \frac{\delta^2}{1+\delta\gamma}\left\{\frac{2+2\gamma(\delta-2)}{(\delta-2)^2} + \frac{(1+\gamma(\delta-1))(\gamma\delta-\gamma\delta^2-2\delta+1)}{(1+\delta\gamma)(\delta-1)^4}\right\}.$$

Proposition 2.3. *Let $Y \sim GLL(\delta,\gamma)$. Then, the cumulative distribution function of the random variable Y is given by*

$$F(y;\delta,\gamma) = 1 - \frac{\delta^2}{1+\delta\gamma}\mathrm{B}(\delta, y+2)\left\{\gamma + \psi(\delta+y+2) - \psi(\delta)\right\}\ , \quad y \in \mathbb{N}_0.$$

Proof. The cumulative distribution function (cdf) $F(y) = P(Y \leq y)$ of the random variable Y, considering the method used in (3) is given by

$$\begin{aligned} F(y;\delta,\gamma) &= \int_0^1 P(Y \leq y|p)g(p)dp \\ &= 1 - \int_0^1 (1-p)^{y+1}g(p)dp \\ &= 1 - \frac{\delta^2}{1+\delta\gamma}\int_0^1 (\gamma - \log p)p^{\delta-1}(1-p)^{y+1}dp \\ &= 1 - \frac{\delta^2}{1+\delta\gamma}\mathrm{B}(\delta, y+2)\left\{\gamma + \psi(\delta+y+2) - \psi(\delta)\right\}\ , \quad y \in \mathbb{N}_0. \end{aligned}$$

□

Proposition 2.4. *Let $Y \sim GLL(\delta, \gamma)$. Then, the survival function of the random variable Y is given by*

$$S(y; \delta, \gamma) = \frac{\delta^2}{1+\delta\gamma} \mathrm{B}(\delta, y+1) \left\{\gamma + \psi(\delta + y + 1) - \psi(\delta)\right\} \quad , \quad y \in \mathbb{N}_0. \tag{6}$$

Proof. The survival function $S(y) = P(Y > y)$ of the random variable Y is given by

$$\begin{aligned} S(y; \delta, \gamma) &= \int_0^1 P(Y > y|p) g(p) dp \\ &= \int_0^1 (1-p)^y g(p) dp \\ &= \frac{\delta^2}{1+\delta\gamma} \int_0^1 (\gamma - \log p) p^{\delta-1} (1-p)^y dp \\ &= \frac{\delta^2}{1+\delta\gamma} \mathrm{B}(\delta, y+1) \left\{\gamma + \psi(\delta + y + 1) - \psi(\delta)\right\} \quad , \quad y \in \mathbb{N}_0. \end{aligned}$$

□

The survival function (6) indicate that for fixed δ and γ, the survival function decreases as y increases meaning the probability of observing large values reduce over time.

Proposition 2.5. *Let $Y \sim GLL(\delta, \gamma)$. Then, the hazard function of the randm variable Y is given by*

$$h(y; \delta, \gamma) = \frac{\delta}{\delta + y + 1} \frac{\gamma + \psi(\delta + y + 2) - \psi(\delta + 1)}{\gamma + \psi(\delta + y + 1) - \psi(\delta)} \quad , \quad y \in \mathbb{N}_0. \tag{7}$$

Proof. The hazard function $h(y) = \frac{P(y)}{S(y)}$ of the random variable Y is given by dividing (4) and (6) which resulted in

$$h(y; \delta, \gamma) = \frac{P(y; \delta, \gamma)}{S(y; \delta, \gamma)} = \frac{\mathrm{B}(\delta + 1, y + 1) \left\{\gamma + \psi(\delta + y + 2) - \psi(\delta + 1)\right\}}{\mathrm{B}(\delta, y + 1) \left\{\gamma + \psi(\delta + y + 1) - \psi(\delta)\right\}} \quad , \quad y \in \mathbb{N}_0.$$

After some algebraically computations, (7) will be obtained. □

The cumulative hazard function of the $GLL(\delta, \gamma)$ is given by

$$\begin{aligned} \mathrm{H}(y; \delta, \gamma) &= -\log S(y; \delta, \gamma) = -2 \log \delta - \log(1 + \delta\gamma) \\ &\quad - \log \mathrm{B}(\delta, y + 2) - \log \left\{\gamma + \psi(\delta + y + 2) - \psi(\delta)\right\} \quad , \quad y \in \mathbb{N}_0. \end{aligned}$$

Proposition 2.6. *Let $Y \sim GLL(\delta, \gamma)$. Then, the recursive formula for calculation probabilities of the GLL model satisfies*

$$P(j) = \frac{j}{\delta + j + 1} \left\{ \frac{\delta^2}{1+\delta\gamma} \frac{\mathrm{B}(\delta + 1, j)}{\delta + j + 1} + P(j-1) \right\}, \quad j = 1, 2, 3, \ldots. \tag{8}$$

