Letters in Nonlinear Analysis and its Applications **4** (2026) No. 1, 14-28
https://doi.org/10.66147/lnaa.20264169
Available online at www.lettersinnonlinearanalysis.com
Research Article

Letters in Nonlinear Analysis and its Applications
Peer Review Scientific Journal
ISSN: 2958-874x

A mixed gamma distribution model: Theoretical insights and applications to medical data

Nawarat Ekkarntrong[a,*], Kullathida Ngamsa-nga[a], Parichat Sopa[a]

[a]*Department of Mathematics, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand*

Abstract

This paper proposes a novel mixed gamma distribution (NMGD) and systematically investigates its statistical properties, including moments, shape parameters, coefficient of variation, moment-generating function, survival function, and hazard function. Parameter estimation is addressed using both the method of moments and maximum likelihood estimation. The practical utility of the NMGD is demonstrated through applications to three empirical datasets: (i) recovery times (in minutes) for patients administered pain relievers, (ii) recovery durations (in months) for breast cancer patients treated with Trastuzumab, and (iii) monthly tax revenue data from Egypt spanning 59 months. Goodness-of-fit is evaluated using the Kolmogorov-Smirnov test, while model selection is guided by Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC). The results indicate that the NMGD provides a superior fit compared to conventional distributions across all datasets, underscoring its potential as a flexible and robust model for medical and financial data analysis.

Keywords: lifetime distribution, waiting time data, cumulative distribution, medical data
2010 MSC: 62E15, 62F10, 62N01, 62P05, 60E05

*Corresponding author
Email addresses: nawaek@kku.ac.th (Nawarat Ekkarntrong), kullathida.n@kkumail.com (Kullathida Ngamsa-nga), parichat.sop@kkumail.com (Parichat Sopa)

Received November 25, 2025, Accepted: March 23, 2026, Online: March 25, 2026

1. Introduction

Over the past few decades, the statistical literature has seen the development of numerous new probability distributions, many constructed via the finite mixture method. Notable examples include the Akask Distribution [1], Shanker Distribution [2], Aradhana Distribution [3], Lindley Distribution [6], Amarendra Distribution [9], and Sujatha Distribution [8]. These are typically one-parameter models and have found applications in reliability, survival, and lifetime data analysis.

Gamma distributions are widely utilized in diverse fields such as queueing systems (e.g. customer waiting times in banks), hydrology (rainfall modelling), medical statistics, and in modeling the waiting time for the arrival of k customers at a store. For example, Ghitany, Atieh, and Nadarajah (2008) [6] introduced the Lindley Distribution, derived from a mixture of $Gamma(1, \theta)$ and $Gamma(2, \theta)$, to model customer waiting times at banks; they demonstrated that the Lindley Distribution outperformed the Exponential Distribution. In 2017, Shanker [10] introduced the Rani Distribution, derived from a mixture of $Gamma(1, \theta)$ and $Gamma(5, \theta)$, which was applied to the strength data of airplane window glass. The Rani Distribution was shown to be more suitable than the Akask, Rama, and Amarendra Distributions.

Further, Messaadia and Zeghdoudi (2018) [5] developed the Zeghdoudi Distribution from a mixture of $Gamma(2, \theta)$ and $Gamma(3, \theta)$ for engineering and medical lifespan data, outperforming the Aradhana, Akask, and Shanker Distributions. Onyekwere and Obulezi (2022) [7] proposed the Chris-Jerry Distribution, a mixture of $Gamma(1, \theta)$ and $Gamma(3, \theta)$, for modeling life insurance premiums in Nigeria, yielding better results than Aradhana and Akask Distributions. More recently, Echebiri and Mbegbu (2022) [4] introduced the Juchez Probability Distribution, arising from a mixture $Gamma(1, \theta)$, $Gamma(2, \theta)$, and $Gamma(4, \theta)$, showing superior fit compared to Sujatha, Aradhana, and Amarendra Distributions.

Motivated by this literature, the present research proposes a New Mixed Gamma Distribution constructed through a finite mixture of $Gamma(4, \theta)$ and $Gamma(5, \theta)$. This work aims to systematically develop and characterize the distribution, including the probability density function (pdf), cumulative distribution function (cdf), shape analysis, hazard function, moments(origin and central), variation, skewness, kurtosis, and moment-generating function (mgf). Parameter estimation is performed using the method of moments and the maximum likelihood method. The utility of the NMGD is subsequently demonstrated through its application to three empirical datasets:

1. Recovery time (in minutes) from pain for 20 patients who received painkillers.
2. Recovery time (in months) from breast cancer for 50 women treated with trastuzumab.
3. Monthly tax revenue in Egypt (in billion Egyptian pounds) from January 2006 to November 2010, covering 59 months.

The goodness-of-fit of the model will be assessed using the Kolmogorov-Smirnov Test (K-S test), and its comparative performance is evaluated using Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

The remainder of this paper is organized as follows: Section 2 presents the structural properties of the proposed distribution, including its probability density function (pdf), cumulative distribution function (cdf), survival and hazard functions, moments, coefficient of variation, skewness, kurtosis, and moment-generating function (mgf). Section 3 discusses parameter estimation methods, specifically the method of moments and maximum likelihood estimation. Section 4 contains the results of the simulation study and real-data applications, while Section 5 concludes the paper and summarizes the main findings.

