EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
2025, Vol. 18, Issue 4, Article Number 6653
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

The Mgamma Distribution: Statistical Properties and Application to Real-Life Data

M. I. Khan[1], Molay Kumar Ruidas[2,*]

[1] *Department of Mathematics, Islamic University of Madinah, Madinah, Kingdom of Saudi Arabia, 42351*
[2] *Department of Statistics, Trivenidevi Bhalotia College, Kazi Nazrul University, Raniganj, West Bengal, India, 713347*

Abstract. We propose a new one-parameter continuous distribution, termed the *Mgamma distribution*, obtained as a finite mixture of two Gamma components having different shapes. Key statistical properties are derived, including moments, generating functions, stochastic ordering, and reliability measures such as hazard rate and mean residual life. Parameter estimation is addressed using the method of moments and maximum likelihood, with simulation studies confirming the consistency and efficiency of estimators. The applicability of the Mgamma distribution is demonstrated on four real datasets from engineering, hydrology, medicine, and queueing systems. Model selection criteria (AIC, BIC, CAIC) and goodness-of-fit tests (KS, Cramér–von Mises, Anderson–Darling) consistently show that Mgamma provides a superior fit compared to classical models such as Exponential, Lindley, Xgamma, Shanker, and Akash distributions. These findings establish Mgamma as a flexible and robust alternative for modeling skewed and heavy-tailed lifetime data.

2020 Mathematics Subject Classifications: 60E05, 62E99

Key Words and Phrases: Finite mixture, Gamma distribution, maximum likelihood estimation, survival properties, stochastic ordering

1. Introduction

In recent years, research has focused on developing continuous distributions that are more flexible in fitting real lifetime data sets. Some distributions that are finite mixtures of Gamma distributions introduced by Geoffrey et al. (1988) [1], mixture of Exponential and Gamma distributions introduced by Ghitany et al. (2008) [2] and Maiti et al. (2024) [3] used to model lifetime data. Recently, Sen et al. (2016) [4] introduced Xgamma, the Lindley distribution by DV Lindley (1958) [5], Shanker (2015) [6] introduced the Shanker distribution, Gamma-Lindley (Zeghdoudi and Nedjar, (2016) [7]) and several other mixture distributions are derived using the literature of Johnson, Kotz and Balakrishnan

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v18i4.6653

Email addresses: khanizhar@iu.edu.sa (M. I. Khan),
molayruidas@gmail.com (M. K. Ruidas)

(1994) [8] to fit real-life data. However, these distributions are not as sophisticated in modeling datasets more accurately. Various classical distributions are modified and extended to find new distributions like extended Gamma introduced by Bakouch (2012) [9], two-parameter Xgamma Weibull by Yousof et al. (2021) [10], Exponentiated survival function of Beta introduced by Mahmoud et al. (2024)[11], new beta power flexible Weibull (NBPF-Weibull) by Tang et al. (2024)[12], GE Weibull by kamal et al. (2023)[13] and extended Xgamma by Saha et al. (2019)[14], which serve the purpose of flexible data modelling. In this study, we introduce a new one-parameter continuous type of distribution using a similar approach used by Messaadia and Zeghdoudi (2018) [15] but taking a different mixing probability.

The rest of the paper is organized as follows. In Section 3, we define the probability density (pdf) and cumulative distribution (cdf) functions of the proposed distribution. Section 4 derives its statistical properties, including raw moments, mean deviation, moment and cumulant generating functions, skewness, and kurtosis. The stochastic ordering that characterizes the structural behavior of the distribution is discussed in Section 5. Reliability properties, such as the hazard function and mean residual life, are presented in Section 7. Section 8 provides results for the order statistics. Parameter estimation methods, including the method of moments and maximum likelihood estimation, are described in Section 9. A Monte Carlo simulation study investigating bias and mean square error of the estimators is given in Section 10. Applications of the proposed distribution to four real-life datasets are discussed in Section 11. Finally, Section 12 concludes the paper with a summary of findings and potential directions for future research.

2. Methodology

The new distribution is a finite mixture of two Gamma distributions having different shape parameters with certain mixing probabilities. The form of the finite mixture distribution is given by

$$g(x) = \sum_{i=1}^{k} \pi_i g(x_i).$$

Let $g_1(x) \sim Gamma(2, \theta)$ and $g_2(x) \sim Gamma(3, \theta)$ and $\pi_1 = \dfrac{\theta}{1+\theta}$ be the mixing probability, then $g(x)$ is our desired probability distribution. Although the Mgamma distribution may be a particular form of one-parameter polynomial Exponential (OPPE) distribution introduced by Bouchahed and Zeghdoudi (2018) [16] taking $r = 2$ and $a_0 = 0, a_1 = 1, a_2 = 1/2$, it is easy to use by non-statisticians due to its flexible model. Whereas for OPPE, it is difficult to determine which values of a_i yield a good model, i.e., choice of a_i's are very difficult to estimate. Moreover, it is a general form, difficult for non-statisticians to use for modeling. Also, Bouchahed and Zeghdoudi (2018) did not test the bias and MSE, they only provided the formula.