Proof. The pmf of GLL for $(Y = j)$ can be rewritten as

$$\begin{aligned} P(j) &= \frac{\delta^2}{1+\delta\gamma}\mathrm{B}(\delta+1, j+1)\left\{\gamma+\psi(\delta+j+2)-\psi(\delta)\right\} \\ &= \frac{\delta^2}{1+\delta\gamma}\frac{j}{\delta+j+1}\mathrm{B}(\delta+1, j)\left\{\gamma+\frac{1}{\delta+j+1}+\psi(\delta+j+1)-\psi(\delta)\right\} \\ &= \frac{j}{\delta+j+1}\left\{\frac{\delta^2}{1+\delta\gamma}\frac{\mathrm{B}(\delta+1,j)}{\delta+j+1}+P(j-1)\right\}, \quad j=1,2,3,\ldots, \end{aligned}$$

with starting value $P(0) = \frac{\delta^2}{(1+\delta)^2}\frac{1+\gamma+\delta\gamma}{1+\delta\gamma}$. □

2.2 Zero-inflated and zero-truncated GLL distribution

In certain contexts, count data exhibit an excessive number of zero outcomes compared to what is expected in the Poisson model. For example , the proportion of zero claims in motor insurance data may contain a large number of zeros due to the small number of accidents among drivers and to the lack of reports for minor claims (Yip and Yau, 2005). Both zero-inflated and zero-truncated models are useful for modeling count data. Zero-inflated models, such as the zero-inflated Poisson and zero-inflated negative binomial (Ridout et al., 2001) zero-inflated generalized Poisson (Zamani and Ismail, 2014; Famoye and Singh, 2006), have been widely applied to analyze zero-inflated count data. Zero-truncated count data, conversely, are characterized by the complete absence of zero outcomes in the dataset. This type of data typically arises when zeros are not possible or observable for a given count variable.

Zero-inflated GLL

The zero-inflated version of a probability function $P(y;\theta)$ is defined by

$$P(y;\theta) = \big[\omega + (1-\omega)P(0;\theta)\big]I_{\{y=0\}} + \big[(1-\omega)P(y;\theta)\big]I_{\{y>0\}}$$

Considering $\psi(\delta+2) - \psi(\delta+1) = \frac{1}{\delta+1}$, the pmf of zero-inflated GLL (ZIGLL) which we denote by $P^{(0)}(y;\delta,\gamma)$ is given by

$$P^{(0)}(y;\delta,\gamma) = \begin{cases} \omega + (1-\omega)\frac{\delta^2}{1+\delta\gamma}\frac{1}{\delta+1}\left\{\gamma+\frac{1}{\delta+1}\right\}, & y=0, \\ (1-\omega)\,P(y;\delta,\gamma)\ , & y=1,2,3,\ldots \end{cases}$$

where $0 \le \omega < 1$ and $P(y;\delta,\gamma)$ is the pmf of GLL given in (4). The $P^{(0)}(y;\delta,\gamma)$ reduces to $P(y;\delta,\gamma)$ when $\omega = 0$.

Zero-truncated GLL

The zero-truncated GLL (ZTGLL) distribution can be obtained by replacing the zero-based geometric with the traditional geometric distribution in (3). Following the same the methodology applien in (3), the pmf of ZTGLL, denoted by $P^{(1)}(y;\delta,\gamma)$ is given by

$$P^{(1)}(y;\delta,\gamma)=\frac{\delta^2}{1+\delta\gamma}\mathrm{B}(\delta{+}1,y)\left\{\gamma+\psi(\delta+y+1)-\psi(\delta+1)\right\},\qquad y=1,2,3,\ldots.$$