2. Structural properties

Let the random variable X follow a mixture of Gamma distributions defined as follows:

$X_1 \sim Gamma(4, \theta)$ with pdf $f_1(x) = \dfrac{\theta^4 x^3 e^{-\theta x}}{6}$ $, x > 0, \theta > 0,$

$X_2 \sim Gamma(5, \theta)$ with pdf $f_2(x) = \dfrac{\theta^5 x^4 e^{-\theta x}}{24}$ $, x > 0, \theta > 0.$

Using the finite mixtures approach, we define the mixture probability $p = \dfrac{\theta}{\theta+24}$ and thus the pdf of the random variable X is

$$f_X(x) = pf_1(x) + (1-p)f_2(x) = \frac{\theta}{\theta+24}\frac{\theta^4 x^3 e^{-\theta x}}{6} + \frac{24}{\theta+24}\frac{\theta^5 x^4 e^{-\theta x}}{24}.$$

Simplifying, we obtain

$$f_X(x) = \frac{\theta^5}{6(\theta+24)}(x^3+6x^4)e^{-\theta x} \quad ,x>0,\theta>0. \tag{1}$$

The cumulative distribution function (cdf) of X is derived as:

$$F_X(x) = 1 - \left(1 + \frac{\theta x(144+6\theta+(72+3\theta)\theta x+(24+\theta)\theta^2x^2+6\theta^3x^3)}{6(\theta+24)}\right)e^{-\theta x} \quad ,x>0,\theta>0.$$

This defines the New Mixed Gamma Distribution (NMGD), denoted by $X \sim NMGD(\theta)$, $\theta > 0$. Figures 1 and 2 illustrate the shapes of the pdf and cdf, respectively, of the NMGD for varying values of θ.

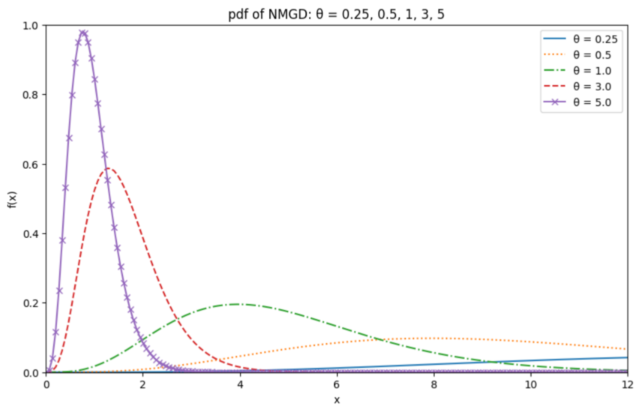

Figure 1: Graph of the probability density function of the new mixed Gamma distribution with varying parameters.

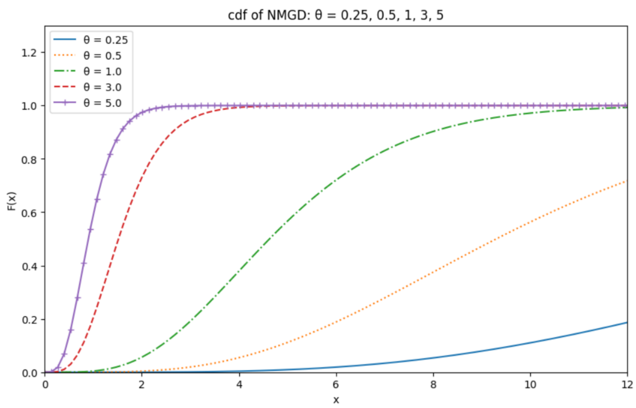

Figure 2: Graph of the cumulative distribution function of the new mixed Gamma distribution with varying parameters.

2.1. Survival Function and Hazard Function

For the NMGD, the survival function is given by

$$
\begin{aligned}
S(x) &= 1 - F(x) \\
&= \left(1 + \frac{\theta x(144 + 6\theta + (72 + 3\theta)\theta x + (24 + \theta)\theta^2 x^2 + 6\theta^3 x^3)}{6(\theta + 24)}\right) e^{-\theta x}.
\end{aligned}
$$

Figure 3 displays the survival function $S(x)$ for different values of θ.

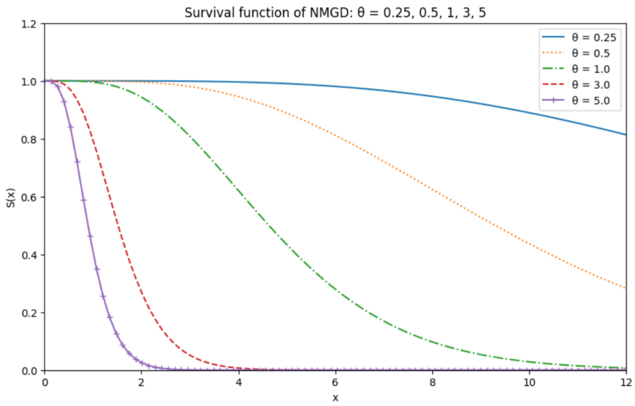

Figure 3: Graph of the survival function of the New Mixed Gamma Distribution for different parameters

For a continuous distribution with pdf $f(x)$ and survival function $S(x)$, the hazard function is defined by

$$
h(x) = \lim_{\Delta x \to 0} \frac{P(X < x + \Delta x \mid X > x)}{\Delta x} = \frac{f(x)}{S(x)}.
$$

Proposition 2.1 (Hazard Function)**.** *For $X \sim \mathrm{NMGD}(\theta)$,*

$$
h(x) = \frac{\theta^5 (x^3 + 6x^4)}{6\theta + 144 + \theta x\big(144 + 6\theta + (72 + 3\theta)\theta x + (24 + \theta)\theta^2 x^2 + 6\theta^3 x^3\big)}.
$$

Moreover, $h(0) = 0$, $h(x)$ is increasing in x, and $0 < h(x) < \theta$ for $x > 0$.

Proof. As $h'(x) > 0$ for all $x > 0$, h(x) is strictly increasing; together with $\lim_{x\to 0^+} h(x) = 0$ and $\lim_{x\to\infty} h(x) = \theta$, this yields $0 < h(x) < \theta$ for all $x > 0$. □

Figure 4 shows the behavior of the hazard function $h(x)$ for several choices of θ.

2.2. Moments and Related Measures

Having characterized the basic distributional functions of the NMGD, we now derive its general k-th raw and central moments together with the corresponding moment generating function. These moment expressions and the MGF provide the basis for the subsequent analysis of key shape measures, namely the coefficient of variation, skewness, and kurtosis.