3. The Mgamma distribution

Most of the one-parameter continuous distributions, such as Lindley, Shanker and Xgamma etc., are derived directly from simple transformations or mixtures of Exponential and Gamma with a single mixing scheme, whereas Mgamma is a finite mixture of two Gamma distributions with different shape parameters but a single scale parameter, which makes it different and richer than other classical distributions. A random variable X is said to have the Mgamma distribution with parameter θ if its pdf is defined as

$$f(x) = \frac{\theta^3 x \left(1 + \dfrac{x}{2}\right) e^{-\theta x}}{1+\theta}; \quad x > 0, \theta > 0 \tag{1}$$

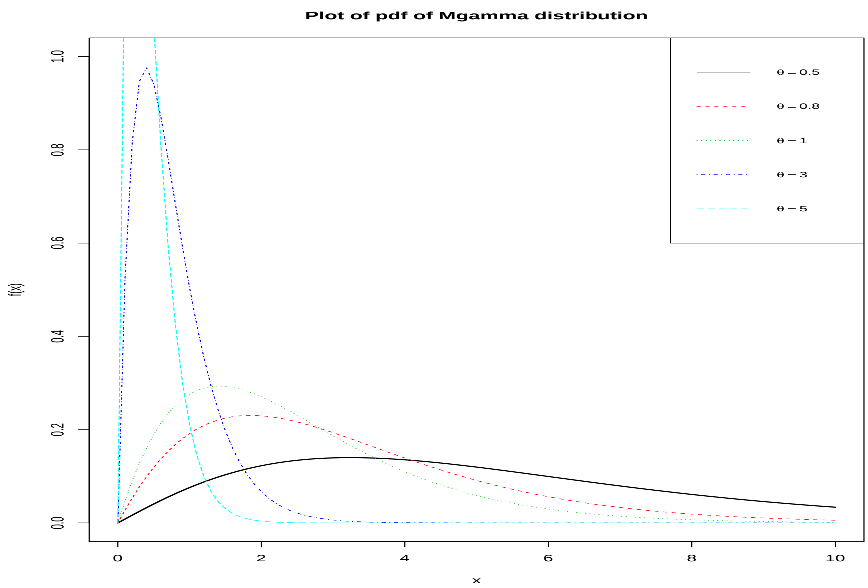

Figure 1: p.d.f. plot of Mgamma distribution for different θ

denoted by x $\sim$ M(θ), its cdf is given by

$$F(x) = 1 - \left[1 + \frac{\theta x}{1+\theta}\left(1 + \theta + \frac{\theta x}{2}\right)\right] e^{-\theta x}; \quad x > 0, \theta > 0 \tag{2}$$

To find the mode of the distribution, we first differentiate equation (1), w. r. t θ

$$\frac{d}{dx} f(x;\theta) = \frac{\theta^3}{1+\theta}\left[1 + x - \theta x(1 + x/2)\right] e^{-\theta x} \tag{3}$$

We observe that for $\theta < 1$, $\dfrac{(1-\theta) + \sqrt{\theta^2+1}}{\theta}$ is the unique critical point for which f(x) is maximized. And for $\theta > 1$, f(x) is decreasing.

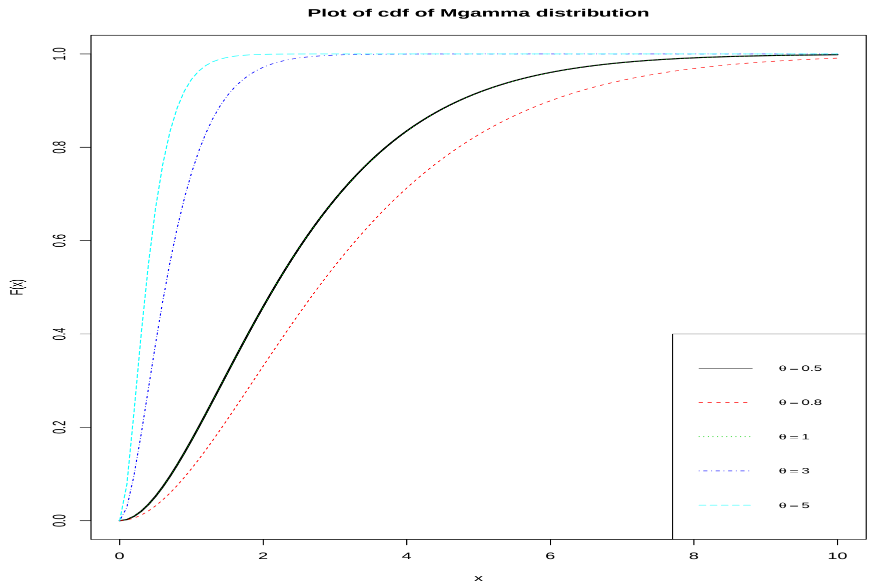

Figure 2: c.d.f. plot of Mgamma distribution for different θ

So, the mode(x) of the new distribution is given by

$$mode(x) = \frac{(1-\theta) + \sqrt{\theta^2+1}}{\theta}; \quad \theta < 1 \tag{4}$$

Remark 1. *From the above it is seen that* $mode(x) < median(x) < mean(x)$ *which is same as in the Xgamma and Exponential distribution. We get value of mode for different values of* θ