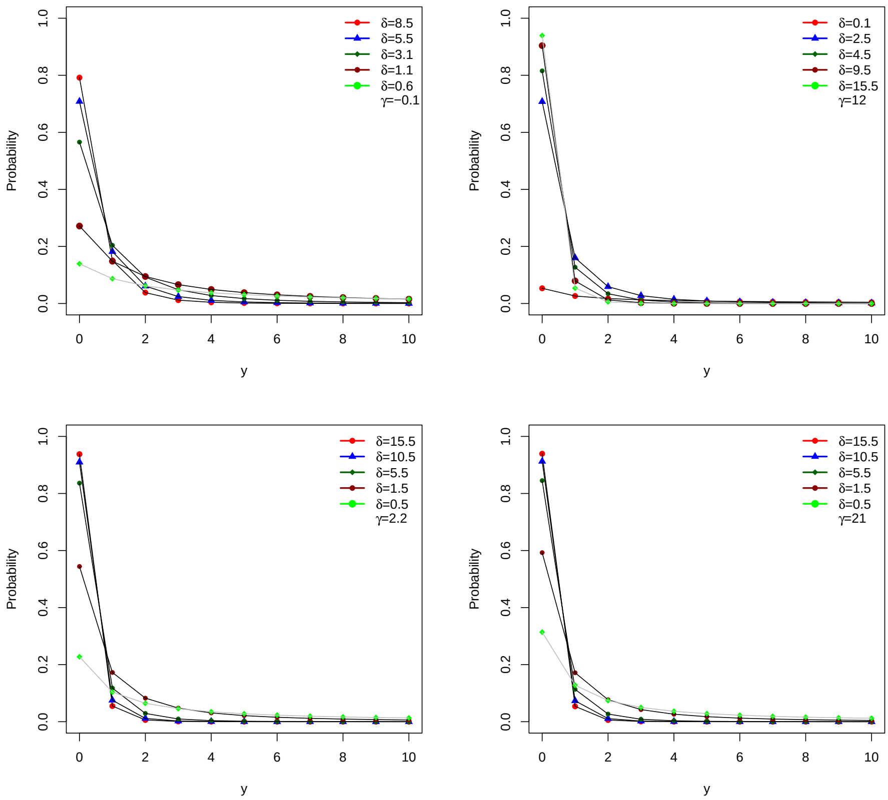

Figure 1: Probability plot of the GLL distribution for several selected values of parameters.

Figure 1 illustrates the behavior of the probability mass function of the GLL distribution for several selected values of (δ,γ). As can be concluded from the pmf of the GLL model and the behavior of the probability function indicated in Figure 1, Increasing δ concentrates more probability mass at $Y=0$, resulting in stronger zero-inflation. In plots with large δ values, the bar at $Y=0$, is much higher and the probability mass function decays rapidly for $Y\geq 1$. Conversely, small δ values shift mass away from zero and produce a heavier right tail. Thus, δ controls the decay rate of the distribution, with smaller δ corresponding to a greater probability of larger counts.

On the other hand, increasing γ shifts probability from zero toward nonzero counts and lengthens the tail; larger γ values yield relatively higher probabilities for larger Y, producing a heavier tail. A negative γ reduces tail weight and skewness, concentrating more probability near zero-though the effect of γ interacts with δ. Intuitively, γ modulates tail behavior and skewness: positive γ increases the chance of extreme counts, while negative γ reduces it.

There is also an observable interaction between δ and γ. For a fixed γ, varying δ reallocates mass between zero and the tail as described: a large δ combined with a negative or small γ produces a sharply peaked distribution at zero, whereas a small δ with a positive γ yields a heavy-tailed distribution.

As with real-world data, for example insurance claim counts, δ can be interpreted as a measure of risk intensity where a smaller δ suggests a higher chance of observing large number of claims. Regarding γ, it can be interpreted a measure of variability in risk, specifically, positive values of γ indicate that the data have a longer tail, meaning there is a higher probability of extreme events. For general count data, δ can be considered a measure of the decay rate, such that a small δ corresponds to slower decay, making large observations are more likely. Conversely, γ can be interpreted as a measure of skewness or tail behavior, with positive and negative values indicating longer and shorter tails, respectively. In summary, a small δ and positive γ indicate that the data are over-dispersed with a longer tail while a large δ and a negative γ lead to a less over-dispersed distribution with a shorter tail.

3 Model estimation

In this section, we discuss different methods of GLL parameters estimation.

3.1 Method of moments

Given the sample $y_1, \ldots, y_n$ of size n from the GLL distribution the moment estimates $\tilde{\delta}$ and $\tilde{\gamma}$ of δ and γ considering (5) are given through the solution of equations

$$\begin{cases} \frac{\delta^2}{1+\delta\gamma}\frac{1+\gamma(\delta-1)}{(\delta-1)^2} = \bar{x}, \\ 2\frac{\delta^2}{1+\delta\gamma}\frac{1+\gamma(\delta-2)}{(\delta-2)^2} = \bar{x^2} + \bar{x}, \end{cases}$$

solving both equations for γ leads to

$$\begin{cases} \gamma = \frac{\bar{x}(\delta-1)^2-\delta^2}{\delta^2(\delta-1)-\bar{x}(\delta-1)^2}, \\ \gamma = \frac{(\bar{x}+\bar{x^2})(\delta-2)^2-2\delta^2}{2\delta^2(\delta-2)-(\bar{x}+\bar{x^2})(\delta-2)^2}. \end{cases} \tag{9}$$

Now, to estimate δ, we can set the equations equal to one another. Given the complexity of the resulting expression, the resulting equation is typically solved numerically. Therefore, a numerical solution of (9) yields $\tilde{\delta}$; substituting this value into one of the equations in (9) gives $\tilde{\gamma}$.

3.2 Maximum likelihood estimation

In this part we consider the maximum likelihood estimators of GLL parameters. Suppose $y_1, \ldots, y_n$ is a random sample of size n from the GLL distribution with pmf given in (4). The likelihood function can be expressed as

$$\ell(\delta, \gamma) = 2n \log \delta - n \log(1 + \delta\gamma) + \sum_{i=1}^{n} \log \Gamma(y_i + 1) + n \log \Gamma(\delta + 1)$$

$$-\sum_{i=1}^{n}\log\Gamma(\delta+y_i+2)-\sum_{i=1}^{n}\log\left\{\gamma+\psi(\delta+1)-\psi(\delta+y_i+2)\right\}.$$