Theorem 2.2 (Raw Moments)**.** *Let $X \sim NMGD(\theta)$ where $\theta > 0$. The r-th raw moment of the new mixed Gamma distribution is given by:*

$$
\mu_r{}' = E[X^r] = \frac{(r + 3)!(\theta + 6r + 24)}{6\theta^r(\theta + 24)}, \quad r = 1, 2, 3, \ldots \tag{2}
$$

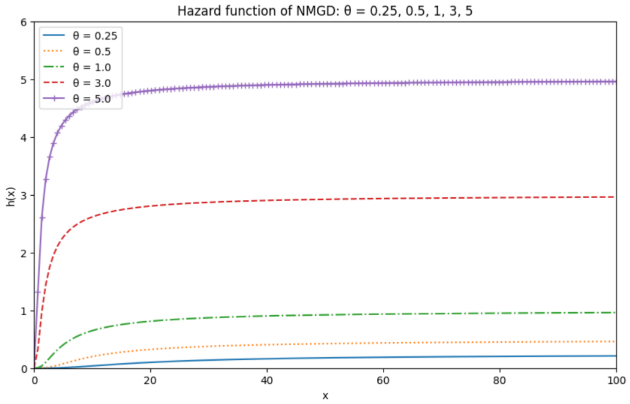

Figure 4: Graph of the hazard function of the New Mixed Gamma Distribution for different parameters

Proof. Using the definition of the r-th raw moment,

$$\mu_r{}' = E[X^r] = \int_0^{\infty} x^r \frac{\theta^5}{6(\theta+24)}(x^3+6x^4)e^{-\theta x}dx, \quad r = 1, 2, 3, \dots$$

and the identity $\int_0^{\infty} x^k e^{-\theta x}\, dx = \Gamma(k+1)/\theta^{k+1}$, we obtain

$$\mu_r' = \frac{\theta^5}{6(\theta+24)}\left(\frac{\Gamma(r+4)}{\theta^{r+4}} + 6\frac{\Gamma(r+5)}{\theta^{r+5}}\right) = \frac{(r+3)!\,(\theta+6r+24)}{6\,\theta^r\,(\theta+24)}.$$

□

For $r = 1, 2, 3, 4$, the raw moments are derived as follows:

$$\mu_1{}' = \frac{4(\theta+30)}{\theta(\theta+24)} = \mu \tag{3}$$

$$\mu_2{}' = \frac{20(\theta+36)}{\theta^2(\theta+24)} \tag{4}$$

$$\mu_3{}' = \frac{120(\theta+42)}{\theta^3(\theta+24)} \tag{5}$$

$$\mu_4{}' = \frac{840(\theta+48)}{\theta^4(\theta+24)} \tag{6}$$

Theorem 2.3 (Central Moments)**.** *Let $X \sim NMGD(\theta)$ where $\theta > 0$. The k-th central moment of the new mixed Gamma distribution is given by:*

$$\mu_k = E[(X-\mu)^k] = \sum_{r=0}^{k}\binom{k}{r}(-\mu)^{k-r}\frac{(r+3)!(\theta+6r+24)}{6\theta^r(\theta+24)} \quad k = 2, 3, 4, \dots. \tag{7}$$

Proof. Using the definition of the r-th central moment,

$$\mu_r{}' = E[(X-\mu)^r] = \int_0^{\infty} (x-\mu)^r \frac{\theta^5}{6(\theta+24)}(x^3+6x^4)e^{-\theta x}dx,$$

and the binomial expansion, we obtain

$$\begin{aligned}
{\mu_r}' &= \sum_{r=0}^{k} \binom{k}{r} (-\mu)^{k-r} \frac{\theta^5}{6(\theta+24)} \left(\frac{\Gamma(r+4)}{\theta^{r+4}} + \frac{6\Gamma(r+5)}{\theta^{r+5}} \right) \\
&= \sum_{r=0}^{k} \binom{k}{r} (-\mu)^{k-r} \frac{(r+3)!(\theta+6r+24)}{6\theta^r(\theta+24)}, \quad k = 2, 3, 4, \ldots
\end{aligned}$$

□

For $k = 2, 3, 4$, the central moments are derived as follows:

$$\mu_2 = \frac{4(\theta^2 + 60\theta + 4,095)}{\theta^2(\theta+24)^2} = \sigma^2, \tag{8}$$

$$\mu_3 = \frac{8(\theta^3 + 90\theta^2 + 2,160\theta + 17,280)}{\theta^3(\theta+24)^3}, \tag{9}$$

$$\mu_4 = \frac{72(\theta^4 + 120\theta^3 + 4,800\theta^2 + 80,640\theta + 483,840)}{\theta^4(\theta+24)^4}. \tag{10}$$

Theorem 2.4 (Moment Generating Function)**.** *For $X \sim \mathrm{NMGD}(\theta)$, the MGF for $t < \theta$ is*

$$M_X(t) = E[e^{tX}] = \frac{\theta^5(\theta - t + 24)}{(\theta+24)(\theta-t)^5}$$

Proof. By the definition of the moment generating function,

$$\begin{aligned}
M_X(t) &= E[e^{tX}] \\
&= \int_0^\infty e^{tx} \frac{\theta^5}{6(\theta+24)} (x^3 + 6x^4) e^{-\theta x} dx \\
&= \frac{\theta^5}{6(\theta+24)(\theta-t)^5} [(\theta-t)\Gamma(4) + 6\Gamma(5)], \qquad ,t < \theta \\
&= \frac{\theta^5(\theta - t + 24)}{(\theta+24)(\theta-t)^5}.
\end{aligned}$$

□

2.2.1. Coefficient of Variation, Kurtosis and Skewness

The explicit expressions for the second, third, and fourth central moments yield closed-form expressions for the classical shape measures: coefficient of variation, skewness, and kurtosis.