4. Moments and related measures

Let x ∼ Mgamma(θ), then the k^{th} raw moment is given by

$$\begin{aligned}
E(X^k) &= \int_0^\infty x^k \frac{\theta^3 x \left(1+\frac{x}{2}\right) e^{-\theta x}}{\theta+1}\, dx. \\
&= \frac{\theta^3}{\theta+1} \int_0^\infty x^{k+1} \left(1+\tfrac{x}{2}\right) e^{-\theta x}\, dx \\
&= \frac{\theta^3}{\theta+1} \left[\int_0^\infty x^{k+1} e^{-\theta x}\, dx + \tfrac{1}{2} \int_0^\infty x^{k+2} e^{-\theta x}\, dx \right] \\
&= \frac{\theta^3}{\theta+1} \left[\frac{\Gamma(k+2)}{\theta^{k+2}} + \tfrac{1}{2} \frac{\Gamma(k+3)}{\theta^{k+3}} \right] \\
&= \frac{\theta^3}{\theta+1} \left[\frac{(k+1)!}{\theta^{k+2}} + \tfrac{1}{2} \frac{(k+2)(k+1)!}{\theta^{k+3}} \right] \\
&= \frac{\theta^3 (k+1)!}{\theta^{k+3}(\theta+1)} \left[\theta + \tfrac{k+2}{2}\right]
\end{aligned}$$

$$= \frac{(k+1)!}{\theta^k(\theta+1)}\left(\theta+\frac{k+2}{2}\right). \tag{5}$$

Thus, the k^{th} order raw moment is given by

$$E(X^k)=\frac{(k+1)!}{\theta^k(\theta+1)}\left(\theta+\frac{k}{2}+1\right), \quad \theta>0,\ k>-2.$$

Therefore putting $k = 1, 2, 3$ and 4, we get the first, second, third and fourth order moments about the origin as:

$$\mu_1{}' = \frac{3+2\theta}{\theta^(1+\theta)} = \text{mean(x)},$$

$$\mu_2{}' = \frac{6(2+\theta)}{\theta^2(1+\theta)},$$

$$\mu_3{}' = \frac{12(5+2\theta)}{\theta^3(1+\theta)}$$

and

$$\mu_4{}' = \frac{120(3+\theta)}{\theta^4(1+\theta)}$$

.

The variance is given by

$$\begin{aligned}
\mu_2 = \text{Var}(X) &= \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2 \\
&= \frac{6(\theta+2)}{\theta^2(\theta+1)} - \left(\frac{2\theta+3}{\theta(\theta+1)}\right)^2 \\
&= \frac{6(\theta+2)}{\theta^2(\theta+1)} - \frac{(2\theta+3)^2}{\theta^2(\theta+1)^2} \\
&= \frac{6(\theta+2)(\theta+1)}{\theta^2(\theta+1)^2} - \frac{(2\theta+3)^2}{\theta^2(\theta+1)^2} \\
&= \frac{6(\theta+2)(\theta+1)-(2\theta+3)^2}{\theta^2(\theta+1)^2} \\
&= \frac{6(\theta^2+3\theta+2)-(4\theta^2+12\theta+9)}{\theta^2(\theta+1)^2} \\
&= \frac{6\theta^2+18\theta+12-4\theta^2-12\theta-9}{\theta^2(\theta+1)^2} \\
&= \frac{2\theta^2+6\theta+3}{\theta^2(\theta+1)^2}.
\end{aligned}$$

$$\begin{aligned}
\mu_3 &= \mathbb{E}(X^3) - 3\mathbb{E}(X)\mathbb{E}(X^2) + 2(\mathbb{E}(X))^3 \\
&= \frac{12(2\theta+5)}{\theta^3(\theta+1)} - 3\cdot\frac{2\theta+3}{\theta(\theta+1)}\cdot\frac{6(\theta+2)}{\theta^2(\theta+1)} + 2\cdot\frac{(2\theta+3)^3}{\theta^3(\theta+1)^3} \\
&= \frac{12(2\theta+5)(\theta+1)^2 - 18(2\theta+3)(\theta+2)(\theta+1) + 2(2\theta+3)^3}{\theta^3(\theta+1)^3} \\
&= \frac{2\big(2\theta^3+9\theta^2+9\theta+3\big)}{\theta^3(\theta+1)^3}
\end{aligned}$$

$$\begin{aligned}
\mu_4 &= \mathbb{E}(X^4) - 4\mathbb{E}(X^3)\mathbb{E}(X) + 6(\mathbb{E}(X^2)\mathbb{E}(X)) - 3\mathbb{E}(X)^4 \\
&= \frac{24\theta^4+144\theta^3+252\theta^2+180\theta+45}{\theta^4(1+\theta)^4}
\end{aligned}$$