Therefore, the normal equations are given by

$$\begin{aligned}\frac{\partial}{\partial\delta}\ell(\delta,\gamma) &= \frac{2n}{\delta}-\frac{n\gamma}{1+\delta\gamma}+n\,\psi(\delta+1)-\sum_{i=1}^{n}\psi(\delta+y_i+2)\\ &\quad-\sum_{i=1}^{n}\frac{\psi'(\delta+1)-\psi'(\delta+y_i+2)}{\gamma+\psi(\delta+1)-\psi(\delta+y_i+2)}=0,\\ \frac{\partial}{\partial\gamma}\ell(\delta,\gamma) &= \frac{n\delta}{1+\delta\gamma}-\sum_{i=1}^{n}\frac{1}{\gamma+\psi(\delta+1)-\psi(\delta+y_i+2)}=0.\end{aligned}$$

These non-linear equations have no closed form solutions and must be solved numerically using methods such as the Newton-Raphson method or an EM-algorithm, using the method-of-moments estimates of δ and γ as initial values for iteration. In this study we used the `optim()` function in `R` extended by R Core Team (2014) for optimization purpose. The `optim` function provides various optimization algorithms, such as Nelder-Mead, BFGS, CG (conjugate gradient), and L-BFGS-B. Among these, the BFGS algorithm works well for smooth functions, and L-BFGS-B performs better when there are constraints on the parameters. Here we used the L-BFGS-B algorithm to optimize the loglikelihood function because of its suitability for smooth functions and the presence of constraints on the parameters. Moreover, the method-of-moments estimates were used as the starting values in the `optim`() function.

Furthermore, the log-likelihood function is non-linear, involves complex terms, and is non-convex in the parameters δ and γ, implying it may have multiple local maxima. Therefore, to address potential convergence issues, several methods can be employed. The first approach is to add a penalty term to the log-likelihood function to discourage extreme parameter values. Adding such a penalty term leads to a penalized likelihood function, which can be defined as

$$\ell_p(\delta,\gamma)=\ell(\delta,\gamma)-\lambda(\delta^2+\gamma^2),$$

where λ is the regularization parameter. The second approach is to transfer the constrained parameters into an unconstrained space. For example, we can set $\delta=e^{\alpha}$ where $\alpha\in\mathbb{R}$, which ensures that $\delta>0$. As another method, we may use the profile likelihood function for optimization purpose.

In this study, the method of moments (MOM) estimates $(\tilde{\delta},\tilde{\gamma})$ were used as the starting values for the numerical optimization. This choice is supported by two key considerations. First, MOM estimates provide consistent, data driven starting points that are typically in the neighborhood of the true parameters, thereby promoting stable convergence. Second, and most critically, the simulation study in Section 4 where this exact procedure (MOM initials followed by MLE optimization) demonstrates its practical reliability. The consistent reduction in bias and RMSE, along with the convergence of coverage probabilities to their nominal levels as the sample size increases (see Tables 1 and 2), provides strong empirical evidence that the MLE procedure initialized with MOM estimates is numerically stable and produces valid inferences for the GLL distribution.

3.3 Asymptotic normality of maximum likelihood estimators

The maximum likelihood estimator is asymptotically consistent and normally distributed as the sample size increases. For $\hat{\Theta} = (\hat{\delta}, \hat{\gamma})^{\mathrm{T}}$, which is the maximum likelihood estimator of $\Theta = (\delta, \gamma)^{\mathrm{T}}$, as n tends to ∞, the sampling distribution of $(\hat{\delta}, \hat{\gamma})$ is asymptotically multivariate normal, that is, $\hat{\Theta} \sim N_2(\Theta, \boldsymbol{I}^{-1}(\Theta)$ where $\boldsymbol{I}(\Theta)$ is the Fisher information matrix computed as

$$\boldsymbol{I}(\Theta) = \begin{bmatrix} \boldsymbol{I}_{\boldsymbol{\delta\delta}} & \boldsymbol{I}_{\boldsymbol{\delta}\gamma} \\ \boldsymbol{I}_{\gamma\boldsymbol{\delta}} & \boldsymbol{I}_{\beta\gamma} \end{bmatrix} = E[-\frac{\partial}{\partial \Theta} U_{\Theta}(\delta, \gamma)],$$

and $U_{\Theta}(\delta, \gamma)$ is the score vector that is given by

$$U_{\Theta}(\delta, \gamma) = \Big(\frac{\partial}{\partial \delta} \ell(\delta, \gamma), \frac{\partial}{\partial \gamma} \ell(\delta, \gamma) \Big).$$

