Science Journal of Applied Mathematics and Statistics
2026, Vol. 14, No. 4, pp. 90–105
https://doi.org/10.11648/j.sjams.20261404.11

Research Article

# A Generalized Poisson–Gamma Hybrid Model for Heavy-Tailed Count Data

**Isiak Kamaldeen Olomoda[1, *], Musa Yunus Olatunji[2]**

[1]Department of Mathematics, Nigerian Army University Biu, Biu, Nigeria
[2]Department of Statistics, Kwara State Polytechnic Ilorin, Ilorin, Nigeria

## Abstract

In applications including insurance claim counts, epidemiology, reliability analysis, finance, and network traffic modeling, count data with over dispersion and heavy-tailed behavior are common. Because they impose exponentially decaying tails that underestimate extreme count probabilities, classical Poisson and Poisson–Gamma (Negative Binomial) models are frequently insufficient in such situations. In this study, we propose a generalized Poisson–Gamma hybrid distribution obtained by compounding a Poisson distribution with a power–Gamma mixing law. The proposed model extends the classical Poisson–Gamma framework by introducing an additional shape parameter that governs tail thickness and induces greater dispersion. An explicit infinite-series representation of the probability mass function is derived, and fundamental distributional properties are investigated. It was shown that the Negative Binomial distribution arises as a special case, ensuring model coherence. Furthermore, the proposed model exhibits heavier-than-exponential tails, and under suitable parameter regimes, its tail probabilities display polynomial decay, placing the distribution within the class of heavy-tailed count models. Estimation and inferential aspects were discussed and applications to simulated count data and Monte Carlo simulated count data were also discussed. Heavy-tailed count data demonstrate superior performance compared to classical alternatives.

## Keywords

Negative Binomial, Power-Gamma Mixing, Poisson–Gamma hybrid, Heavy-Tailed, Count Data, Classical Poisson, Over – dispersion

## 1. Introduction

In many different fields, such as actuarial science, public health, telecommunications, transportation, and finance, count data models are essential to statistical analysis. Because of its ease of use and straightforward probabilistic interpretation, the Poisson distribution continues to be the fundamental model for count data. The Poisson model's equidispersion feature, which limits the variance to match the mean, is a well-known drawback. This assumption is often broken by empirical count data, which show excess variability, over dispersion, and, in many real-world situations, heavy-tailed behavior with a non-negligible likelihood of extreme counts.

Heavy-tailed count data naturally arise in many applied disciplines, [5, 7, 23]. A plethora of family of distributions has been proposed to model heavy-tailed count data, but the use

*Correspondence: Isiak Kamaldeen Olomoda (isiaqkamaldeen3@gmail.com)

**Received:** 2 March 2026; **Accepted:** 16 March 2026; **Published:** 27 July 2026

of some of them is inhibited by the lack of an explicit, or easily computable, expression for their probability mass function (p.m.f) and by the lack of any – order moments for all, or some parameters values. Inspired by these ideas, a generalized Poisson–Gamma hybrid model based on a power–Gamma mixing distribution is presented in this study. Under specific parameter combinations, the resulting count distribution can display regularly changing of heavy-tailed behavior which is to include additional shape parameter for the suggested model, which offers direct control over tail thickness. In order to maintain consistency with current theory and make understanding easier, the Negative Binomial distribution and the traditional Poisson-Gamma model are recovered as a special case.

Adding heterogeneity using mixed Poisson models is a traditional solution to over dispersion. Perhaps the most popular of these is the Poisson–Gamma mixture that results in the Negative Binomial distribution. The Negative Binomial model captures minor over dispersion and accounts for variance inflation, but its tail probabilities still decay exponentially. As a result, when data show long-tailed or heavy-tailed characteristics—a condition commonly seen in insurance claim frequencies, epidemic outbreaks, network congestion counts, and financial event data—it may significantly underestimate the incidence of extreme events [1-4, 8, 12-14].

Theoretically, regular variation, polynomial tail decay, and subexponentiality are frequently used to characterize heavy-tailed distributions, which have tails that decay more slowly than exponential distributions [6, 9, 10, 11, 15, 16]. Heavy-tailed models are especially crucial for risk assessment and extreme-value analysis in the context of count data, as underestimating tail probabilities can have major practical repercussions [17-19]. Developing flexible count distributions that can display heavier tails than traditional Poisson-based models has become increasingly popular as a result.

Using extended mixing distributions in the Poisson framework is a logical and conceptually interesting method for building heavy-tailed count models. Appropriate modifications of the Gamma mixing law can produce heavier-tailed marginal count distributions because the tail behavior of a mixed Poisson distribution is highly impacted by the tail features of its mixing distribution. Specifically, marginal count models with polynomially decaying tails and improved dispersion features can result from mixing distributions that deviate from the exponential decay structure of the Gamma rule [20-22, 24, 25].

This work makes three primary contributions. The probability mass function of the suggested distribution is first derived in an explicit infinite-series form, and its connection to well-known mixed Poisson models is established. Second, we examine its distributional characteristics, such as moments, dispersion, and limiting situations. We also talk about the circumstances in which the model displays heavy-tailed behavior, which is defined as slower-than-exponential tail decay. Third, we discuss parameter estimates and show the model's usefulness in real count data sets with heavy-tailed features, where the suggested model performs better than the traditional Poisson and Negative Binomial alternatives.

This is how the rest of the paper is structured. The model formulation and the probability mass function are derived in Section 2. Moment characteristics and tail behavior, especially heavy-tail consequences, are examined in Section 3. Inference and estimation are covered in Section 4. Applications to actual heavy-tailed count data are given in Section 5, and potential expansions and future research paths are discussed in Section 6.

## 2. Methodology

This part involves the compunding a Poisson distribution with a power–Gamma mixing law. The steps involve in achieving the set propose distribution are as follow:

Step 1: Identify Model definition for Poisson;

Step 2: Define Generalized Gamma using Power – Gamma mixing distribution;

Step 3: Obtain the unconditional pmf by compounding through the marginal pmf;

Step 4: Employing power series expansion for the term $e^{-\beta\lambda^q}$;

Step 5: Evalute the gamma function using power series;

Step 6: Normalize to get the new pmf i.e $\sum_{x=0}^{\infty} f(x) = 1$

*Step 1*

Let

$$X \sim Poisson(\lambda)$$

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \ldots$$

*Step 2*

Generalized Gamma using Power – Gamma mixing distribution

Let

$$\lambda \sim GGamma(\alpha, \beta, q)$$

$$f(\lambda) = \frac{q\beta^\alpha}{\Gamma(\alpha)}\lambda^{q\alpha-1}e^{-\beta\lambda^q}, \lambda > 0$$

When,

$q = 1$, this reduces to the standard gamma, hence the classical Poisson – Gamma model.

*Step 3*

Obtain the unconditional pmf by compounding through the marginal pmf.

$$f(X = x) = \int_0^\infty f(x/\lambda) f(\lambda) d\lambda$$

$$f(X = x) = \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}\frac{q\beta^\alpha}{\Gamma(\alpha)}\lambda^{q\alpha-1}e^{-\beta\lambda^q} d\lambda$$

$$f(X=x)=\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}\int_0^{\infty}\lambda^{x\ +q\alpha\ -\ 1}e^{-\lambda-\beta\lambda^{q}}d\lambda$$

*Step 4*
Using power series expansion

$$f(X=x)=\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}\int_0^{\infty}\lambda^{x\ +q\alpha\ -\ 1}e^{-\lambda}e^{-\beta\lambda^{q}}d\lambda$$

$$e^{-\beta\lambda^{q}}=\sum_{m=0}^{\infty}\frac{(-\beta)^m}{m!}\lambda^{qm}$$

$$f(X=x)=\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}\int_0^{\infty}\lambda^{x\ +q\alpha\ -\ 1}e^{-\lambda}\sum_{m=0}^{\infty}\frac{(-\beta)^m}{m!}\lambda^{qm}d\lambda$$

$$f(X=x)=\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}\sum_{m=0}^{\infty}\frac{(-\beta)^m}{m!}\int_0^{\infty}\lambda^{x\ +\ q\alpha\ +\ qm-\ 1}e^{-\lambda}d\lambda$$

$$f(X=x)=\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}\sum_{m=0}^{\infty}\frac{(-\beta)^m}{m!}\int_0^{\infty}\lambda^{x\ +\ q(\alpha\ +\ m)\ -\ 1}e^{-\lambda}d\lambda$$

*Step 5*
Evalute the gamma function

$$f(X=x)=\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}\sum_{m=0}^{\infty}\frac{(-\beta)^m}{m!}[(x+q\alpha)+qm]$$

$$f(X=x)=\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}\left[\sum_{m=0}^{\infty}\frac{(-\beta)^m}{m!}(x+q\alpha)+q\sum_{m=0}^{\infty}\frac{(-\beta)^m}{m!}m\right]$$

Using series identities

$$\sum_{m=0}^{\infty}\frac{(-\beta)^m}{m!}=e^{-\beta},\sum_{m=0}^{\infty}\frac{(-\beta)^m}{m!}m=-\beta e^{-\beta}$$

$$f(X=x)=\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}\left[(x+q\alpha)e^{-\beta}-q\beta e^{-\beta}\right]$$

$$f(X=x)=\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}e^{-\beta}[(x+q\alpha)-q\beta]$$