Theorem 2.5 (Coefficient of variation)**.** *If $X \sim \mathrm{NMGD}(\theta)$, then its coefficient of variation is*

$$\gamma = \frac{\sqrt{\theta^2 + 60\theta + 4095}}{2(\theta+30)},$$

where γ is strictly decreasing in $\theta > 0$ and satisfies

$$\frac{1}{2} < \gamma < \frac{\sqrt{455}}{20}.$$

Proof. By definition, $\gamma = \dfrac{\sigma}{\mu}$, which for $X \sim \mathrm{NMGD}(\theta)$ yields

$$\gamma = \frac{\sqrt{\theta^2 + 60\theta + 4095}}{2(\theta + 30)}.$$

A direct differentiation shows that $\dfrac{d\gamma}{d\theta} < 0$ for all $\theta > 0$, hence γ is strictly decreasing in θ. Moreover,

$$\lim_{\theta \to 0^+} \gamma = \frac{\sqrt{455}}{20}, \qquad \lim_{\theta \to \infty} \gamma = \frac{1}{2}.$$

Since γ is continuous and strictly decreasing on $(0, \infty)$, it follows that

$$\frac{1}{2} < \gamma < \frac{\sqrt{455}}{20}, \qquad \theta > 0.$$

□

Theorem 2.6 (Skewness)**.** *If $X \sim \mathrm{NMGD}(\theta)$, then its skewness coefficient is*

$$\beta_1 = \frac{\theta^3 + 90\theta^2 + 2160\theta + 17280}{\left(\theta^2 + 60\theta + 4095\right)^{3/2}},$$

where β_1 is strictly increasing in $\theta > 0$ and satisfies

$$\frac{128}{91}\sqrt{\frac{1}{455}} < \beta_1 < 1.$$

Proof. By the definition, $\beta_1 = \dfrac{\mu_3}{\sigma^3}$, we obtain

$$\beta_1 = \frac{\theta^3 + 90\theta^2 + 2,160\theta + 17,280}{(\theta^2 + 60\theta + 4,095)^{3/2}}.$$

Differentiating this expression with respect to θ and simplifying shows that $\beta_1'(\theta) > 0$ for all $\theta > 0$, so β_1 is strictly increasing in θ. Furthermore,

$$\lim_{\theta \to 0^+} \beta_1(\theta) = \frac{128}{91\sqrt{455}}, \qquad \lim_{\theta \to \infty} \beta_1(\theta) = 1.$$

Since β_1 is continuous and increasing on $(0, \infty)$, the stated bounds follow. □

Theorem 2.7 (Kurtosis)**.** *If $X \sim \mathrm{NMGD}(\theta)$, then its kurtosis coefficient is*

$$\beta_2 = \frac{9\left(\theta^4 + 120\theta^3 + 4800\theta^2 + 80640\theta + 483840\right)}{2\left(\theta^2 + 60\theta + 4095\right)^2},$$

where β_2 is strictly increasing in $\theta > 0$ and satisfies

$$\frac{768}{5915} < \beta_2 < \frac{9}{2}.$$

Proof. Using $\beta_2 = \mu_4/\sigma^4$ for $X \sim \mathrm{NMGD}(\theta)$ gives the stated closed form for $\beta_2(\theta)$. A straightforward differentiation shows that $\beta_2'(\theta) > 0$ for all $\theta > 0$, hence β_2 is strictly increasing in θ. Furthermore,

$$\lim_{\theta \to 0^+} \beta_1(\theta) = \frac{768}{5915}, \qquad \lim_{\theta \to \infty} \beta_1(\theta) = \frac{9}{2}.$$

which, together with monotonicity, yields the claimed bounds. □

Figure 5 illustrates the kurtosis coefficient $\beta_2(\theta)$ of the NMGD, confirming its monotone increase in θ and the bounds stated in Theorem 2.7.

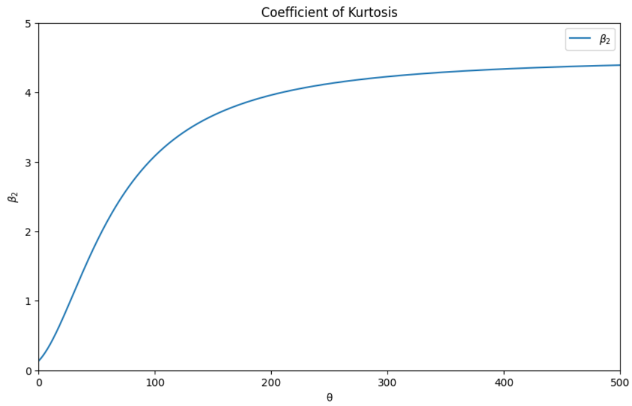

Figure 5: Coefficient of Kurtosis for the New Mixed Gamma Distribution

3. Parameter Estimation

Let the random sample $X_1, X_2, \ldots, X_n$ follow the New Mixed Gamma Distribution, and let $x_1, x_2, \ldots, x_n$ be the observed values. In this section, we develop point estimators for the parameter θ using both the method of moments and maximum likelihood, and then investigate the bias of the resulting estimator.

3.1. Method of Moments (MOME)

The first sample moment is given by:

$$M_1' = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}.$$

From Equation (3), the population mean is:

$$\mu_1' = E[X] = \frac{4(\theta+30)}{\theta(\theta+24)}.$$

Equating the sample and population moments, we solve the following equation for $\hat{\theta}$, the estimate of θ using the method of moments:

$$\overline{X} = \frac{4(\hat{\theta}+30)}{\hat{\theta}(\hat{\theta}+24)}.$$

Thus, the method of moments estimate of θ is:

$$\hat{\theta} = \frac{-(24\overline{X}-4)+\sqrt{(24\overline{X}-4)^2+480\overline{X}}}{2\overline{X}}, \quad \overline{X} > 0.$$

3.2. Maximum Likelihood Estimation (MLE)

As an alternative to the method of moments, we now derive the maximum likelihood estimator of θ. The likelihood function is given by:

$$L(\theta; x) = \prod_{i=1}^{n} \frac{\theta^5}{6(\theta+24)}(x_i^3 + 6x_i^4)e^{-\theta x_i}, \quad x, \theta > 0.$$

The maximum likelihood estimate of θ is:

$$\hat{\theta} = \frac{-(24\overline{X}-4)+\sqrt{(24\overline{X}-4)^2+480\overline{X}}}{2\overline{X}}, \quad \overline{X} > 0.$$

Remark 3.1. The estimates of θ obtained using both the method of moments (MOME) and maximum likelihood estimation (MLE) are identical:

$$\hat{\theta} = \frac{-(24\overline{X} - 4) + \sqrt{(24\overline{X} - 4)^2 + 480\overline{X}}}{2\overline{X}}, \quad \overline{X} > 0.$$

This common estimator will be the focus of our subsequent bias analysis.