Since, $\boldsymbol{I}(\Theta)$ has no closed form for GLL distribution, we used the estimated observed Fisher information matrix,

$$\boldsymbol{I}(\hat{\Theta}) = -\frac{\partial}{\partial \Theta} U_{\Theta}(\delta, \gamma)|_{\Theta = \hat{\Theta}},$$

is used to estimate $\boldsymbol{I}(\Theta)$. The observed information matrix of $\hat{\Theta}$ is given in the appendix. Therefore, the asymptotic $100(1-\eta)\%$ confidence interval for θ_j is given by

$$\theta_j : \hat{\theta}_j \pm z_{\eta/2} se(\hat{\theta}_j), \quad j = 1, 2.$$

4 Simulation study

In this section, we conduct a Monte Carlo simulation to study the finite sampe behavior of maximum likelihood estimates based on the finite sample sizes for the parameters of the GLL distribution. For sample generation from the GLL distribution, we used the pmf of GLL (4) and the weighted random sampling techniques Motwani (1995). The simulation study is carried out N=10,000 times for each triple (δ, γ) and n=50, 100, 200, 300, 500, $\gamma = 0.05$, -0.1, 0.1, 0.3 and δ=1.5, 3.5. For evaluation purposes, we used the average bias ($Bias$), root mean square error ($RMSE$), and coverage probability (CP) criteria, which are defined as

$$\begin{aligned} Bias(\hat{\theta}) &= \frac{1}{N} \sum_{i=1}^{N} (\hat{\theta}_i - \theta), \\ RMSE(\hat{\theta}) &= \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\hat{\theta}_i - \theta\right)^2}, \\ CP_{\theta}(n) &= \frac{1}{N} \sum_{i=1}^{N} I\left\{ \hat{\theta}_i - 1.96\, \hat{s}_{\hat{\theta}_i} < \theta < \hat{\theta}_i + 1.96\, \hat{s}_{\hat{\theta}_i} \right\}. \end{aligned}$$

Table 1 shows that the average bias for maximum likelihood estimators of $\hat{\delta}$, $\hat{\gamma}$, can be either positive or negative. The average bias decreases as the sample size n increases. The table also presents the root mean square error (RMSE) of the estimators, which similarly decreases with larger sample sizes. Moreover, the table reports the coverage probability at the 90% and 95% confidence levels. The results indicate that the coverage probability approaches the intended confidence level as the sample size n increases. This is may be due to the overestimation of the standard error in small samples or to deviation from normality in the parameters' distribution for small samples. Additionally, for small samples, the parameter estimates may frequently approach the boundary region ($\delta > 0, 1 + \delta\gamma > 0$), resulting in skewed distributions and wider confidence intervals. The simulation results presented in Figure 2 and Tables 1 and 2 can be summarized as follows:

a) The highest bias and RMSE for the parameter estimates occur for $\delta = 3.5$ and $\gamma = 0.3$ at $n = 50$, while the lowest values occur for $\delta = 3.5$ and $\gamma = -0.1$ at $n = 500$.

b) Both the bias and RMSE of the parameter estimates decrease as the sample size increases.

c) The coverage probabilities converge toward the nominal level 95 % as n increases.

Table 1: Estimated bias, RMSE, and coverage probability for γ and $\delta = 1.5$.

γ		Bias		RMSE		CP 90%		CP 95%	
	n	δ	γ	δ	γ	δ	γ	δ	γ
0.05	50	1.360	-0.207	1.340	0.390	0.914	0.976	0.955	0.965
	100	0.968	-0.192	0.788	0.376	0.906	0.975	0.953	0.980
	200	0.515	-0.176	0.224	0.152	0.903	0.940	0.951	0.955
	300	0.249	-0.063	0.201	0.124	0.904	0.937	0.947	0.954
	500	0.143	-0.056	0.176	0.100	0.897	0.934	0.949	0.961
0.1	50	1.385	-0.207	1.481	0.227	0.929	0.991	0.956	0.958
	100	0.950	0.156	1.303	0.207	0.908	0.977	0.951	0.984
	200	0.748	-0.053	0.830	0.191	0.905	0.944	0.949	0.959
	300	0.270	-0.051	0.205	0.147	0.903	0.935	0.953	0.956
	500	0.147	0.049	0.171	0.107	0.901	0.922	0.951	0.952
0.3	50	1.229	-0.211	1.478	0.338	0.920	0.984	0.957	0.987
	100	0.887	-0.144	0.850	0.335	0.910	0.966	0.952	0.968
	200	0.772	-0.143	0.297	0.330	0.907	0.965	0.952	0.966
	300	0.316	-0.127	0.226	0.345	0.903	0.965	0.952	0.979
	500	0.155	-0.126	0.196	0.203	0.897	0.93	0.951	0.954
-0.1	50	1.175	0.115	1.352	0.480	0.912	0.963	0.953	0.999
	100	0.615	-0.109	0.756	0.132	0.903	0.945	0.952	0.973
	200	0.238	-0.121	0.299	0.078	0.901	0.934	0.951	0.963
	300	0.140	-0.029	0.181	0.063	0.901	0.925	0.951	0.957
	500	0.136	-0.021	0.162	0.052	0.901	0.919	0.950	0.955

5 Application to real data

5.1 Automobile insurance data

The data in Table 3 were used to study the performance of the GLL distribution compared to alternative models. The data, collected by Tröbliger, consists of 23,589 automobile insurance records that report the number of accidents per driver over a one-year period (Klugman et al., 2012). We also fitted to the Poisson and negative binomial distributions as alternatives to the GLL model. The negative binomial distribution considered here, introduced by Greene (2008), is obtained by mixing a Poisson

Table 2: Estimated bias, RMSE, and coverage probability for γ and $\delta = 3.5$.