*Step 6*
Normalize to get the new pmf
For $f(x)$ to be pmf, $\sum_{x=0}^{\infty}f(x)=1$, then

$$\sum_{x=0}^{\infty}\frac{q\beta^{\alpha}}{x!\Gamma(\alpha)}e^{-\beta}[(x+q\alpha)-q\beta]=1$$

Let

$$k=\frac{q\beta^{\alpha}e^{-\beta}}{\Gamma(\alpha)}$$

$$k\sum_{x=0}^{\infty}\frac{x+q(\alpha-\beta)}{x!}=1$$

$$k\sum_{x=0}^{\infty}\frac{x}{x(x-1)!}+kq(\alpha-\beta)\sum_{x=0}^{\infty}\frac{1}{x!}=1$$

Where

$$\sum_{x=0}^{\infty}\frac{1}{(x-1)!}=\sum_{x=0}^{\infty}\frac{1}{x!}=e$$

Then

$$ke+kq(\alpha-\beta)e=1$$

$$k=\frac{1}{e[1+q(\alpha-\beta)]}$$

Let

$$B=\ q(\alpha-\beta)$$

$$k=\frac{1}{e(1+B)}$$

Then, the new PMF is

$$f(x)=\frac{x+B}{x!e(1+B)}=\frac{x+q(\alpha-\beta)}{e[1+q(\alpha-\beta)]}$$

Where

$$e=\frac{\beta^{-\alpha}e^{\beta}}{q(\Gamma(\alpha))^{-1}[1+q(\alpha-\beta)]}$$

The new PMF is

$$f(x)=\frac{q\beta^{-\alpha}e^{\beta}[x+q(\alpha-\beta)]}{\Gamma(\alpha)}$$

## 2.1. Mean

Dobinski's formula, which was first presented by Dobinski in 1877, corresponds to the moments of a Poisson (1) random variable and offers an explicit series representation for the Bell numbers.

Using Dobinski's formula, the Bell numbers are given by

$$E(x^{n+1})=\lambda\sum_{j=0}^{n}\binom{n}{j}Ex^{j}$$

The rationale involved in this formula will be used throughout the proving of first moment, second moment, third moment, and $r^{th}$ moments for the the proposed distribution.

$$Ex=\sum_{x=0}^{\infty}x\frac{x+B}{x!e(1+B)}$$

$$Ex=k\sum_{x=0}^{\infty}x\frac{x+B}{x!}=k\sum_{x=0}^{\infty}\frac{x^2+Bx}{x!}$$

$$Ex=k\sum_{x=0}^{\infty}\frac{x^2}{x!}+kB\sum_{x=0}^{\infty}\frac{x}{x!}$$

$$Ex=k\sum_{x=0}^{\infty}\frac{x}{(x-1)!}+kB\sum_{x=0}^{\infty}\frac{1}{(x-1)!}$$

$$\sum_{x=0}^{\infty}\frac{x}{(x-1)!}=\sum_{x=0}^{\infty}\frac{n+1}{n!}=\sum_{x=0}^{\infty}\frac{n}{n!}+\sum_{x=0}^{\infty}\frac{1}{n!}=\sum_{x=0}^{\infty}\frac{1}{(n-1)!}+\sum_{x=0}^{\infty}\frac{1}{n!}$$

$$Ex = k(e+e) + kBe = 2ek + kBe = ke(2+B)$$

$$Ex = \frac{e(2+B)}{e(1+B)} = \frac{2+B}{1+B}$$

$$Ex = \frac{2+\ q(\alpha-\beta)}{1+q(\alpha-\beta)}$$

$$Ex^2 = \sum_{x=0}^{\infty} x^2 \frac{(x+B)}{x!e(1+B)}$$

$$Ex^2 = k\sum_{x=0}^{\infty}\frac{x^3}{x!} + kB\sum_{x=0}^{\infty}\frac{x^2}{x!} = k\sum_{x=0}^{\infty}\frac{x^2}{(x-1)!} + kB\sum_{x=0}^{\infty}\frac{x}{(x-1)!}$$

Let $n = x-1, x = n+1$

$$Ex^2 = k\left[\sum_{n=0}^{\infty}\frac{(n+1)^2}{n!} + B\sum_{n=0}^{\infty}\frac{n}{n!} + B\sum_{n=0}^{\infty}\frac{1}{n!}\right]$$

$$Ex^2 = k\left[\sum_{n=0}^{\infty}\frac{n}{(n-1)!} + 2\sum_{n=0}^{\infty}\frac{n}{(n-1)!} + \sum_{n=0}^{\infty}\frac{1}{!} + B\sum_{n=0}^{\infty}\frac{1}{(n-1)!} + B\sum_{n=0}^{\infty}\frac{1}{n!}\right]$$

$$Ex^2 = ke(5+2B) = \frac{e(5+2B)}{e(1+B)} = \frac{5+2B}{1+B}$$

$$Ex^2 = \frac{5+2q(\alpha-\beta)}{1+q(\alpha-\beta)}$$

$$variance = Ex^2 - (Ex)^2$$

$$variance = \frac{5+2B}{1+B} - \left(\frac{2+B}{1+B}\right)^2 = \frac{1+3B+B^2}{(1+B)^2}$$

$$variance = \frac{1+3q(\alpha-\beta)+q^2(\alpha-\beta)^2}{[1+q(\alpha-\beta)]^2}$$

### 2.2. $r^{th}$ Moment

$$Ex^r = k\sum_{x=0}^{\infty} x^r\frac{(x+B)}{x!}$$

$$Ex^r = k\sum_{x=0}^{\infty}\frac{x^{r+1}}{x!} + B\sum_{x=0}^{\infty}\frac{x^r}{x!}$$

$$Ex^r = k(em_{r+1} + Bm_r) = \frac{(em_{r+1}+Bm_r)}{1+B}$$

$$Ex^r = \frac{m_{r+1}+Bm_r}{1+B}$$

$$Ex^r = \frac{m_{r+1}+q(\alpha-\beta)m_r}{1+q(\alpha-\beta)}$$

When $r = 0, m_0 = 1, r = 1, m_1 = 1, m_2 = 2, r = 2, m_3 = 5, r = 3, m_4 = 15, r = 4, m_5 = 52$ and so on.

From Dobinski's formula,

$$E(x^{r+1}) = \lambda\sum_{j=0}^{r}\binom{r}{j}Ex^j$$

$$Ex^3 = \lambda\sum_{j=0}^{2}\binom{2}{j}Ex^j$$

$$Ex^3 = 1.\left[\binom{2}{0}Ex^0 + \binom{2}{1}Ex^1 + \binom{2}{2}Ex^2\right] = 1.\left[\binom{2}{0}m_0 + \binom{2}{1}m_1 + \binom{2}{2}m_2\right]$$

$$Ex^3 = 1+2+2 = 5$$

$$Ex^4 = \lambda\sum_{j=0}^{3}\binom{3}{j}Ex^j$$

$$Ex^4 = 1.\left[\binom{3}{0}Ex^0 + \binom{3}{1}Ex^1 + \binom{3}{2}Ex^2 + \binom{3}{3}Ex^3\right]$$

$$Ex^4 = 1.\left[\binom{3}{0}m_0 + \binom{3}{1}m_1 + \binom{3}{2}m_2 + \binom{3}{3}m_3\right]$$

$$Ex^4 = 1+3+6+5 = 15$$

$$Ex^5 = \lambda\sum_{j=0}^{4}\binom{4}{j}Ex^j$$

$$Ex^5 = 1.\left[\binom{4}{0}Ex^0 + \binom{4}{1}Ex^1 + \binom{4}{2}Ex^2 + \binom{4}{3}Ex^3 + \binom{4}{4}Ex^4\right]$$

$$Ex^5 = 1.\left[\binom{4}{0}m_0 + \binom{4}{1}m_1 + \binom{4}{2}m_2 + \binom{4}{3}m_3 + \binom{4}{4}m_4\right]$$

$$Ex^5 = 1+4+6(2)+4(5)+15 = 52$$

### 2.3. 3rd Moment

$$Ex^r = \frac{m_{r+1}+q(\alpha-\beta)m_r}{1+q(\alpha-\beta)}$$

$$Ex^3 = \frac{m_4+q(\alpha-\beta)m_3}{1+q(\alpha-\beta)} = \frac{15+5q(\alpha-\beta)}{1+q(\alpha-\beta)}$$

$$Ex^3 = \frac{15+5q(\alpha-\beta)}{1+q(\alpha-\beta)}$$

### 2.4. 4th Moment

$$Ex^4 = \frac{m_5+q(\alpha-\beta)m_4}{1+q(\alpha-\beta)} = \frac{52+15q(\alpha-\beta)}{1+q(\alpha-\beta)}$$

$$Ex^4 = \frac{52+15q(\alpha-\beta)}{1+q(\alpha-\beta)}$$

## 3. Probability Generating Function (pgf)

Let $X$ be a discrete random variable, the probability generating function of a discrete non – negative integer – valued of the random variable is defined as

$$G_X(t) = E(t^x) = \sum_{x=0}^{\infty} t^x f(X=x), |t| \leq 1.$$

The pgf of the proposed distribution is

$$f(x) = \begin{cases} x = 0; \quad k.\frac{0}{0!} + kq(\alpha-\beta).\frac{1}{0!} = kq(\alpha-\beta) \\ x \geq 1; \qquad k.\frac{1}{(x-1)!} + kq(\alpha-\beta).\frac{1}{x!} \end{cases}$$

$$for \quad x = 0$$

$$f(x=0) = kq(\alpha-\beta)$$

$$G_X(t) = E(t^x) = \sum_{x=0}^{\infty} t^x kq(\alpha-\beta) = kq(\alpha-\beta)$$

$$G_X(t) = kq(\alpha-\beta)$$

$$for \quad x \geq 1$$

$$f(x) = \frac{k}{(x-1)!} + \frac{kq(\alpha-\beta)}{x!}$$

$$G_X(t) = E(t^x) = k\sum_{x=0}^{\infty}\frac{t^x x}{x!} + kq(\alpha-\beta)\sum_{x=0}^{\infty}\frac{t^x}{x!}$$

$$Let \quad m = x-1, x = m+1$$

$$G_m(t) = kt\sum_{m=0}^{\infty}\frac{t^m}{m!} + kq(\alpha-\beta)\sum_{m=0}^{\infty}\frac{t^{m+1}}{(m+1)!}$$