Theorem 3.2. *The estimator $\hat{\theta}$ of θ is positively biased, that is, $E[\hat{\theta}] - \theta > 0$.*

Proof. Let $\hat{\theta} = g(\overline{X})$, where

$$g(t) = \frac{-(24t - 4) + \sqrt{(24t - 4)^2 + 480t}}{2t}, \qquad t > 0.$$

Since $g''(t) > 0$ for all $t > 0$, the function g is strictly convex. By Jensen's inequality,

$$E[g(\overline{X})] > g(E[\overline{X}]).$$

Using $E[\overline{X}] = \mu$ and $g(\mu) = \theta$, we obtain $E[\hat{\theta}] = E[g(\overline{X})] > \theta$. Hence $\hat{\theta}$ is positively biased. □

4. Simulation and Applications

4.1. Simulation Study

To complement the theoretical properties derived in the previous sections, we conduct a Monte Carlo study to assess the finite-sample performance of the maximum likelihood estimator of θ for the NMGD. The design explores a range of parameter values and sample sizes to evaluate the bias, variability, and overall accuracy of the estimator in practical settings.

Simulation Setup:

- Synthetic datasets are generated from the NMGD.
- The true parameter values considered are $\theta = 0.1, 0.25, 0.5, 1, 3, 5, 10$.
- The sample sizes examined are $n = 20, 40, 70, 100, 200, 500$.
- The unknown parameter θ is estimated from each simulated dataset using the Maximum Likelihood Estimation (MLE) method.
- For each combination of parameter value and sample size, the following summary measures are computed over multiple simulation replications:
  - The mean of the estimated parameter.
  - The mean bias, defined as the average deviation of the estimator from the true parameter value.
  - The mean squared error (MSE) of the estimator, reflecting both bias and estimation variability.
- The purpose of this simulation is to investigate the finite-sample behavior, bias characteristics, and efficiency of the MLE estimator for the NMGD across a range of conditions.

The simulation study evaluated the performance of maximum likelihood estimation (MLE) for the parameter θ of the proposed New Mixed Gamma Distribution (NMGD) across various sample sizes and true parameter values. The results are summarized in Table 1, which presents the mean estimates, biases, and mean squared errors (MSEs) for θ under different scenarios.

The findings reveal that the MLE exhibits a slight positive bias in all examined cases. As the sample size increases, the bias diminishes notably, approaching negligible levels for larger samples, which suggests

Sample size	$\theta = 0.1$			$\theta = 0.25$			$\theta = 0.5$			$\theta = 1$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
20	0.1010	0.0010	0.0001	0.2527	0.0027	0.0007	0.5058	0.0058	0.0026	1.0109	0.0109	0.0103
40	0.1005	0.0005	0.0001	0.2512	0.0012	0.0003	0.5026	0.0026	0.0013	1.0062	0.0062	0.0051
70	0.1002	0.0002	0.0000	0.2506	0.0006	0.0002	0.5014	0.0014	0.0007	1.0029	0.0029	0.0029
100	0.1002	0.0002	0.0000	0.2505	0.0005	0.0001	0.5010	0.0010	0.0005	1.0013	0.0013	0.0020
200	0.1001	0.0001	0.0000	0.2504	0.0004	0.0001	0.5002	0.0002	0.0003	1.0012	0.0012	0.0010
500	0.1001	0.0001	0.0000	0.2501	0.0001	0.0000	0.5002	0.0002	0.0001	1.0003	0.0003	0.0004

(a) Mean, Bias, and MSE for $\theta = 0.1, 0.25, 0.5, 1$ and various sample sizes.

Sample size	$\theta = 3$			$\theta = 5$			$\theta = 10$		
	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
20	3.0318	0.0318	0.0942	5.0547	0.0547	0.2674	10.1098	0.1098	1.0560
40	3.0213	0.0213	0.0480	5.0287	0.0287	0.1292	10.0510	0.0510	0.5164
70	3.0113	0.0113	0.0264	5.0131	0.0131	0.0724	10.0205	0.0205	0.2953
100	3.0073	0.0073	0.0187	5.0106	0.0106	0.0505	10.0179	0.0179	0.2040
200	3.0026	0.0026	0.0092	5.0084	0.0084	0.0257	10.0119	0.0119	0.1012
500	3.0014	0.0014	0.0036	5.0009	0.0009	0.0101	10.0051	0.0051	0.0402

(b) Mean, Bias, and MSE for $\theta = 3, 5, 10$ and various sample sizes.

Table 1: Mean, Bias, and MSE for different θ and sample sizes (two-part presentation).

desirable asymptotic properties consistent with classical statistical theory. For example, at $n = 20$, the average estimated θ exceeds the true value by a small margin, e.g., approximately 0.0318 for $\theta = 3$, whereas at $n = 500$, the bias reduces to around 0.0014.

In terms of accuracy, the MSE consistently decreases as the sample size grows, indicating improved estimator precision. This pattern is evident across all parameter values—both low and high—highlighting the robustness of the MLE in finite samples. For instance, when $\theta = 1$, the MSE drops from approximately 0.0103 at $n = 20$ to less than 0.0005 at $n = 500$.

Furthermore, the bias appears relatively stable across the range of θ values, with a tendency to be minimally positive, which could be attributed to the inherent properties of the estimation process or the distribution's effects. Nonetheless, the decreasing trend of bias and MSE with increasing sample size underscores the consistency and efficiency of the MLE for the NMGD parameters.