Again, putting the value of (μ_2) and (μ_4) we get kurtosis (γ_2) as,

$$\begin{aligned}
\gamma_2 = \frac{\mu_4}{\mu_2^2} &= \frac{\dfrac{24\theta^4+144\theta^3+252\theta^2+180\theta+45}{\theta^4(\theta+1)^4}}{\dfrac{(2\theta^2+6\theta+3)^2}{\theta^4(\theta+1)^4}} \\
&= \frac{3(8\theta^4+48\theta^3+84\theta^2+60\theta+15)}{4\theta^4+24\theta^3+48\theta^2+36\theta+9} \\
&= \frac{24\theta^4+144\theta^3+252\theta^2+180\theta+45}{(2\theta^2+6\theta+3)^2}
\end{aligned}$$

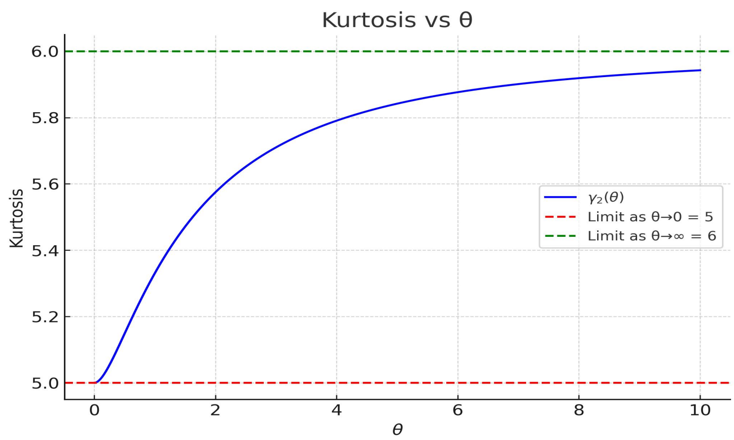

Figure 3: The graph of kurtosis for different θ

Now, putting the value of (μ_2) and (μ_3) we get the value of skewness (γ_1) as ,

$$\begin{aligned}\gamma_1 &= \frac{\mu_3}{\mu_2^{3/2}} \\ &= \frac{\frac{2\left(2\theta^3+9\theta^2+9\theta+3\right)}{\theta^3(\theta+1)^3}}{\frac{(2\theta^2+6\theta+3)^{3/2}}{\theta^3(\theta+1)^3}} \\ &= \frac{2\left(2\theta^3+9\theta^2+9\theta+3\right)}{(2\theta^2+6\theta+3)^{3/2}}.\end{aligned}$$

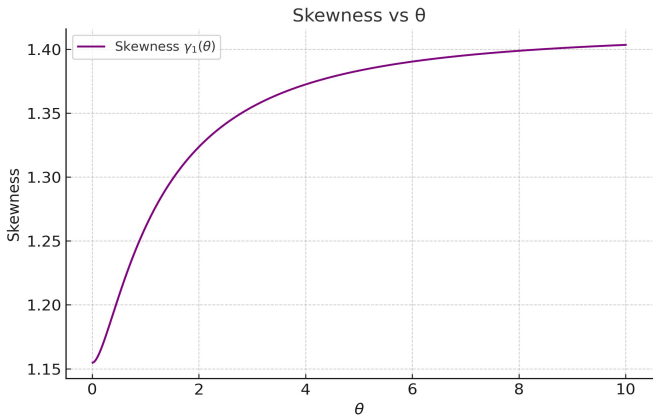

Figure 4: The graph of skewness for different θ

From the Fig. 3 and Fig. 4 we can conclude the following observations,

- As the numerator and denominator are positive for all $\theta > 0$ in (γ_1), so $\gamma_1 > 0$ i.e. the distribution is right-skewed.
- Again, for kurtosis, we found that ($\gamma_2 > 3$), $\forall\theta > 0$, so the distribution is leptokurtic (heavy-tailed, more peaked than normal).

4.1. Mean Deviation

The mean deviation about mean of random variable X, having pdf in equation (1) is obtained as

$$\begin{aligned} MD_{\bar{x}} &= E(|x-\mu|) \\ &= \int_0^\infty |x-\mu| f(x)dx \\ &= \int_0^\mu |\mu - x| f(x)dx + \int_\mu^\infty |x-\mu| f(x)dx \\ &= 2\mu F(\mu) - \mu + 2\int_\mu^\infty x f(x)dx \\ &= 2\mu F(\mu) - \mu + \frac{1}{\theta(1+\theta)}\left(2\theta(\theta^2\mu^2 + 2\theta\mu + 2) + (\theta^3\mu^3 + 3\theta^2\mu^2 + 6\theta\mu + 6)\right)e^{-\theta\mu} \end{aligned}$$

4.2. Moment and Cumulant Generating Function

It is also possible to express the moment generating function in terms of the moment

$$\begin{aligned} M_X(t) = E(e^{tX}) &= \int_0^\infty e^{tx}\left(\frac{\theta^3 x(1+\frac{x}{2})e^{-\theta x}}{1+\theta}\right)dx \\ &= \frac{\theta^3}{1+\theta}\int_0^\infty x\left(1+\frac{x}{2}\right)e^{-(\theta-t)x}dx \\ &= \frac{\theta^3}{1+\theta}\left(\int_0^\infty x e^{-(\theta-t)x}dx + \int_0^\infty \frac{x}{2}e^{-(\theta-t)x}\right)dx \\ &= \frac{\theta^3}{1+\theta}\left(\frac{\Gamma 2}{(\theta-t)^2} + \frac{\Gamma 2}{2(\theta-t)^2}\right) \\ &= \frac{\theta^3(1+\theta-t)}{(1+\theta)(\theta-t)^2} \end{aligned}$$