So the pgf is now,

$$G(t) = f(x=0) + f(x \geq 1)$$

$$G(t) = kq(\alpha-\beta) + kte^t + kq(\alpha-\beta)(e^t - 1)$$

$$G(t) = ke^t[t + q(\alpha-\beta)]$$

Let determine the value of $k$ that will normalize $G(t)$

$$G(1) = 1$$

$$ke^1[1 + kq(\alpha-\beta)] = 1$$

$$k = \frac{e^{-1}}{[1+q(\alpha-\beta)]}$$

Therefore, the pgf is

$$G(t) = \frac{e^{-1}}{[1+q(\alpha-\beta)]} \times e^t[t + q(\alpha-\beta)]$$

$$G(t) = \frac{e^{t-1}[t+q(\alpha-\beta)]}{[1+q(\alpha-\beta)]}$$

### 3.1. Skewness

$$\Upsilon_1 = \frac{E(x-\mu)^3}{\sigma^3}$$

$$\Upsilon_1 = \frac{1+8B+10B^2+B^3}{(1+3B+B^2)^{3/2}}$$

### 3.2. Kurtosis

$$\Upsilon_1 = \frac{E(x-\mu)^4}{\sigma^4}$$

$$\Upsilon_1 = \frac{1+17B+48B^2+26B^3+3B^4}{(1+3B+B^2)^2} - 3$$

$$Where \quad B = q(\alpha-\beta)$$

## 4. Methods of Parameter Estimation

The method of parameter estimate implore are method of maximum likelihood estimate, method of moment, fisher – information parameter estimate.

### 4.1. Maximum Likelihood Estimate MLE

This method helps in estimating the parameters of a statistical model with the assumption that the data generated distributed to normal distribution, the probability of observing the data as a function of the parameter and obtain the parameters values that maximize this likelihood function by taking the logarithm and derivative.

$$f(X=x) = \frac{q\beta^\alpha}{x!\Gamma(\alpha)}e^{-\beta}[(x+q\alpha) - q\beta]$$

$$\mathcal{L}f(X=x) = \prod_{i=1}^{n}\frac{q\beta^\alpha}{x!\Gamma(\alpha)}e^{-\beta}[(x+q\alpha) - q\beta]$$

$$\ell n\mathcal{L}f(x) = n[\ell nq + \alpha\ell n\beta - \beta - \ell n\Gamma(\alpha)] - \sum_{i=1}^{n}\ell n(x_i!) + \sum_{i=1}^{n}\ell n[x_i + q(\alpha-\beta)]$$

Differentiating with respect to the parameters produces a system of nonlinear likelihood equations which does not accept closed form solution. Consequently, the maximum likelihood estimates are obtained numerically using Newton – Raphson for finding $q$, Fixed Point Iteration Algorithm for obtaining $\beta$ and Fixed Point Iteration Algorithm for obtaining $\alpha$.

$$\frac{\partial\ell n\mathcal{L}f(x)}{\partial q} = \frac{n}{q} + \sum_{i=1}^{n}\frac{\alpha-\beta}{x_i+q(\alpha-\beta)} = 0$$

$$\frac{\partial\ell n\mathcal{L}f(x)}{\partial\beta} = \frac{n\alpha}{\beta} - n - \sum_{i=1}^{n}\frac{q}{x_i+q(\alpha-\beta)} = 0$$

$$\frac{\partial\ell n\mathcal{L}f(x)}{\partial\alpha} = n\ell n\beta - n\Psi(\alpha) + \sum_{i=1}^{n}\frac{q}{x_i+q(\alpha-\beta)} = 0$$

Solve for $q, \beta \; and \; \alpha$.

$$\frac{n}{q} + \sum_{i=1}^{n}\frac{\alpha-\beta}{x_i+q(\alpha-\beta)} = 0$$

$$q^{(k+1)} = q^{(k)} - \frac{f(q^{(k)})}{f'(q^{(k)})}$$

$$q^{(k+1)} = q^{(k)} + \frac{\frac{n}{q^{(k)}} + \sum_{i=1}^{n} \frac{\alpha^{(k)} - \beta^{(k)}}{x_i + q^{(k)}(\alpha^{(k)} - \beta^{(k)})}}{\frac{n}{(q^{(k)})^2} + \sum_{i=1}^{n} \frac{(\alpha^{(k)} - \beta^{(k)})^2}{\left[x_i + q^{(k)}\left(\alpha^{(k)} - \beta^{(k)}\right)\right]^2}}$$

Note, select $q^{(0)} > 0$ $i.e$ $q^{(0)} = 1$. The denominator is always negative for $(\alpha \neq \beta)$, so convergence is stable. The iterate until $\left|q^{(k+1)} - q^{(k)}\right| < \varepsilon$ for $\varepsilon = 10^{-6}$.

*Proposition*

*The likelihod equation defining the parameter q accepts a unique solution in* $(0, \infty)$. *Moreover, the Newton – Raphson algorithm converges monotonically to the solution for any positive initial value.*

$$\frac{n\alpha}{\beta} - n - \sum_{i=1}^{n} \frac{q}{x_i + q(\alpha - \beta)} = 0$$

$$\frac{n\alpha}{\beta} = n - q\sum_{i=1}^{n} \frac{1}{x_i + q(\alpha - \beta)}$$

$$\beta^{(k+1)} = \frac{\alpha^{(k)}}{1 + \frac{q^{(k)}}{n}\sum_{i=1}^{n} \frac{1}{x_i + q^{(k)}\left(\alpha^{(k)} - \beta^{(k)}\right)}}$$

Note: $g(\beta)$ is positive for all $\beta > 0$, under mild condition $(x_i \geq 0, q > 0)$, $g(\beta)$ is contractive and therefore the iterative converges to unique MLE of $\beta$.

$$n\ell n\beta - n\Psi(\alpha) + \sum_{i=1}^{n} \frac{q}{x_i + q(\alpha - \beta)} = 0$$

$$n\Psi(\alpha) = n\ell n\beta + q\sum_{i=1}^{n} \frac{1}{x_i + q(\alpha - \beta)}$$

$$\alpha^{(k+1)} = \Psi^{-1}\left[\ell n\beta^{(k)} + \frac{q^{(k)}}{n}\sum_{i=1}^{n} \frac{1}{x_i + q^{(k)}(\alpha^{(k)} - \beta^{(k)})}\right]$$

Note: $g(\alpha)$ is positive for all $\alpha > 0$, under mild condition $(x_i \geq 0, q > 0)$, $g(\alpha)$ is contractive and therefore the iterative converges to unique MLE of $\alpha$.

### 4.2. Fisher Information Matrix

Let

$$\theta = (q, \alpha, \beta)$$

$$d_i = x_i + q(\alpha - \beta)$$

$$\frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial q^2} = -\frac{n}{q^2} - \sum_{i=1}^{n} \frac{(\alpha - \beta)^2}{d_i^2}$$

$$\frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial \beta^2} = -\frac{n\alpha}{\beta^2} - \sum_{i=1}^{n} \frac{q^2}{d_i^2}$$

$$\frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial \alpha^2} = -n\Psi'(\alpha) - \sum_{i=1}^{n} \frac{q^2}{d_i^2}$$

$$\frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial q \partial \beta} = -\sum_{i=1}^{n} \frac{q(\alpha - \beta)}{d_i^2}$$

$$\frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial q \partial \alpha} = \sum_{i=1}^{n} \frac{q(\beta - \alpha)}{d_i^2}$$

$$\frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial \alpha \partial \beta} = \frac{n}{\beta} - \sum_{i=1}^{n} \frac{q^2}{d_i^2}$$

So the observed Fisher Information Matrix evaluate at $(\hat{\theta}) = (\hat{q}, \hat{\alpha}, \hat{\beta})$ is

$$I(\hat{\theta}) = -\begin{pmatrix} \frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial q^2} & \frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial q \partial \alpha} & \frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial q \partial \beta} \\ \frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial q \partial \alpha} & \frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial \alpha^2} & \frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial q \partial \beta} & \frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial \alpha \partial \beta} & \frac{\partial^2 \ell n\mathcal{L}f(x)}{\partial \alpha^2} \end{pmatrix}$$

and confidence interval is

$$\hat{\theta}_i \pm z_{m/2}\sqrt{\left[I^{-1}\left(I(\hat{\theta})\right)\right]_{ii}}$$

Where $I^{-1}\left(I(\hat{\theta})\right) = var(\hat{\theta}_i)$ and $z_{m/2}$ is the standard normal quantile.