Overall, these simulation results demonstrate that the MLE for θ under the NMGD is asymptotically unbiased and consistent, with improved finite-sample performance as the sample size increases. The practical implications suggest that the estimator performs reliably even for moderate sample sizes, reinforcing the utility of the proposed distribution in empirical applications involving medical and financial data.

4.2. Application Part

To demonstrate the practical usefulness of the proposed NMGD, we now fit the model to three real datasets from medical and economic contexts and compare its performance with several competing lifetime distributions. In all cases, the parameter θ is estimated by maximum likelihood.

Dataset 1 This first dataset comprises recovery times (in minutes) until pain relief for 20 patients treated with painkillers:

1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0.

Dataset 2 The second dataset contains recovery durations (in months) for 50 female breast cancer patients undergoing Trastuzumab treatment:

50, 74, 35, 39, 21, 37, 27, 35, 30, 35, 26, 38, 34, 34, 26, 41, 61, 33, 33, 26, 25, 41, 35, 34, 34, 33, 60, 61, 42, 30, 80, 31, 24, 49, 26, 31, 28, 41, 37, 41, 61, 33, 26, 34, 50, 73, 45, 80, 39, 21.

Dataset 3 The third dataset records monthly tax revenue (in billion Egyptian pounds) for Egypt from January 2006 to November 2010, covering 59 months:

5.9, 20.4, 14.9, 16.2, 17.2, 7.8, 6.1, 9.2, 10.2, 9.6, 13.3, 8.5, 21.6, 18.5, 5.1, 6.7, 17, 8.6, 9.7, 39.2, 35.7, 15.7, 9.7, 10, 4.1, 36, 8.5, 8, 9.2, 26.2, 21.9, 16.7, 21.3, 35.4, 14.3, 8.5, 10.6, 19.1, 20.5, 7.1, 7.7, 18.1, 16.5, 11.9, 7, 8.6, 12.5, 10.3, 10.3, 11.2, 6.1, 8.4, 11, 11.6, 11.9, 5.2, 6.8, 8.9, 7.1, 10.8.

Basic descriptive statistics for these datasets, including maximum, minimum, mean, and standard deviation, are summarized in Table 2.

Dataset	**Maximum**	**Minimum**	**Mean**	**Standard Deviation**
Dataset 1	4.1000	1.1000	1.9000	0.7041
Dataset 2	80.0000	21.0000	39.6000	14.8681
Dataset 3	39.2000	4.1000	13.4881	8.0515

Table 2: Summary statistics (maximum, minimum, mean, standard deviation) for the three datasets

4.2.1. Kolmogorov-Smirnov Test (K-S Test)

This section assesses the goodness-of-fit of the proposed New Mixed Gamma Distribution (NMGD) to the three real datasets using the Kolmogorov-Smirnov (K-S) test. The hypotheses for the test are defined as follows:

H_0: The data follows the NMGD with parameter $\hat{\theta}$.

H_1: The data does not follow the NMGD with parameter $\hat{\theta}$.

Table 3 summarizes the K-S test statistics, critical values at a 5% significance level, estimated parameters, and sample means for each dataset.

Dataset	$\hat{\theta}$	**Test Statistic** D_n	**Critical Value** c	$\hat{\mu}$
Dataset 1 (n =20)	2.5805	0.1737	0.3041	1.9000
Dataset 2 (n =50)	0.1261	0.1545	0.1923	39.6130
Dataset 3 (n =59)	0.3696	0.1588	0.1771	12.5570

Table 3: Kolmogorov-Smirnov goodness-of-fit test results for the NMGD applied to three datasets

The critical values c were calculated using the formula $c = \frac{1.36}{\sqrt{n}}$ at the significance level $\alpha = 0.05$, where n denotes the sample size. The results indicate that, across all datasets, the test statistic D_n is less than the corresponding critical value c, so the null hypothesis H_0 is not rejected. Specifically:

- For Dataset 1, $D_{20} = 0.1737 < 0.3041$, supporting the adequacy of the NMGD fit with $\hat{\theta} = 2.5805$.
- For Dataset 2, $D_{50} = 0.1545 < 0.1923$, indicating a good fit of the NMGD with $\hat{\theta} = 0.1261$.
- For Dataset 3, $D_{59} = 0.1588 < 0.1771$, confirming the NMGD suitably models the data with $\hat{\theta} = 0.3696$.

Overall, the K-S test results strongly support that the new mixed gamma distribution provides an adequate fit for all three empirical datasets at the 5% significance level.

4.2.2. Model Selection

This section presents a comparative analysis of the New Mixed Gamma Distribution (NMGD) against alternative candidate models for three empirical datasets. Model fit is evaluated using the Log-likelihood, Akaike's Information Criterion (AIC), and Bayesian Information Criterion (BIC).

The probability density functions (pdf) of the competing distributions for Dataset 1 are summarized in Table 4.

Table 5 reports the estimated parameters along with the Log-likelihood, AIC, and BIC values for each model fitted to Dataset 1.

The NMGD attains the highest Log-likelihood (-19.8850) and the lowest AIC (41.7700) and BIC (42.7658) values, indicating superior fit over competing models for Dataset 1. Figure 6 displays the frequency histogram of Dataset 1 alongside the pdf curves of all six fitted models, highlighting the NMGD's superior alignment with the data.