The cumulant generating function of the random variable X is given by

$$k_X(t) = \log M_X(t)$$

$$\begin{aligned} K_X(t) &= \log\left(\frac{\theta^3(1+\theta-t)}{(1+\theta)(\theta-t)^2}\right) \\ &= 3\log\theta + \log(1+\theta-t) - \log(1+\theta) - 2\log(\theta-t) \end{aligned}$$

5. Stochastic interpretation of the parameter θ

Stochastic orders are important measures for comparing the behavior of random variables shown by Shaked and Santikumar (1995) [17]) Let X and Y are the two random variables with cumulative distribution functions $F_X(.)$ and $F_Y(.)$ respectively. Then X is said to be smaller than Y in the

- Stochastic order $(X \prec_s Y), if F_X(t) \geq F_Y(t), \forall$t
- Convex order $(X \prec_{cx} Y)$ if for all convex functions ϕ and provided expectation exist, $E[\phi(X) \leq E[\phi(Y)]$.
- Hazard rate order $(X \prec_{hr} Y), if h_X(t) \geq h_Y(t), \forall$t.
- Likelihood Ratio order $(X \prec_{lr} Y), if \frac{f_X(t)}{f_Y(t)}$ is decreasing in t.

Remark 2. *Likelihood ratio order* $\Rightarrow$ *Hazard rate order* $\Rightarrow$ *Stochastic order if E[X] = E[Y] , then Convex order* $\iff$ *Stochastic order.*

Theorem 1. *Let* $X_i \sim M(\theta_i), i = 1, 2$ *be two random variables. If* $\theta_1 \leq \theta_2, then X_1 \prec_{lr} X_2, X_1 \prec_{hr} X_2, X_1 \prec_s X_2 and X_1 \prec_{cx} X_2$

Proof. We have

$$\frac{f_{X_1}(t)}{f_{X_2}(t)} = \frac{\theta_1^3(1+\theta_2)}{\theta_2^3(1+\theta_1)} e^{-(\theta_1-\theta_2)t}$$

differentiating w.r.t t, we get

$$\frac{d}{dt}\frac{f_{X_1}(t)}{f_{X_2}(t)} = -(\theta_1-\theta_2)e^{-(\theta_1-\theta_2)t} \quad < 0, \quad \forall \theta_1 \geq \theta_2$$

Hence, we see that $\frac{f_{X_1}(t)}{f_{X_2}(t)}$ decreases in t, which holds the above theorem. (proved)

6. Incomplete moments and Inequality curves

The r^{th} incomplete moment $\mu_r(t)$ for a random variable X with pdf f(x) and cdf F(x) is given by Sen et al. (2018) [18])

$$\mu_r(t) = \int_0^t x^r f(x) dx$$

So, when $x \sim Mgamma(\theta)$, the r^{th} incomplete moment is obtained as

$$\begin{aligned}\mu_r(t) &= \int_0^t x^r \left(\frac{\theta^3 x(1+\frac{x}{2})e^{-\theta x}}{1+\theta}\right) dx \\ &= \frac{\theta^3}{1+\theta}\bigg(\int_0^\infty x^{r+1}e^{-\theta x}dx + \int_0^\infty \frac{1}{2}x^{r+2}e^{-\theta x}dx\bigg) \\ &= \frac{\theta^3}{1+\theta}\left[\gamma(r+2,\theta t) + \frac{1}{2}\gamma(r+3,\theta t)\right]\end{aligned}$$

where $\gamma(\alpha, x) = \int_0^x u^{\alpha-1}e^{-u}\, du$ is lower incomplete Gamma function. When a non-negative continuous random variable X has pdf f (x) and cdf F (x), the Lorenz and Bonferroni curves are defined by

$$l(p) = \frac{1}{\mu}\int_0^q xf(x)dx$$

$$B(p) = \frac{1}{p\mu}\int_0^q xf(x)dx$$

respectively, where $\mu = E(X)$ and $q = F^1(p)$ for $0 < p < 1$. So, when x ∼ Mgamma(θ) the Lorenz and Bonferroni curves are given by

$$l(p) = \frac{\theta^4}{3+\theta}\left[\gamma(3,\theta q) + \frac{1}{2}\gamma(4,\theta q)\right]$$

$$B(p) = \frac{\theta^4}{p(3+\theta)}\left[\gamma(3,\theta q) + \frac{1}{2}\gamma(4,\theta q)\right]$$

The Lorenz and Bonferroni curves confirm that Mgamma can capture income/wealth inequality patterns due to its flexible tail behavior, offering better interpretability compared to simpler one-parameter models

7. Reliability Properties

In reliability engineering, a batch of products may contain both weaker items (e.g., Gamma$(2, \theta)$) and stronger items (long-lasting Gamma$(3, \theta)$). Similarly, in the medical field, the patients' survival times may be a mixture of fast progressors and slow progressors, etc. A mixture allows non-monotone hazard rates, which are often observed in practice but impossible with simple one-parameter models. So Mgamma distribution is very much suitable in this field with single- parameter and mixture distributions.