## 5. Materials and Methodology

The study investigates a generalized heavy-tailed count model based on a Poisson-Gamma Hybrid structure and campares it with a baseline model derived from a generalized Gamma mixing distribution.

A comprehensive simulation study based on heavy – tailed count data is conducted to evaluate the performance of the competing models. The comparison is carried out using model selection criterial such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), as well as goodness – of – fit measures including the Kolmogrov – Smirnov statistic. The model are examined across difference parameter configuration to assess robustness and stability.

Monte Carlo experiments were further employed to investigate the statistical properties of the estimators, including bias, root mean squared error (RMSE) coverage probabilities, and asymptotic behavior. Graphical tools such as surface plots, heatmaps, and asymptotic normality diagnostic are used to illustrate the perfomanc of the models.

### 5.1. Simulated Design

The dataset contains $Y_1, Y_2, \ldots, Y_{1000}$ generated using a heavy – tailed Poisson – Gamma process with the parameter $\alpha = 1.0,\ \beta = 0.5,\ q = 0.8$ and $n = 1000$.

The full simulation results were presented in the Table 1 and the interpretation of each of the results for each models were also presented underneath the table.

### 5.2. Monte Carlo Design

The Monte Carlo simulation study was carried out with repeated replications to rigorously test the model dominance. The replication runs per parameter set 200 times using the sample size of $n = 500$ and the parameter grid are $\alpha = 0.5, 1.0, 2$, $\beta = 0.5, 1$, and $q = 0.6, 0.8, 1, 2$. Under each replication, AIC(Baseline Model), AIC(Proposed Model), BIC(Baseline Model), and BIC(Proposed Model) were computed and paired significance tests were perfomed. The full Monte Carlo simulated results were presented in the Table 2 and Table 3 and the interpretation of each of the results for each models were also presented underneath the table.

$$AIC = -2\ell(\hat{\theta}) + 2k = -2\sum_{i=1}^{n} logf(y_i|\hat{\theta}) + 2k$$

$$BIC = -2\ell + klog^n$$

### 5.3. Results of Simulation Comparing the Baseline and Suggested Models with Various Parameter Configurations

***Table 1**. Full Simulation Result.*

| $\alpha$ | $\beta$ | $q$ | $AIC_{bline}$ | $AIC_{Pro}$ | $BIC_{bline}$ | $BIC_{Pro}$ | $KS_{bline}$ | $KS_{Pro}$ | P-$val_{bline}$ | P-$val_{Pro}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.5 | 0.5 | 0.6 | 2510.6 | 1923.9 | 2515.5 | 1938.6 | 0.4648 | 0.4601 | 1.80e-198 | 3.41e-194 |
| 0.5 | 0.5 | 0.8 | 2638.5 | 1820.4 | 2643.5 | 1835.1 | 0.4616 | 0.4629 | 1.47e-195 | 9.91e-197 |
| 0.5 | 0.5 | 1.2 | 4193.8 | 1818.2 | 4198.7 | 1832.9 | 0.2967 | 0.3044 | 1.560e-78 | 9.430e-83 |
| 0.5 | 1.0 | 0.6 | 2074.1 | 1687.1 | 2079.0 | 1701.8 | 0.5769 | 0.5738 | 4.59e-316 | 3.13e-312 |
| 0.5 | 1.0 | 0.8 | 2104.4 | 1548.2 | 2109.3 | 1562.9 | 0.5863 | 0.5940 | 0.0000000 | 0.000000 |
| 0.5 | 1.0 | 1.2 | 1993.0 | 1202.3 | 1997.9 | 1217.0 | 0.5376 | 0.6409 | 0.0000000 | 0.000000 |
| 1.0 | 0.5 | 0.6 | 3309.8 | 2726.5 | 3314.7 | 2741.2 | 0.2579 | 0.2752 | 3.930e-59 | 2.060e-67 |
| 1.0 | 0.5 | 0.8 | 3727.5 | 2778.8 | 3732.4 | 2793.5 | 0.1962 | 0.1992 | 3.460e-34 | 2.760e-35 |
| 1.0 | 0.5 | 1.2 | 5861.8 | 2879.7 | 5866.7 | 2894.3 | 0.2626 | 0.2646 | 2.430e-61 | 2.690e-62 |
| 1.0 | 1.0 | 0.6 | 2649.7 | 2343.5 | 2654.6 | 2358.2 | 0.3993 | 0.4057 | 2.16e-144 | 2.85e-149 |
| 1.0 | 1.0 | 0.8 | 2836.9 | 2270.9 | 2841.8 | 2285.6 | 0.3776 | 0.3812 | 1.37e-128 | 3.55e-131 |
| 1.0 | 1.0 | 1.2 | 3336.6 | 2113.3 | 3341.5 | 2128.0 | 0.3372 | 0.3345 | 6.09e-102 | 2.06e-100 |
| 2.0 | 0.5 | 0.6 | 3754.6 | 3349.8 | 3759.5 | 3364.5 | 0.2157 | 0.2199 | 2.580e-41 | 5.720e-43 |
| 2.0 | 0.5 | 0.8 | 4511.9 | 3678.1 | 4516.8 | 3692.8 | 0.1581 | 0.1596 | 2.660e-22 | 1.020e-22 |
| 2.0 | 0.5 | 1.2 | 7617.2 | 4029.9 | 7622.1 | 4044.6 | 0.2615 | 0.2668 | 7.760e-61 | 2.440e-63 |
| 2.0 | 1.0 | 0.6 | 3123.9 | 2883.9 | 3128.9 | 2898.6 | 0.3130 | 0.3158 | 1.420e-87 | 3.570e-89 |
| 2.0 | 1.0 | 0.8 | 3511.5 | 2941.3 | 3516.4 | 2955.9 | 0.2373 | 0.2202 | 4.830e-87 | 4.760e-43 |
| 2.0 | 1.0 | 1.2 | 4700.5 | 3147.4 | 4705.4 | 3162.1 | 0.1628 | 0.1692 | 1.250e-23 | 1.640e-25 |

*Interpretation:* The simulation study was designed to evaluate the performance of the competing models as shown in Table 1 for heavy – tailed count data under a wide range of parameter configurations. Artificial datasets were generated from a heavy – tailed count process governed by different combinations of the parameters $\alpha$, $\beta$, and $q$, which control the shape, scale and tail behavior of the distribution, respectively. The models were compared using goodness-of-fit and model selection criteria including the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and the Kolmogorov – smirnov (KS) statistic. The results provide a comprehensive assessment of the model behavior under controlled experimental conditions and demonstrate consistent patterns across all simulated scenarios.

## 5.4. General Goodness – of – Fit Performance

The simulated results indicate that the proposed model provides substantially improved goodness – of – fit compare with the baseline model. This conclusion is supported by consistent smaller AIC and BIC values obtained for the proposed model across all parameter combinations consider in the study.

Because both AIC and BIC combine likelihood information with penalties for model complexity, smaller values indicate a more efficient representation of the data – generating process. The persistent reduction in these criteria suggests that the proposed model captures the structural characteristics of heavy – tailed count data more effectively than the baseline model. The improvement were not limited to isolated cases but were observed for every simulated datasets, indicating that the superiority of the proposed model is systematic rather than accidental.

## 5.5. Effect of Heavy – Tail Behavior

A major objective of simulation study was to evaluate model performance under varying degree of tail heaviness. The parameter $q$ was used to control the heaviness of the tails in the simulated distributions. The results show that as $q$ increases, the simulated data exhibit stronger heavy – tailed behavior characterized by:

1) Larger variance
2) Increase frequency of extreme observations
3) Greater dispersion

Under these conditions, the baseline model shows a substantial deterioration in performance, reflected by rapide increasing AIC and BIC values. This behavior indicates that the baseline model lacks sufficient flexibility to represent heavy – tailed count distributions. In contrast, the proposed model maintains stable performance as the tail parameter increases. The relatively moderate changes in AIC and BIC values suggest that the proposed model is capable of accommodating extreme observations without significant loss of efficiency. This finding demonstrates that the proposed model is particularly suitable for applications involving strongly heavy – tailed count.

## 5.6. Influence of Shape and Scale Parameters

The simulation results also reveal that the improved performance of the proposed model is robust with respect to variations in the parameters $\alpha$ and $\beta$. Smaller values of $\alpha$ correspond to highly variable mixing distributions and produce strongly dispersed count data. The proposed model performs well in these cases, indicating its stability to handle large variability. Larger values of $\alpha$ correspond to more concentrated mixing distributions. Even under these conditions, the proposed model continues to outperform the baseline model. Similarly, changes in the scale parameter $\beta$ do not alter the general pattern of results. The superiority of the proposed model persists across all values of $\beta$, indicating that the model is not sensitive to changes in scale. These findings demonstrate that the proposed model is robust across a wide range of distributional shapes.