Model	Probability Density Function (pdf)
Size-biased Sechdistrib distribution (SBZD)	$f(x;\theta) = \dfrac{\theta^4}{2(\theta+3)}(x^2+x^3)e^{-\theta x}, \quad x,\theta > 0$
Sechdistrib distribution (ZD)	$f(x;\theta) = \dfrac{\theta^3}{\theta+2}(x+x^2)e^{-\theta x}, \quad x,\theta > 0$
Shanker distribution	$f(x;\theta) = \dfrac{\theta^2}{\theta^2+1}(\theta+x)e^{-\theta x}, \quad x,\theta > 0$
Xgamma distribution	$f(x;\theta) = \dfrac{\theta^2}{\theta+1}\left(1+\dfrac{\theta}{2}x^2\right)e^{-\theta x}, \quad x,\theta > 0$
Exponential distribution	$f(x;\theta) = \theta e^{-\theta x}, \quad x,\theta > 0$

Table 4: Probability density functions of models compared for Dataset 1

Model	$\hat{\theta}$	Log-likelihood	**AIC**	**BIC**
NMGD	2.5805	-19.8850	41.7700	42.7658
SBZD	1.9011	-22.1043	46.2087	47.2044
ZD	1.3654	-24.8594	51.7188	52.7145
Shanker	0.8387	-29.9329	61.8658	62.8616
Xgamma	1.1075	-31.5082	65.0165	66.0122
Exponential	0.5263	-32.8371	67.6742	68.6699

Table 5: Model comparison for Dataset 1 based on Log-likelihood, AIC, and BIC

The pdfs of the candidate distributions for Dataset 2 are listed in Table 6.

Table 7 provides the parameter estimates and model selection criteria for Dataset 2.

Again, NMGD outperforms all competitors in terms of the highest Log-likelihood and the lowest AIC and BIC values, confirming its superior flexibility and fit for Dataset 2. Figure 7 illustrates the histogram of Dataset 2 with overlaid pdfs of all nine models.

For Dataset 3, the pdfs of the candidate distributions are summarized in Table 8.

Table 9 summarizes parameter estimates and model selection criteria for Dataset 3.

Once more, the NMGD achieves the best fit, indicated by the highest Log-likelihood and lowest AIC and BIC values, outperforming all alternative models tested for Dataset 3. Figure 8 shows the frequency histogram of Dataset 3 with pdf curves of the six candidate models.

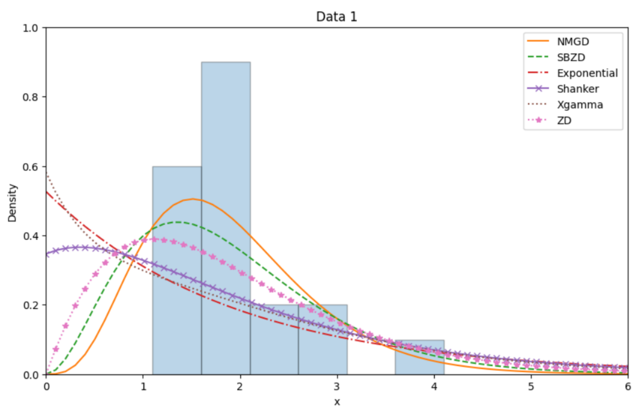

Figure 6: Frequency histogram of Dataset 1 with pdf overlays of candidate models

Model	Probability Density Function (pdf)
Juchez Distribution	$f(x;\theta) = \dfrac{\theta^4}{\theta^3+\theta^2+6}(1+x+x^3)e^{-\theta x}, \quad x,\theta > 0$
Amarendra Distribution	$f(x;\theta) = \dfrac{\theta^4}{\theta^3+\theta^2+2\theta+6}(1+x+x^2+x^3)e^{-\theta x}, \quad x,\theta > 0$
Akask Distribution	$f(x;\theta) = \dfrac{\theta^3}{\theta^2+2}(1+x^2)e^{-\theta x}, \quad x,\theta > 0$
Sujatha Distribution	$f(x;\theta) = \dfrac{\theta^3}{\theta^2+\theta+2}(1+x+x^2)e^{-\theta x}, \quad x,\theta > 0$
Aradhana Distribution	$f(x;\theta) = \dfrac{\theta^3}{\theta^2+2\theta+2}(1+x)^2e^{-\theta x}, \quad x,\theta > 0$
Shanker Distribution	$f(x;\theta) = \dfrac{\theta^2}{\theta^2+1}(\theta+x)e^{-\theta x}, \quad x,\theta > 0$
Lindley Distribution	$f(x;\theta) = \dfrac{\theta^2}{\theta+1}(1+x)e^{-\theta x}, \quad x,\theta > 0$
Exponential Distribution	$f(x;\theta) = \theta e^{-\theta x}, \quad x,\theta > 0$

Table 6: Probability density functions of models compared for Dataset 2

Model	$\hat{\theta}$	Log-likelihood	**AIC**	**BIC**
NMGD	0.1261	-202.2757	406.5515	408.4635
Juchez	0.1009	-205.1416	412.2831	414.1952
Amarendra	0.1001	-205.3995	412.7989	414.7109
Akask	0.0756	-209.7882	421.5764	423.4884
Sujatha	0.0747	-210.2409	422.4818	424.3938
Arahana	0.0738	-210.6499	423.2999	425.2119
Shanker	0.0504	-217.6236	437.2472	439.1592
Lindley	0.0493	-218.6144	439.2288	441.1409
Exponential	0.0252	-233.9416	469.8831	471.7952

Table 7: Model comparison for Dataset 2 based on Log-likelihood, AIC, and BIC

Model	Probability Density Function (pdf)
Chris-Jerry Distribution	$f(x;\theta) = \dfrac{\theta^2}{\theta+2}\left(1+\theta x^2\right)e^{-\theta x}, \quad x,\theta > 0$
Shanker Distribution	$f(x;\theta) = \dfrac{\theta^2}{\theta^2+1}(\theta+x)e^{-\theta x}, \quad x,\theta > 0$
Xgamma Distribution	$f(x;\theta) = \dfrac{\theta^2}{\theta+1}\left(1+\dfrac{\theta}{2}x^2\right)e^{-\theta x}, \quad x,\theta > 0$
Lindley Distribution	$f(x;\theta) = \dfrac{\theta^2}{\theta+1}(1+x)e^{-\theta x}, \quad x,\theta > 0$
Exponential Distribution	$f(x;\theta) = \theta e^{-\theta x}, \quad x,\theta > 0$

Table 8: Probability density functions of models compared for Dataset 3

Taken together, the goodness-of-fit tests and model selection criteria across three datasets indicate that the proposed NMGD is highly competitive and often superior to several well-known one-parameter lifetime models in real applications.