γ		Bias		RMSE		CP 90%		CP 95%	
	n	δ	γ	δ	γ	δ	γ	δ	γ
0.05	50	1.337	0.012	2.518	0.287	0.947	0.979	0.961	0.970
	100	0.976	-0.012	0.288	0.193	0.906	0.975	0.954	0.976
	200	0.824	-0.046	0.224	0.152	0.903	0.940	0.951	0.955
	300	0.743	-0.003	0.201	0.124	0.904	0.937	0.947	0.954
	500	0.686	-0.002	0.825	0.087	0.903	0.925	0.949	0.954
0.1	50	1.525	-0.213	2.860	0.313	0.948	0.937	0.963	0.980
	100	1.002	-0.156	1.303	0.297	0.928	0.977	0.957	0.968
	200	0.927	-0.098	0.988	0.191	0.905	0.944	0.952	0.963
	300	0.873	-0.086	0.903	0.147	0.903	0.935	0.953	0.961
	500	0.712	-0.078	0.866	0.119	0.901	0.922	0.949	0.958
0.3	50	2.300	-0.160	4.547	0.402	0.955	0.926	0.967	0.964
	100	0.694	0.002	1.063	0.045	0.921	0.948	0.955	0.960
	200	0.628	-0.004	0.817	0.024	0.913	0.936	0.954	0.957
	300	0.578	0.003	0.728	0.210	0.910	0.929	0.954	0.953
	500	0.496	-0.003	0.677	0.014	0.903	0.925	0.951	0.953
-0.1	50	0.857	0.018	1.534	0.096	0.933	0.955	0.956	0.965
	100	0.149	0.037	0.796	0.198	0.903	0.945	0.952	0.973
	200	0.070	0.012	0.537	0.051	0.901	0.934	0.954	0.960
	300	0.042	0.007	0.425	0.035	0.907	0.940	0.952	0.953
	500	0.024	0.004	0.323	0.025	0.905	0.925	0.952	0.959

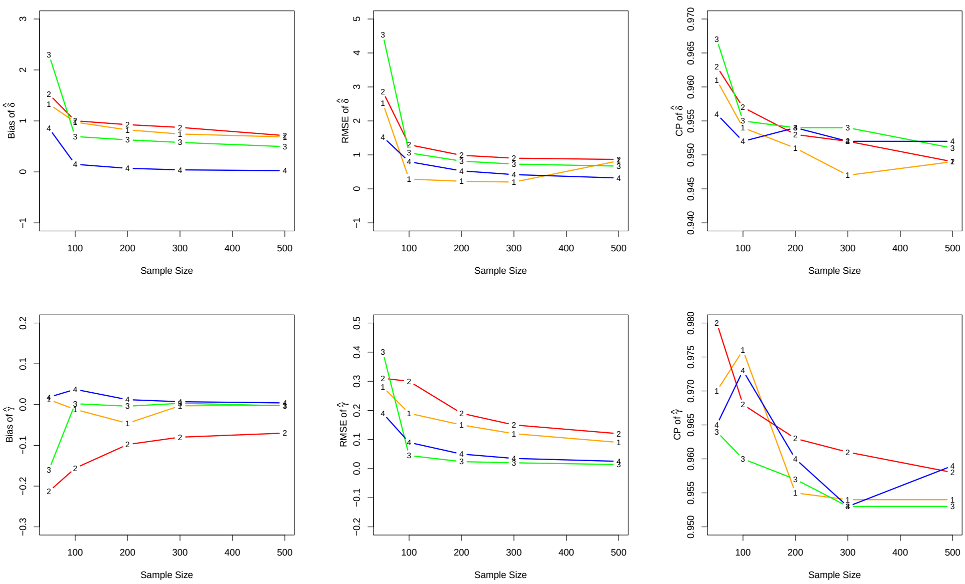

Figure 2: Top: estimated bias, RMSE and CP of $\hat{\delta}$. Bottom: estimated bias, RMSE and CP of $\hat{\gamma}$ for Table 2: $\delta = 3.5$, 1: $\gamma = 0.05$, 2: $\gamma = 0.1$, 3: $\gamma = 0.3$ and 4: $\gamma = -0.1$ values.

distribution with a gamma distribution and has the probability mass function (pmf)

$$P(y; r, \beta) = \binom{y + r - 1}{y} \left(\frac{1}{1+\beta}\right)^r \left(\frac{\beta}{1+\beta}\right)^y, \quad y = 0, 1, 2, \ldots, \quad r > 0, \quad \beta > 0.$$

Based on the chi-square p-value and the log likelihood, the negative binomial and the GLL distributions represent a significant improvement over the Poisson distribution. However, a direct comparison between the negative binomial and the GLL distributions indicates that the GLL model provides a beeter fit than the negative binomial distribution.

Table 3: Fitted Poisson, negative binomial and GLL to automobile insurance data.