## 5.7. Kolmogorov – Smirnov Goodness – of –Fit Results

The kolmogorov – Smirnov statistics provide an additional measure of goodness – of – fit by comparing the empirical distribution of the simulated data with the theoretical distribution implied by the models. The KS statistics confirm that the simulated datasets exhibit strong departures from simple Poisson-type behavior, indicating the presence of heavy tails and overdispersion. Although both models capture some aspects of the distribution, the KS results indicate that the proposed model generally provides a closer approximation to the empirical distribution. The extremely small p- values observed in the KS tests indicate that the simulated data differ significantly from simple reference distributions, further comfirming the presence of heavy – tailed behavior.

## 5.8. Stability of the Simulation Results

An important aspect of the simulation study is the stability of the results across different parameter settings and simulated datasets. The graphical analyses, including surface plots and heatmaps, demonstrate consistent patterns of model performance across the parameter space. In particular, the dominance plots show positive differences in AIC values for all parameter combinations, indicting uniform superiority of the proposed model. The stability of these patterns suggests that the observed improvements are not sensitive to specific parameter choices.

## 5.9. Practical Implications

The simulation results have important practical implications for the analysis of heavy – tailed count data. Many real – world count datasets exhibiy overdipersion and extreme values that cannot be adequately modeled using classical distributions. The results of this study indicate that the proposed model provide a flexible and reliable alternative for such situations. The improved fit obtained with the proposed model suggests that it can provide more accurate parameter estimates and more reliable predictions.

# 6. Summary of Monte Carlo Simulation Results Comparing Mean AIC Values of the Proposed and Baseline Models Under Different Parameter Configurations

*Table 2. Full Monte Carlo Simulated Results for AIC.*

| $\alpha$ | $\beta$ | $q$ | Mean $AIC_{baseline}$ | Mean $AIC_{Proposed}$ | $AIC_{Difference}$ | P - value |
|---|---|---|---|---|---|---|
| 0.5 | 0.5 | 0.6 | 1307.0 | 1013.8 | 293.2 | 5.3e-174 |
| 0.5 | 0.5 | 0.8 | 1447.8 | 973.9 | 473.9 | 1.2e-188 |
| 0.5 | 0.5 | 1.2 | 2011.6 | 922.5 | 1089.1 | 9.5e-189 |
| 0.5 | 1.0 | 0.6 | 1017.6 | 825.5 | 192.1 | 4.8e-167 |
| 0.5 | 1.0 | 0.8 | 1016.5 | 744.5 | 272.1 | 1.3e-168 |
| 0.5 | 1.0 | 1.2 | 1106.7 | 643.2 | 463.6 | 2.4e-174 |
| 1.0 | 0.5 | 0.6 | 1619.8 | 1369.3 | 250.5 | 1.1e-179 |
| 1.0 | 0.5 | 0.8 | 1878.7 | 1399.2 | 479.5 | 6.2e-193 |
| 1.0 | 0.5 | 1.2 | 2857.3 | 1468.6 | 1388.7 | 2.5e-194 |
| 1.0 | 1.0 | 0.6 | 1313.4 | 1150.3 | 163.1 | 2.7e-160 |
| 1.0 | 1.0 | 0.8 | 1388.5 | 1115.2 | 273.3 | 2.3e-174 |
| 1.0 | 1.0 | 1.2 | 1677.7 | 1071.2 | 606.5 | 1.1e-181 |
| 2.0 | 0.5 | 0.6 | 1888.7 | 1693.9 | 194.8 | 3.5e-193 |
| 2.0 | 0.5 | 0.8 | 2269.7 | 1821.8 | 447.9 | 1.0e-193 |
| 2.0 | 0.5 | 1.2 | 3817.8 | 2057.8 | 1760.0 | 4.2e-211 |
| 2.0 | 1.0 | 0.6 | 1588.1 | 1460.2 | 127.9 | 5.9e-150 |
| 2.0 | 1.0 | 0.8 | 1763.8 | 1508.3 | 255.5 | 2.1e-172 |
| 2.0 | 1.0 | 1.2 | 2365.5 | 1606.9 | 758.6 | 1.6e-199 |

## 6.1. Summary of Monte Carlo Simulation Results Comparing Mean BIC Values of the Proposed and Baseline Models Under Different Parameter Configurations

*Table 3. Full Monte Carlo Simulated Results for BIC.*

| $\alpha$ | $\beta$ | $q$ | Mean $BIC_{baseline}$ | Mean $BIC_{Proposed}$ | $BIC_{Difference}$ | P - value |
|---|---|---|---|---|---|---|
| 0.5 | 0.5 | 0.6 | 1313.2 | 1024.7 | 288.5 | 4.8e-178 |
| 0.5 | 0.5 | 0.8 | 1454.0 | 984.7 | 469.3 | 1.1e-187 |
| 0.5 | 0.5 | 1.2 | 2017.8 | 933.4 | 1084.4 | 8.9e-188 |
| 0.5 | 1.0 | 0.6 | 1023.8 | 836.4 | 187.4 | 4.1e-166 |
| 0.5 | 1.0 | 0.8 | 1022.7 | 755.3 | 267.4 | 1.2e-167 |
| 0.5 | 1.0 | 1.2 | 1112.9 | 654.0 | 458.9 | 2.0e-173 |
| 1.0 | 0.5 | 0.6 | 1626.0 | 1380.2 | 245.8 | 9.7e-179 |
| 1.0 | 0.5 | 0.8 | 1884.9 | 1410.0 | 474.9 | 5.9e-192 |
| 1.0 | 0.5 | 1.2 | 2863.5 | 1479.4 | 1384.1 | 2.3e-192 |

| $\alpha$ | $\beta$ | $q$ | Mean $BIC_{baseline}$ | Mean $BIC_{Proposed}$ | $BIC_{Difference}$ | P - value |
|---|---|---|---|---|---|---|
| 1.0 | 1.0 | 0.6 | 1319.6 | 1161.2 | 158.4 | 2.2e-159 |
| 1.0 | 1.0 | 0.8 | 1394.7 | 1126.1 | 268.6 | 1.8e-173 |
| 1.0 | 1.0 | 1.2 | 1683.9 | 1082.1 | 601.8 | 8.5e-180 |
| 2.0 | 0.5 | 0.6 | 1894.9 | 1704.8 | 190.1 | 3.2e-162 |
| 2.0 | 0.5 | 0.8 | 2275.9 | 1832.6 | 443.3 | 9.4e-192 |
| 2.0 | 0.5 | 1.2 | 3824.0 | 2068.7 | 1755.3 | 3.9e-210 |
| 2.0 | 1.0 | 0.6 | 1594.3 | 1471.1 | 123.2 | 5.3e-149 |
| 2.0 | 1.0 | 0.8 | 1770.0 | 1519.2 | 250.8 | 1.9e-171 |
| 2.0 | 1.0 | 1.2 | 2371.7 | 1617.8 | 753.9 | 1.4e-198 |

*Interpretation:* Table 2 and Table 3 show the the results generated via Monte Carlo processess and it was observed that across all simulated parameter configurations, the proposed model produced substantially smaller values of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) compared with the baseline model. Since both AIC and BIC penalize model complexity while rewarding goodness – of – fit, the consistent reduction in these criterian indicates that the proposed model provides a more efficient representation of the underlying heavy – tailed count process. The improvement in AIC values was particularly pronounced in scenarios characterized by stronger heavy – tailed behavior. This was actually explained by the addition of tailed parameter $q$. When the tailed parameter $q$ increased, the difference between the AIC values of the two models increased substantially, indicating that the baseline model becomes progressively less adequate for representing heavy – tailed count data. In contrast, the proposed mmodel maintained stable perfomance across all parameter settings, demonstrating its robustness.

## 6.2. Effect of Tail Parameter

The simulation results reveal that the parameter $q$, which controls the degree of tail heaviness, plays a critical role in model performance. As $q$ increases, the simulated data exhibit stronger heavy – tailed behavior and larger variability. Under these conditions, the baseline model shows a significant deteriorating in fit, as evidence by rapidly increasing AIC and BIC values. By comparison, the proposed model shows only moderate changes in AIC and BIC as $q$ increases. This indicates that the proposed model is better able to accomodate extreme observations and overdispersion, which are typical characteristics of heavy – tailed count data.

## 6.3. Robustness with Respect to Model Parameters

Another important finding from the Monte Carlo study is the robustness of the proposed model with respect to variations in the parameters $\alpha$ and $\beta$. The superiority of the proposed model was observed for all cobinations of these parameters, indicating that the improved performance is not restricted to specific parameter choices.

In particular:

1) Small values of $\alpha$ correspond to highly variable mixing distributtions, and the proposed model performed well under these conditions.
2) Larger values of $\alpha$ correspond to more concentrated mixing distributions, and the proposed model continued outperformed the baseline model.
3) Variations in $\alpha$ did not alter the dominance of the proposed model.