The hazard rate of Mgamma distribution is

$$h(x) = \frac{f(x)}{1-F(x)}$$

$$= \frac{\theta^3 x(1+\frac{x}{2})e^{-\theta x}}{\left[1+\frac{\theta x}{1+\theta}\left(1+\theta+\frac{\theta x}{2}\right)\right]e^{-\theta x}}$$

$$= \frac{\theta^3 x\left(1+\frac{x}{2}\right)}{\left[1+\theta+\theta x\left(1+\theta+\frac{\theta x}{2}\right)\right]}$$

From Fig. 5, it is evident that unlike the Exponential and Lindley distributions, the hazard rate of the Mgamma distribution is non-monotone.

(i) For small θ, $h(x)$ is increasing.

(ii) For moderate θ, $h(x)$ shows a bathtub shape.

(iii) For large θ, $h(x)$ is monotone increasing like Lindley.

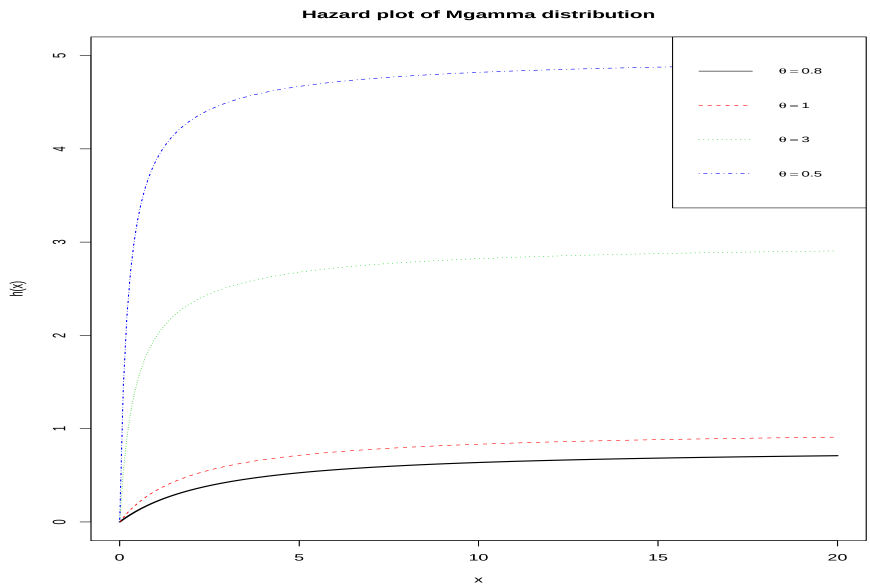

Figure 5: Hazard plot of Mgamma distribution for different values of θ

Mean remaining life

For a continuous random variable X with pdf $f(x)$ and cdf $F(x)$, the mean residual life (mrl) function is defined as

$$m(x) = E[(X-x)|X>x]$$

$$= \frac{1}{1-F(x)}\int_x^{\infty}[1-F(t)]dt$$
$$= \frac{\int_x^{\infty}\bar{F}(t)dt}{\bar{F}(x)}$$

Where,

$$\bar{F}(t) = \left(1 + \frac{\theta t}{1+\theta}(1+\theta+\frac{\theta t}{2}\right) e^{-\theta t}$$

and,

$$\int_x^{\infty}\bar{F}(t)dt = \frac{1}{1+\theta}\left((1+\theta)x + \frac{2(1+\theta)}{\theta} + x + \frac{1}{\theta} + \frac{\theta x^2}{2}\right) \tag{6}$$

For the Mgamma distribution the mrl function is given by

$$m(x) = \frac{(1+\theta)x + \dfrac{2(1+\theta)}{\theta} + x + \dfrac{1}{\theta} + \dfrac{\theta x^2}{2}}{1+\theta+\theta x(1+\theta+\theta x/2)}$$

Figure 6 shows mrl functions for different choices of θ.