No. of accidents	Frequency	Poisson	Negative binomial	GLL
0	20592	20420.9	20596.8	20592.9
1	2651	2945.1	2631	2648.8
2	297	212.4	318.4	299.6
3	41	10.2	37.8	40.3
4	7	0.4	4.4	6.4
5	0	0	0.5	1.3
6+	1	0	0.1	0.7
parameters		λ=0.144	β=0.129	δ=27.280
			r=1.117	γ=-0.023
Chi-squares		196.45	3.66	0.039
p-values		<0.01	0.3	0.99
-Log-likelihood		10297.84	10223.43	10221.42
AIC		20597.68	20450.86	20446.84
BIC		20605.75	20467.00	20462.98

5.2 Accident profile data

The data in Table 4 represent 9,461 automobile insurance policies (Klugman et al., 2012; Tremblay, 1992), for which the number of accidents per policy was recorded. The Poisson, negative binomial, and GLL models were fitted to the data, and the results are presented in Table 4. The results indicate that only the GLL distribution provides an adequate fit at the 95% confidence level.

Table 4: Fitted Poisson, negative binomial and GLL to accident profile data.

No. of accidents	Frequency	Poisson	Negative binomial	GLL
0	7840	7638.4	7850.4	7840.9
1	1317	1634.5	1286.2	1311.2
2	239	174.9	255.7	238.1
3	42	12.5	53.8	51.9
4	14	0.6	11.6	13.2
5	4	0.1	2.7	3.8
6	4	0	0.5	1.2
7+	1	0	0.1	0.7
parameters		λ=0.214	β=0.305	δ=14.67
			r=0.701	γ=-0.030
Chi-squares		294.4	8.85	4.32
p-values		<0.01	0.03	0.23
-Log-likelihood		5490.78	5348.04	5343.14
AIC		10983.56	10700.08	10690.28
BIC		10990.71	10714.39	10704.59

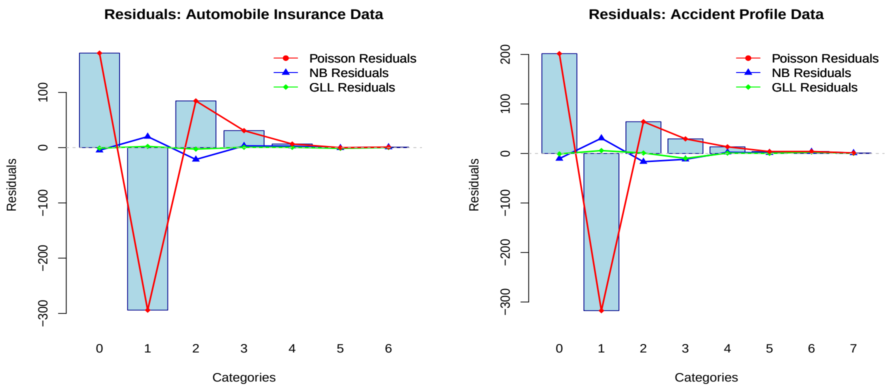

Figure 3: Residual plot comparing observed frequencies and fitted values from Poisson, NB and GLL, Top: Automobile Insurance Data. Bottom: Accident Profile Data.

Figure 3 presents the residual plots, constructed as the difference between observed and predicted frequencies. These plots were used to assess model fit. For both datasets, the residuals of the GLL model were clustered tightly around zero, indicating a close match to the observed data. In contrast, the negative binomial (NB) model residuals showed moderate deviation, while the Poisson residuals exhibited substantial and systematic departure from the zero line, confirming its inadequacy for these over-dispersed counts.

6 Conclusion

In this paper, we introduce the geometric log-Lindley distribution, a mixed distribution obtained by combining the geometric and log-Lindley distributions. The GLL random variable is defined on $\mathbb{N}_0$ and has a closed-form pmf, making it an appropriate choice for modeling over-dispersed count data. We present the statistical properties of the proposed model, including the probability mass function, survival and hazard rate functions, moments, and a recursive formula for computing probabilities. Additionally, the zero-inflated and zero-truncated versions of the GLL distribution are discussed, extending its applicability to a wider range of scenarios.

The model parameters are estimated using the method of MOM and MLE. A simulation study was conducted to evaluate the performance of the estimators based on several criteria, including average bias, RMSE, and coverage probability. The results demonstrate that the MLEs are consistent and that the coverage probability approaches the intended level as the sample size increases. However, for small sample sizes, the coverage probability was higher than the nominal level, which may be attributed to the overestimation of standard errors and to finite-sample and boundary effects. These findings highlight the importance of using large sample sizes for reliable inference with the GLL model.

The performance of the GLL distribution was evaluated using two real insurance datasets. The results indicate that the GLL model provides a better fit than traditional models such as the Poisson and negative binomial distributions, as evidenced by higher log-likelihood values and lower chi-square statistics. This suggests that the GLL distribution is a suitable tool for modeling over-dispersed count data in insurance and other fields.