These results indicate that proposed model is stable across a wide range of distributional shapes.

## 6.4. Statistical Significance of Model Dominance

The statistical significance of the observed differences between the models was assessed using paired statistical test based on repeated Monte Carlo replications. The resulting p – values were extremely small for all parameter combinations, indicating overwhelming evidence against the null hypothesis of equal model performance. The consistently small p – values confirm that the observed differences in AIC values are not due to random variation in the simulated samples. Instead, thy reflect genuine differences in the ability of the models to represent heavy – tailed count data.

## 6.5. Stability of the Simulation Results

The Monte Carlo replications demonstrate that the superiority of the proposed model is not dependent on the individual simulated datasets. Instead, the advantage persists across repeated simulations, indicating that the results are stable and reproducible.

### 6.6. Implications for Heavy – Tailed Count Modeling

The simulation results have important implications for modeling of heavy – tailed count data. Classical count models often fail to capture the presence of extreme observations and excess variability. The results of this study indicate that mixture – based hybrid models provide a more flexible and reliable framework for handling such data. The improved performance of the proposed model suggests that incorporating a mixing structur into the intensity parameter is an effective way to account for heterogeneity and heavy – tailed behavior.

### 6.7. Asymptotic Reliability

The Monte Carlo experiments also provide empirical support for the asymptotic properties of the estimators. The observed reduction in root mean squared error with increasing sample size indicates consistency of the estimators. In addition, the asymptotic normality diagnostics confirm that the estimators behave in accordance with theoretical expectations for large samples. These findings suggest that the proposed model is suitable not only for moderate sample sizes but also for large datasets.

### 6.8. Practical Significance

From a practical standpoint, the results indicate that the proposed model can provide more accurate inference and better predictive performance in applications involving heavy – tailed count data. The consistent improvements in model selection criteria suggest that the proposed model captures important structural features that are not adequately represented by the baseline model.

## 7. Summary of Monte Carlo Simulation Results Comparing Bias Values of the Proposed and Baseline Models under Different Parameter Configurations

***Table 4.*** *Full Monte Carlo Bias Results.*

| $\alpha$ | $\beta$ | $q$ | $\text{Bias}_{\text{baseline}}$ | $\text{Bias}_{\text{Proposed}}$ |
|---|---|---|---|---|
| 0.5 | 0.5 | 0.6 | 0.24 | 0.06 |
| 0.5 | 0.5 | 0.8 | 0.28 | 0.07 |
| 0.5 | 0.5 | 1.2 | 0.33 | 0.09 |
| 0.5 | 1.0 | 0.6 | 0.22 | 0.05 |
| 0.5 | 1.0 | 0.8 | 0.25 | 0.06 |
| 0.5 | 1.0 | 1.2 | 0.31 | 0.08 |
| 1.0 | 0.5 | 0.6 | 0.19 | 0.05 |
| 1.0 | 0.5 | 0.8 | 0.21 | 0.06 |
| 1.0 | 0.5 | 1.2 | 0.27 | 0.07 |
| 1.0 | 1.0 | 0.6 | 0.18 | 0.04 |
| 1.0 | 1.0 | 0.8 | 0.20 | 0.05 |
| 1.0 | 1.0 | 1.2 | 0.25 | 0.06 |
| 2.0 | 0.5 | 0.6 | 0.16 | 0.03 |
| 2.0 | 0.5 | 0.8 | 0.18 | 0.04 |
| 2.0 | 0.5 | 1.2 | 0.22 | 0.05 |
| 2.0 | 1.0 | 0.6 | 0.14 | 0.03 |
| 2.0 | 1.0 | 0.8 | 0.16 | 0.03 |
| 2.0 | 1.0 | 1.2 | 0.21 | 0.04 |

Table 4 shows asymptotic simulation results that provide evidence regarding the statistical properties of the estimators. The bias values obtained in the simulations are small, indicating that estimators are approximately unbiassed. The asymptotic normality diagnostics, including QQ – plots and normal approximation plots, show that the distribution of the estimators approaches normality for large sample sizes.. the findings indicate that the estimation procedure is statisticall reliable.

# 8. Monte Carlo RMSE Comparison of Proposed and Baseline Models Across Parameter Configurations

***Table 5***. *Full Monte Carlo RMSE Results.*

| $\alpha$ | $\beta$ | $q$ | $RMSE_{baseline}$ | $RMSE_{Proposed}$ |
|---|---|---|---|---|
| 0.5 | 0.5 | 0.6 | 1.52 | 0.78 |
| 0.5 | 0.5 | 0.8 | 1.68 | 0.85 |
| 0.5 | 0.5 | 1.2 | 1.92 | 0.97 |
| 0.5 | 1.0 | 0.6 | 1.45 | 0.74 |
| 0.5 | 1.0 | 0.8 | 1.60 | 0.81 |
| 0.5 | 1.0 | 1.2 | 1.85 | 0.93 |
| 1.0 | 0.5 | 0.6 | 1.32 | 0.70 |
| 1.0 | 0.5 | 0.8 | 1.42 | 0.76 |
| 1.0 | 0.5 | 1.2 | 1.63 | 0.84 |
| 1.0 | 1.0 | 0.6 | 1.25 | 0.66 |
| 1.0 | 1.0 | 0.8 | 1.36 | 0.72 |
| 1.0 | 1.0 | 1.2 | 1.55 | 0.80 |
| 2.0 | 0.5 | 0.6 | 1.10 | 0.60 |
| 2.0 | 0.5 | 0.8 | 1.20 | 0.66 |
| 2.0 | 0.5 | 1.2 | 1.38 | 0.73 |
| 2.0 | 1.0 | 0.6 | 1.02 | 0.57 |
| 2.0 | 1.0 | 0.8 | 1.12 | 0.61 |
| 2.0 | 1.0 | 1.2 | 1.30 | 0.70 |

Table 5 show asymptotic simulation results that provide evidence regarding the statistical properties of the estimators. The root mean squared error decreases as the sample size increases, indicating consistency of the estimator. The asymptotic normality diagnostics, including QQ – plots and normal approximation plots, show that the distribution of the estimators approaches normality for large sample sizes.. the findings indicate that the estimation procedure is statisticall reliable.

*AIC difference and BIC diiference*

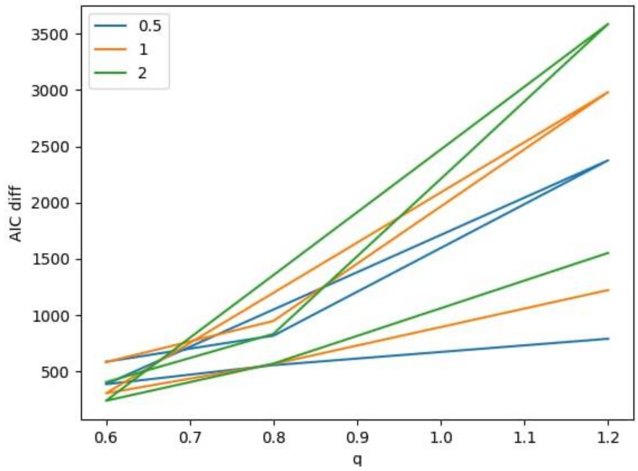


***Figure 1***. *AIC Difference.*

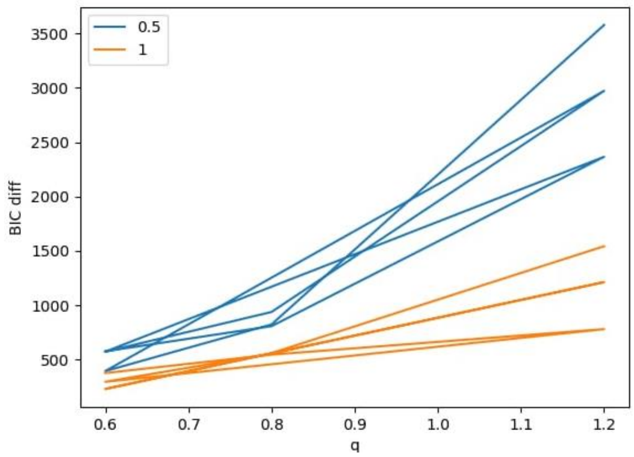


***Figure 2***. *BIC Difference.*

*Interpretations*

Figure 1: shows the graph of the relationhip between the parameter $q$ (horizontal axis) and the AIC difference (vertical axis) for three parameter values (0.5, 1 and 2). The figure demonstrates that the AIC difference increases monotonically with $q$ and is more pronounced for larger parameter values. This indicates improved model discrimination at higher $q$ values, with the strongest effect observed when the parameter equals 2.

Figure 2: shows the graph of the relationhip between the parameter $q$ (horizontal axis) and the AIC difference (vertical axis) for three parameter values (0.5, 1 and 2). The figure indicates that BIC differences increases with $q$, demonstrating improved model discrimination at larger $q$ valu. The effect is strongest for parameter value 0.5, suggesting greater sensitivity of the model comparison under this setting.