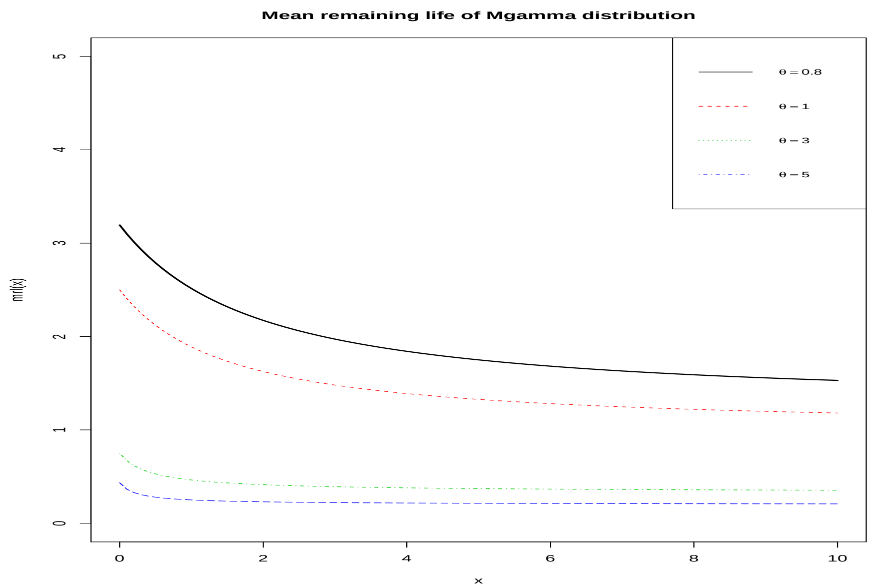

Figure 6: mrl plot of Mgamma distribution for different values of θ

From the plot we found that

(i) For $x = 0$, $m(0) = \mu = \dfrac{2\theta+3}{\theta(1+\theta)}$ which is the mean of Mgamma.

(ii) The mrl(x) decreases with both x and θ, indicating that the expected remaining life becomes shorter.

8. Order Statistics

Let $Y_{(1)}, Y_{(2)}, .., Y_{(n)}$ be the ordered random sample drawn from the Mgamma Distribution, then the c.d.f. and p.d.f. of the r^{th} order statistics is given by

$$G_r(y) = \sum_{i=r}^{n} \binom{n}{i} [G(y;\theta;\lambda)]^i \, [1 - G(y;\theta;\lambda)]^{n-i}$$

$$g_r(y) = \frac{n!}{(r-1)!(n-r)!} [G(y;\theta;\lambda)]^{r-1} \, [1 - G(y;\theta;\lambda)]^{n-r} \, g(y;\theta;\lambda)$$

So, the p.d.f. and c.d.f. of $y_{(1)}$ and $y_{(n)}$ is given by

$$g_{Y_1}(y) = \frac{n\theta^3 x(1+x/2)e^{-\theta x}}{1+\theta} \left[\left(1 + \frac{\theta x}{1+\theta}(1+\theta+\frac{\theta x}{2})\right) e^{-\theta x}\right]^{n-1}$$

$$g_{Y_n}(y) = \frac{n\theta^3 x(1+x/2)e^{-\theta x}}{1+\theta} \left[1 - \left(1 + \frac{\theta x}{1+\theta}(1+\theta+\frac{\theta x}{2})\right) e^{-\theta x}\right]^{n-1}$$

$$G_{y_1}(y) = 1 - \left[\left(1 + \frac{\theta x}{1+\theta}(1+\theta+\frac{\theta x}{2})\right) e^{-\theta x}\right]^{n}$$

$$G_{y_n}(y) = \left[1 - \left(1 + \frac{\theta x}{1+\theta}(1+\theta+\frac{\theta x}{2})\right) e^{-\theta x}\right]^{n}$$

9. Estimation of the parameters

Here we discuss two estimation methods *viz.* (i) Method of moments and (ii) Methods of maximum likelihood estimation.

9.1. Method of moment

Let $x_1, x_2, x_3, .., x_n$ be a random sample of size n drawn from the equation (1). We compare the sample mean $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ with the first order raw moment $\mu_1{}'$ to get the mom estimator of θ as

$$\theta_{mom} = \frac{1}{2\bar{x}} \left(-\bar{x} + \sqrt{\bar{x}^2 + 8\bar{x} + 4} + 2\right), \quad x > 0$$

The method of moment estimator θ_{mom} is positively biased, consistent and asymptotically normal.

For large sample the 100(1-α)% confidence interval of θ is given by $\theta \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{1}{n\sigma^2}}$

9.2. Maximum likelihood estimation

Let $X_1, X_2, ..., X_n$ be a random sample of size n drawn independently from the pdf 1. The maximum likelihood estimator (mle) of θ is given by

$$L(\theta|x_1, x_2, .., x_n) = \prod_{i=1}^{n} \frac{\theta^3 x_i(1 + x_i/2)e^{-\theta x_i}}{1 + \theta}$$

The log likelihood equation is as follows

$$L(\theta|x) = 3nlog(\theta) - nlog(1 + \theta) + \sum_{i}^{n} log\left[x_i(1 + x_i/2)\right] - \theta \sum_{i}^{n} x_i \tag{7}$$

To have the maximum likelihood estimator of θ, we have to solve $I(\theta|x) = \frac{\partial L(\theta|x)}{\partial \theta} = 0$, Hence, we have

$$\frac{3n}{\theta} - \frac{n}{1 + \theta} - \sum_{i=1}^{n} x_i = 0$$

Using Newton-Raphson iteration method, we find the $\hat{\theta}_{mle}$ for which the log-likelihood equation is maximized. The initial solution for the iteration method is $\theta_0 = \frac{2}{\bar{x}}$. We continue the iteration process

$$\theta^{(i)} = \theta^{(i-1)} - \frac{I(\theta^{(i-1)}|x)}{I'(\theta^{(i-1)}|x)}$$

we stop when $\theta^{(i)} \equiv \theta^{(i-1)}$. All computations were performed in the updated version of R software [19], following the procedures book [20]. The MLE is consistent and asymptotically normal under standard regularity conditions (iid samples, differentiability of likelihood). Convergence of Newton–Raphson is generally fast; in our simulation, it stabilized within 4–6 iterations.