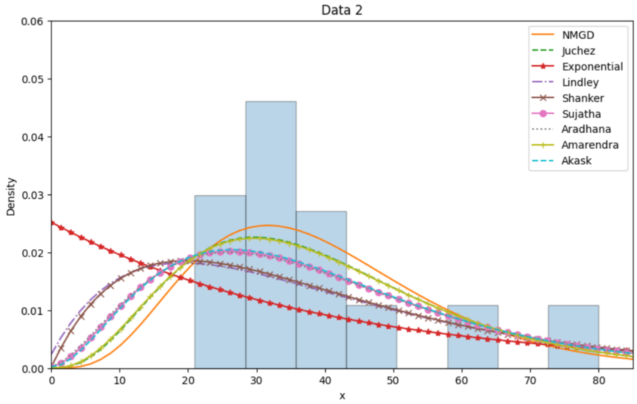

Figure 7: Frequency histogram of Dataset 2 with pdf overlays of candidate models

Model	$\hat{\theta}$	Log-likelihood	**AIC**	**BIC**
NMGD	0.3695	-194.7144	391.4288	393.5064
Chris-Jerry	0.2117	-196.8247	395.6493	397.7268
Shanker	0.1462	-198.5302	399.0605	401.1380
Xgamma	0.2029	-199.2838	400.5676	402.6451
Lindley	0.1392	-200.6293	403.2586	405.3361
Exponential	0.0741	-212.5068	427.0137	429.0912

Table 9: Model comparison for Dataset 3 based on Log-likelihood, AIC, and BIC

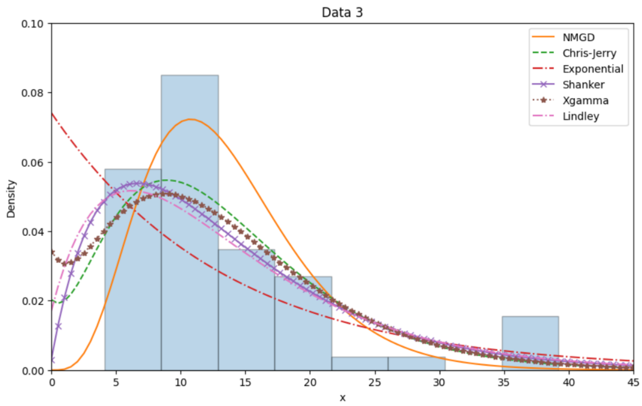

Figure 8: Frequency histogram of dataset 3 and PDF plots of the six distributions

5. Conclusion

This study introduced the New Mixed Gamma Distribution (NMGD). It rigorously examined its statistical properties, including raw moments, central moments, the moment generating function, the mean, the coefficient of variation, skewness, kurtosis, the survival function, and the hazard function.

Goodness-of-fit evaluations using the Kolmogorov-Smirnov (K-S) test confirmed that the NMGD provides an adequate fit for all three real datasets at the 5% significance level. Additionally, model selection criteria—Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC)—consistently favored the NMGD over alternative competing distributions across the datasets.

These results collectively demonstrate the flexibility and robustness of the New Mixed Gamma Distribution, highlighting its strong suitability for modeling recovery times and various medical-related data.

Acknowledgements

The corresponding author would like to thank the Faculty of Science, Khon Kaen University.

References

[1] R. Shanker, *Akash distribution and its applications*, Int. J. Probability and Statistics **4** (2015) 65–75.
[2] R. Shanker, *Shanker distribution and its applications*, Int. J. Statistics and Applications **5** (2015) 338–348.
[3] R. Shanker, *Aradhana distribution and its applications*, Int. J. Statistics and Applications **6** (2016) 23–34.
[4] U.V. Echebiri, J.I. Mbegbu, *Juchez probability distribution: properties and applications*, Asian J. Probability and Statistics **20**(2) (2022) 56–71.
[5] H. Messaadia, H. Zeghdoudi, *Zeghdoudi distribution and its applications*, Int. J. Computing Science and Mathematics **9**(1) (2018) 58–65.
[6] M.E. Ghitany, B. Atieh, S. Nadarajah, *Lindley distribution and its application*, Math. Comput. Simul. **78** (2008) 493–506.
[7] C.K. Onyekwere, O.J. Obulezi, *Chris-Jerry distribution and its applications*, Asian J. Probability and Statistics **20** (2022) 16–30.
[8] R. Shanker, *Sujatha distribution and its applications*, Statistics in Transition new series **17**(3) (2016) 391–410.
[9] R. Shanker, *Amarendra distribution and its applications*, Am. J. Mathematics and Statistics **6** (2016) 44–56.
[10] R. Shanker, *Rani distribution and its application*, Biometrics and Biostatistics Int. J. **6**(1) (2017).
[11] H. Anwar, Joarder, *An induction derivation of Stirling number of the second kind and their application in statistics*, J. Appl. Math. Decision Sci. **1** (1997) 151–157.
[12] J. Riordan, *Introduction to combinatorial analysis*, Courier Corporation, 1958.
[13] A. Rényi, *On measures of entropy and information*, Proc. 4th Berkeley Symp. Math. Stat. Probab. (1961).
[14] S.A. Klugman, H.H. Panjer, G.E. Willmot, *Loss Models*, Wiley, United States, 2019.
[15] N. Mukhopadhyay, *Probability and Statistical Inference*, CRC Press, United States, 2020.
[16] J.E. Peajcariaac, Y.L. Tong, *Convex Functions, Partial Orderings, and Statistical Applications*, Elsevier Science, Amsterdam, 1992.
[17] J. Stewart, *Calculus*, 7th ed., Cengage Learning, 2010.
[18] S. Ross, *A First Course in Probability*, 8th ed., University of Southern California: Library of Congress Cataloging, 2010.