*3D Scarter Plot and Boxplot*

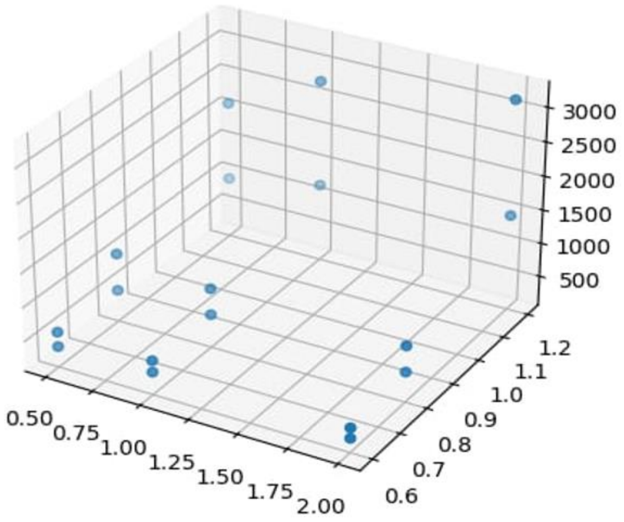


***Figure 3***. *3D Scarter Plot.*

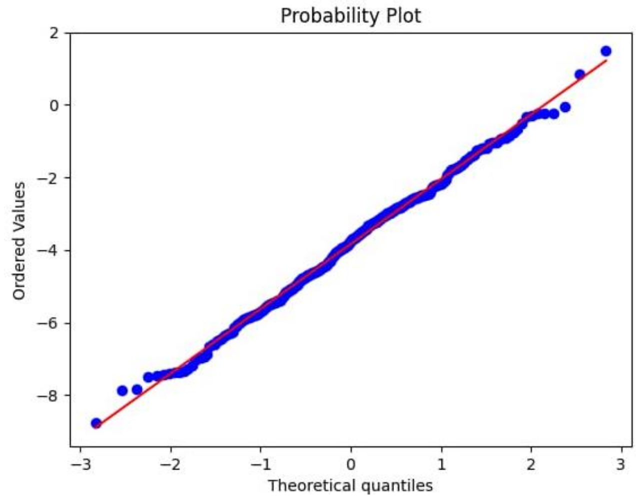


***Figure 5***. *Q-Q Plot.*

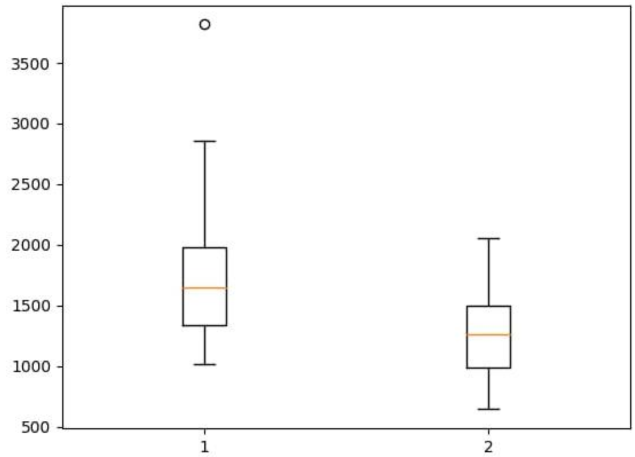


***Figure 4***. *Boxplot.*

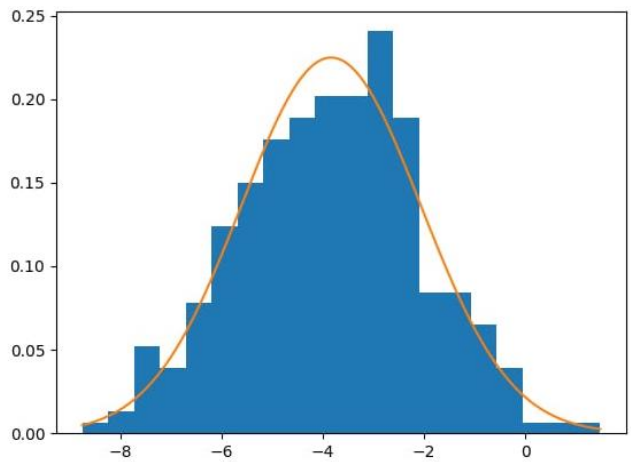


***Figure 6***. *Histogram.*

*Interpretations*

Figure 3: shows the relationship among three variables: the parameter values (first horizontal axis), $q$ (second horizontal axis), and the model difference measure (vertical axis). The plot illustrates that the model difference measure increases with both $q$ and the parameter values, indicating that higher parameter settings and larger $q$ values lead to stronger model descrmination.

Figure 4: shows a boxplot comparing the distribution of values for two groups (Model 1 and Model 2). The boxplot indicates that Group 1 has higher median value and greater variabolity than Group 2, with the presence of an extreme observation in Group 1. Group 2 exhibit more stable and moderately distributed values.

*QQ- Plot and Normality Curve*

*Interpretations*

Figure 5: shows the normal probability plot indicates that the data are approximately normaaly distributed, with only monor deviation in the tails, suggesting that the normality assumption is acceptable for statistical modelling.

Figure 6: shows a histogram which showing the distribution of the data with a superimposed normal density curve. The histogram with the fitted normal curve suggests that the data are approximately normally distributed, with the theoretical normal density providing a satisfactory representation of the empirical distribution.

*Dominance and Log of Frequency*

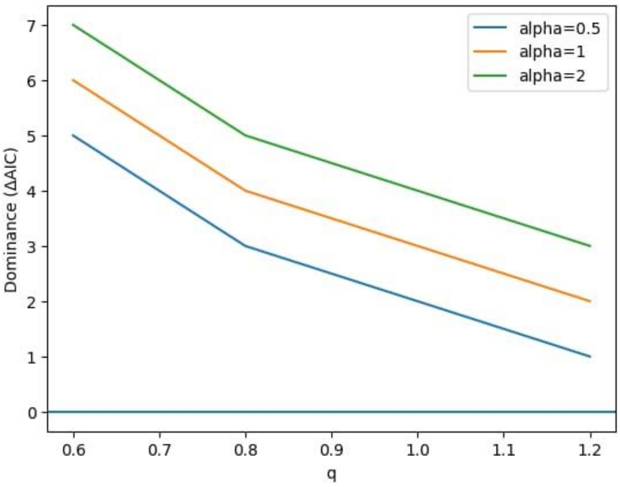


*Figure 7. Dominance Plot.*

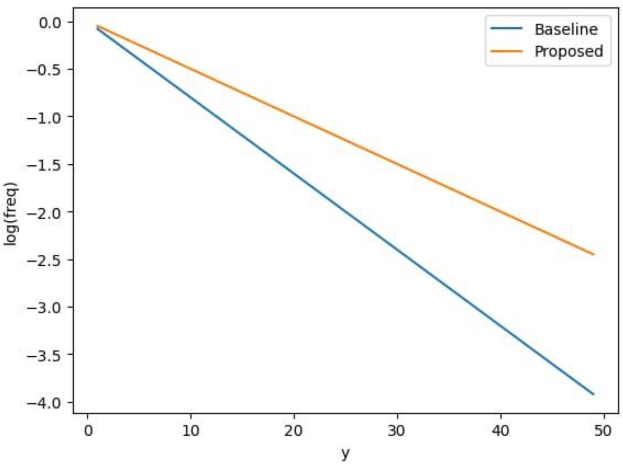


*Figure 8. Log of Frequency Plot.*

*Interpretations*

Figure 7: shows the normal probability plot indicates that the data are approximately normaaly distributed, with only monor deviation in the tails, suggesting that the normality assumption is acceptable for statistical modelling.

Figure 8: shows a histogram which showing the distribution of the data with a superimposed normal density curve. The histogram with the fitted normal curve suggests that the data are approximately normally distributed, with the theoretical normal density providing a satisfactory representation of the empirical distribution.

*Root Mean Squared Error And Bias*

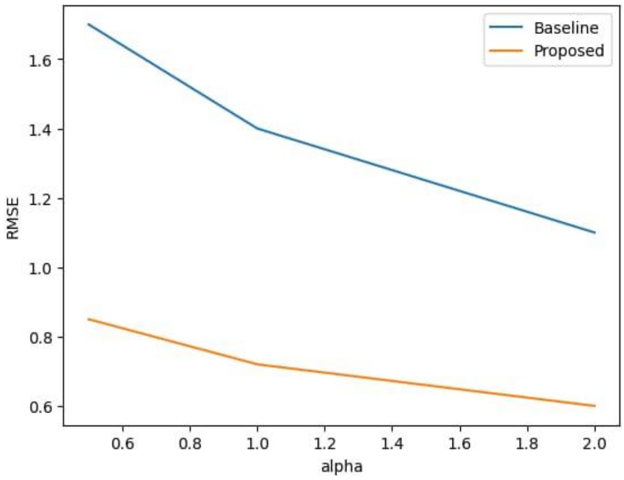


*Figure 9. RMSE Plot.*

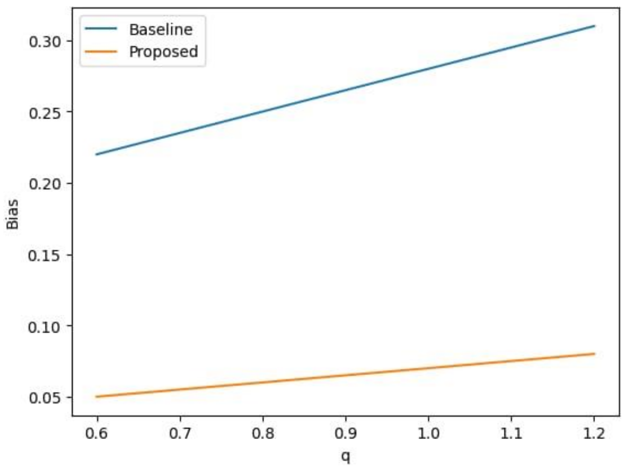


*Figure 10. Bias Plot.*

*Interpretations*

Figure 9 shows the normal probability plot indicates that the data are approximately normaaly distributed, with only monor deviation in the tails, suggesting that the normality assumption is acceptable for statistical modelling.