Beyond insurance, the GLL mode can be applied for many count data where observations contain excess zeros and heavy tails. In biomedical research, examples include recurrent event data such as hospital readmissions, infection episodes, the number of doctor visits, and rare adverse events or in traffic and transportation, counts from sensors or crash data often show similar features. Methodologically, straightforward extensions include GLL regression by including covariates through linking δ or γ, zero-inflated or hurdle versions to explicitly separate structural zeros, multivariate or hierarchical GLL models for clustered/repeated measures, and spatio-temporal GLL processes for dependent count data. These extensions would expand the model's applicability and allow direct modeling of covariate effects, random effects, and dependence.

References

Boucher, J.P., Denuit, M. and Guillen, M. (2007). Risk classification for claim count: a comparative analysis of various zero-inflated mixed Poisson and hurdle models. *North American Actuarial Journal*, **11**(4):110–131.

Denuit, M., Marechal, X., Pitrebois, S. and Walhin, J.F. (2007). *Actuarial Modeling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems*. West Sussex, UK: John Wiley and Sons.

Famoye, F. and Singh, K.P. (2006). Zero-inflated generalized Poisson regression model with an application to domestic violence data. *Journal of Data Science*, **4**(1):117–130.

Gómez-Déniz, E., Sordo, M.A. and Calderín-Ojeda, E. (2014). The Log–Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics*, **54**:49–57.

Greene, W. (2008). Functional forms for the negative binomial model for count data. *Economics Letters*, **99**(3):585–590.

Ismail, N. and Zamani, H. (2013). Estimation of claim count data using negative binomial, generalized Poisson, zero-inflated negative binomial and zero-inflated generalized Poisson regression models. *Casualty Actuarial Society E-Forum*, **41**(20), 1–28.

Izsák, R. (2008). Maximum likelihood fitting of the Poisson lognormal distribution. *Environmental and Ecological Statistics*, **15**(2):143–156.

Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2012). *Loss Models: From Data to Decisions*. John Wiley & Sons.

Meng, S., Wei, Y. and Whitmore, G.A. (1999). Accounting for individual overdispersion in a bonus-malus system. *ASTIN Bulletin: The Journal of the International Actuarial Association*, **29**(2):327–337.

Motwani, R. (1995). *Randomized Algorithms.* Cambridge: Cambridge University Press.

Navarro, D. and Perfors, A. (2005). *An Introduction to the Beta-Binomial Model.* University of Adelaide.

R Core Team (2014). *R: A Language and Environment for Statistical Computing.* Austria, Vienna: R Foundation for Statistical Computing.

Ridout, M.S., Hinde, J.P. and Demétrio, C.G.B. (2001). A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics*, **57**(1):219–223.

Tremblay, L. (1992). Using the Poisson inverse Gaussian in bonus-malus systems. *ASTIN Bulletin: The Journal of the IAA*, **22**(1):97–106.

Weinberg, C.R. and Gladen, B.C. (1986). The beta-geometric distribution applied to comparative fecundability studies. *Biometrics*, **42**:547–560.

Willmot, G.E. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial. *Scandinavian Actuarial Journal*, **1987**(2):113–127.

Wongrin, W. and Bodhisuwan, W. (2017). Generalized Poisson–Lindley linear model for count data. *Journal of Applied Statistics*, **44**(15):2659–2671.

Yip, K.C.H. and Yau, K.K.W. (2005). On modeling claim frequency data in general insurance with extra zeros. *Insurance: Mathematics and Economics*, **36**(2):153–163.

Zakerzadeh, H. and Dolati, A. (2009). Generalized Lindley distribution. *Journal of Mathematical Extension*, **3**(2):13–25.

Zamani, H. and Ismail, N. (2010). Negative binomial-Lindley distribution and its application. *Journal of Mathematics and Statistics*, **6**(1):4–9.

Zamani, H. and Ismail, N. (2014). Functional form for the zero-inflated generalized Poisson regression model. *Communications in Statistics-Theory and Methods*, **43**(3):515–529.

Appendix. The observed information matrix

The observed information matrix is given by replacing δ , γ with the $\hat{\delta}$ and $\hat{\gamma}$ in the following expressions.

$$\frac{\partial^2}{\partial\delta^2}\ell(\delta,\gamma) = \frac{-2n}{\delta^2} + \frac{n\gamma^2}{(1+\delta\gamma)^2} + n\psi'(\delta+1) - \sum_{i=1}^{n}\psi'(\delta+y_i+2) - \sum_{i=1}^{n}\frac{A'B-A^2}{B^2},$$

$$\frac{\partial^2}{\partial\gamma^2}\ell(\delta,\gamma) = \frac{-n\delta^2}{(1+\delta\gamma)^2} + \sum_{i=1}^{n}\frac{1}{B^2},$$

$$\frac{\partial^2}{\partial\delta\partial\gamma}\ell(\delta,\gamma) = \frac{n}{(1+\delta\gamma)^2} - \sum_{i=1}^{n}\frac{A}{B^2},$$

where $A = \psi'(\delta+1) - \psi'(\delta+y_i+2)$ and $B = \gamma + \psi(\delta+1) - \psi(\delta+y_i+2)$.