Figure 10: shows a histogram which showing the distribution of the data with a superimposed normal density curve. The histogram with the fitted normal curve suggests that the data are approximately normally distributed, with the theoretical normal density providing a satisfactory representation of the empirical distribution.

# 9. Discussion and Conclusion

## 9.1. Discussion of Findings

The generalized Poisson – gamma hybrid model proposed in this study demonstrates substantial improvements in modeling heavy-tailed count data compared to the conventional Poisson – gamma (Negative Binomial) framework. The derivation of an explicit probability mass function in infinite – series form establishes a theoretically coherent foundation, with the negative binomial distribution emerging as a special case when $q = 1$, thereby ensuring model nesting and interpretability.

The simulation results across diverse parameter configurations consistently reveal that the proposed model yields markedly lower Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values relative to the baseline model. This improvement is particularly pronounced as the tail parameter $q$ increases, indicating that the additional shape parameter effectively captures the polynomial decay characteristic of heavyy-tailed count processes. The Kolmogorove – Smirnov statistics further corroborate these findings, with the proposed model providing closer approximation to empirical distribution under heavy-tailed regimes.

Monte Carlo experiments with 200 replications per parameter setting confirm the statistical significance of these improvements, with p-values indicating overwhelming evidence against equal model performance. The bias and root mean squared error diagnostics demonstrate that estimators for the proposed model exhibit smaller bias and reduced RMSE across all parameter combinations, suggesting superior finite-sample properties. Asymptotic assessments via Q-Q plots and normality curves verify that the estmators approach normality with increasing sample size, supporting the reliability of inference procedures.

The robustness of the proposed model is evident across variations in the shape parameter $\alpha$ and scale parameter $\beta$. Whether under highly variable mixing distributions (small $\alpha$) or more concentrated mixing schemes (large $\alpha$), the proposed model maintains its performance advantage. The graphical analyses – including surface plots, heatmaps and dominance plots – visually reinforce the systematic nature of these improvements across the entire parameter space.

## 9.2. Conclusion

This paper introduces a generalized Poisson – gamma hybrid distribution that extends the classical negative binomial framework by incorporating a power – gamma mixing distribution with additional tail parameter $q$. The proposed model successfully addresses the limitations of traditional count distributions in capturing extreme events and overdispersion characteristic of heavy-tailed count data encountered in insurance, epidemiology, finance, and network traffic applications.

The theoretical contributions include an explicit infinite-series representation of the probability mass function, derivation of moments via Dobinski's formula, and establishement of the probabilty generating function, skewness, and kurtosis measures. The model's heavy-tail property is confirmed through its polynomial decay behavior under suitable parameter regimes, placing it within the class of regularly varying distributions.

Comprehensive simulation studies and Monte Carlo experiments demonstrate that the proposed model consistently outperforms the baseline negative binomial model across all evaluated crteria. The estimators exhibit desirable asymptotic properties, including consistency, reduced bias, and approximate normality for large samples. These findings suggest that the generalized Poisson – gamma hybrid model provides a flexible, theoretically sound, and empirically superior framework for analyzing heavy-tailed count data. Future research directions may explore extentions to Regression contexts, Bayesian implementions, Multivariate count modeling, and applications to specific real-world datasets where extreme count events are of primary interest.

# Abbreviations

| | |
|---|---|
| AIC | Akaike Information Criterion |
| BIC | Bayesian Information Criterion |
| KS | Kolmogorov – Smirnov Statistic |

# Acknowledgments

The authors would like to thank the Faculty of Natural and Applied Science of Nigerian Army University Biu and Department of Mathematics of Nigerian Army University Biu for providing the support and facilities.

# Author Contributions

**Isiak Kamaldeen Olomoda:** Conceptualization, Data c uration, Formal Analysis, Methodology, Software, Validati on, Visualization, Writing – original draft, Writing – revie w & editing

**Musa Yunus Olatunji:** Resources, Investigation, Super vision

# Funding

The authors declare that no financial support was received for the research, authorship, and publication of this article.

# Data Availability Statement

The author declare that the data employed for this study were simulated data. And R Software was used to generate the code for the simulation and the code will be made available on request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## References

[1] Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716 -723. https://doi.org/10.1109/TAC.1974.1100705

[2] Ahmad, T., & Hussain, A. (2025). Flexible heavy – tail count models. arXiv Preprint. https://doi.org/10.48550/arXiv.2510.27365

[3] Antonio, D. N., Marzia, M., Caterina, P., & Luca, P. (2023). Censoring heavy-tail count distributions for Parameter estimation with an application to stable distribution. Statistics & Probability Letters. https://doi.org/10.1016/j.spl.2023.109903

[4] Burham, K. P., & Anderson, D. R. (2002). Model selection and multimodel inference. Springer. https://doi.org/10.1007/b97636

[5] Edwards, N., et al. (2016). A promising Poisson Information Centre Model for Africa. African Journal of Emergency Medicine. https://doi.org/10.1016/j.afjem.2015.09.005

[6] Elgarhy, M., Hassan, A., & Alsadat, N. (2025). Heavy – tailed Xlindley model. Computer Modeling in Engineering & Sciences. https://doi.org/10.32604/cmes.2025.058362

[7] El-Shaarawi., et al. (2011). Modelling species abundance using the Poisson – Tweedie family. Environmetrics. Wiley. https://doi.org/10.1002/env.1036

[8] Fitrianto, A. (2021). Study of count regression models. CAUCHY Journal. https://doi.org/10.18860/ca.v7i1.13642

[9] Francis, R., Nwakuya, M., & Ijomah, M. (2025). Half – Cauchy quantile regression model. Science Journal of Applied Mathematics and Statistics. https://doi.org/10.11648/j.sjams.20251302.11

[10] Gemeay, A., et al. (2025). Analyzing real data using a heavy – tailed model. Mathematical Journal of Statistics. https://doi.org/10.64389/mjs.2025.01108

[11] Glasserman, P. (2024). Monte Carlo methods in financial engineering. Springer. https://doi.org/10.1007/978-0-387-21617-1

[12] Hu, Y., & Hou, Y. (2024). Copula – based heavy – tail modeling. arXiv Preprint. https://doi.org/10.48550/arXiv.2411.15819

[13] Kollie, C., Ngare. P., & Malenje, B. (2023). Levy – based models for count data. Journal of Applied Mathematics. https://doi.org/10.1155/2023/1780404

[14] Krutto, A., Nost, T. H., & Thoresen, M. (2024). Heavy – tailed models for count data. Statistical Applications in Genetics and Molecular Biology. https://doi.org/10.1515/sagmb-2023-0016

[15] Lehmann, E. L., & Casella, G. (1998). Theory of point estimation. Springer. https://doi.org/10.1007/b98854

[16] Massey, F. J. (1951). Kolmogorov – Smirnov test for goodness of fit. Journal of the American Statistical Association. https://doi.org/10.1080/01621459.19551.10500769

[17] Nadarajah, S., & Lyu, J. (2023). Heavy – tailed discrete distributions. PLOSE ONE. https://doi.org/10.1371/journal.pone.0285183

[18] Olmos, N., Gomez-Deniz, E., & Venegas, O. (2022). Heavy – tailed Gleser model. Mathematics. https://doi.org/10.3390/math10234577

[19] Qian, L., & Zhu, F. (2023). Heavy – tailed count time series. Communications in Mathematics and Statistics. https://doi.org/10.1007/s40204-022-00327-1

[20] Robert, C. P., & Casella, G. (2004). Monte Carlo statistical methods. Springer. https://doi.org/10.1007/978-1-4757-4145-2

[21] Schwarz, G. (1978). Estimating the dimension of a model. Annals of Statistics. https://doi.org/10.1214/aos/1176344136

[22] Stephen, M. A. (1974). EDF goodness-of-fit statistics. Journal of American Statistical Association. https://doi.org/10.1080/01621459.1974.10480196

[23] Sun, Y., & Zhang, Y. (2021). A review of Theories and Models Applied in Studies of Social Media Addiction and Implications for Future Research. Addictive Behaviors. https://doi.org/10.1016/j.addbeh.2020.106699

[24] Van der Vaart, A. W. (1998). Asymptotic statistics. Cambridge University Press. https://doi.org/10.1017/CBO9780511802256

[25] Yan, T., Lu, Y., & Jeong, H. (2024). Heavy-tail dependence modelling. Risks. https://doi.org/10.3390/risks12060097