10. Simulation Study

To generate the random sample from Mgamma distribution, we proceed as follows: Let $F(x) = p$, where p follows $uniform(0, 1)$. We choose the solution that lies in $(0, 1)$. Here $G(x)$ is the cdf of Mgamma random variable. Then, we follow the approach of generating sample from Mgamma distribution. Hence the algorithm is

1. Generate $p_i \sim uniform(0, 1)$, $i = 1, 2, \ldots, n$.
2. Generate $V_i \sim Gamma(2, \theta)$, $i = 1, 2, \ldots, n$.
3. Generate $W_i \sim Gamma(3, \theta)$, $i = 1, 2, \ldots, n$.
4. If $U_i \leq \frac{\theta}{1+\theta}$, then set $X_i = V_i$. Otherwise, set $X_i = W_i$.

We perform Monte Carlo simulation in R programming considering N=10000 times for different values of n=5, 10, 15 and 20 and θ=0.8, 1, 2 and 3 . The following two measures were computed

1. Average bias of the simulated estimates $\hat{\theta}_i$, $i = 1, 2, \ldots, N$: $\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta)$
2. Average Mean Square Error (MSE) of the simulated estimates $\hat{\theta}_i$, $i = 1, 2, \ldots, N$: $\frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta)^2$

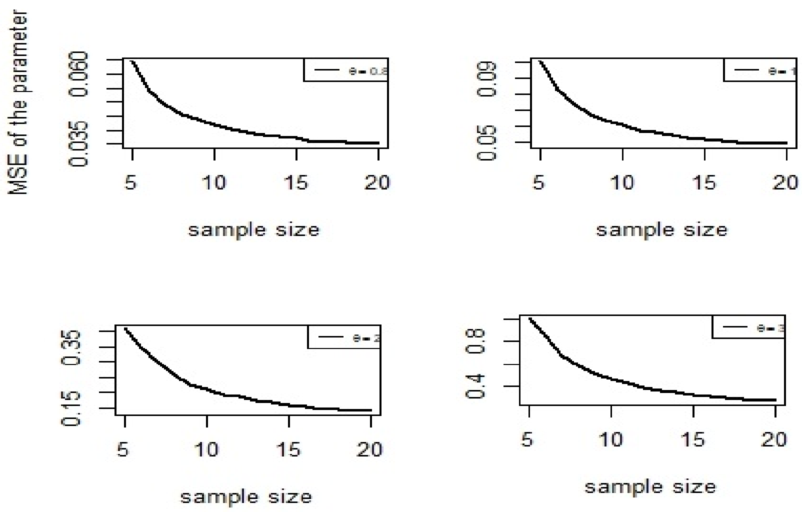

Figure 7: mse plot of Mgamma distribution for different value of θ

It is observed that both the MSE and bias decrease as the sample size increases and θ decreases.

11. Data Analysis

In this section, we have compared our proposed model with the existing one-parameter distributions by fitting some real-life datasets. The first dataset is obtained from Smith and Naylor (1987) [21], which represents the strength of 1.5 cm glass fibres, measured at the National Physical laboratory, England.

The second dataset is taken from Bakouch et. al (2021)[22] which concerns the exceedances of flood peaks (in m^3/s) of the Wheaton River near Carcross in Yukon Territory, Canada that consist of 72 exceedances for the years 1958–1984 and rounded to one decimal place.

The third dataset is about the remission times (in weeks) for 20 leukaemia patients randomly assigned to a certain treatment taken from Al-Babtain et al. (2020) [23].

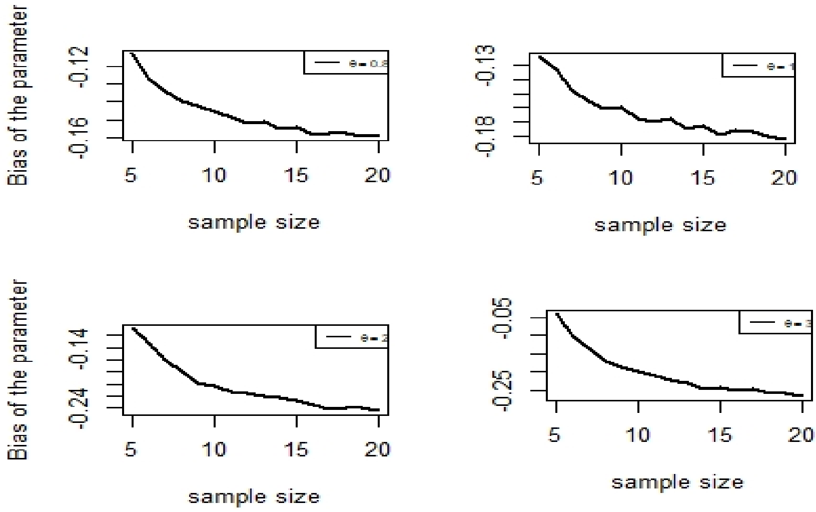

Figure 8: bias plot of Mgamma distribution for different value of θ

The last dataset is obtained from Sharma et al. (2017) [24], which consists of the waiting time (in mins) before service for 100 bank customers.

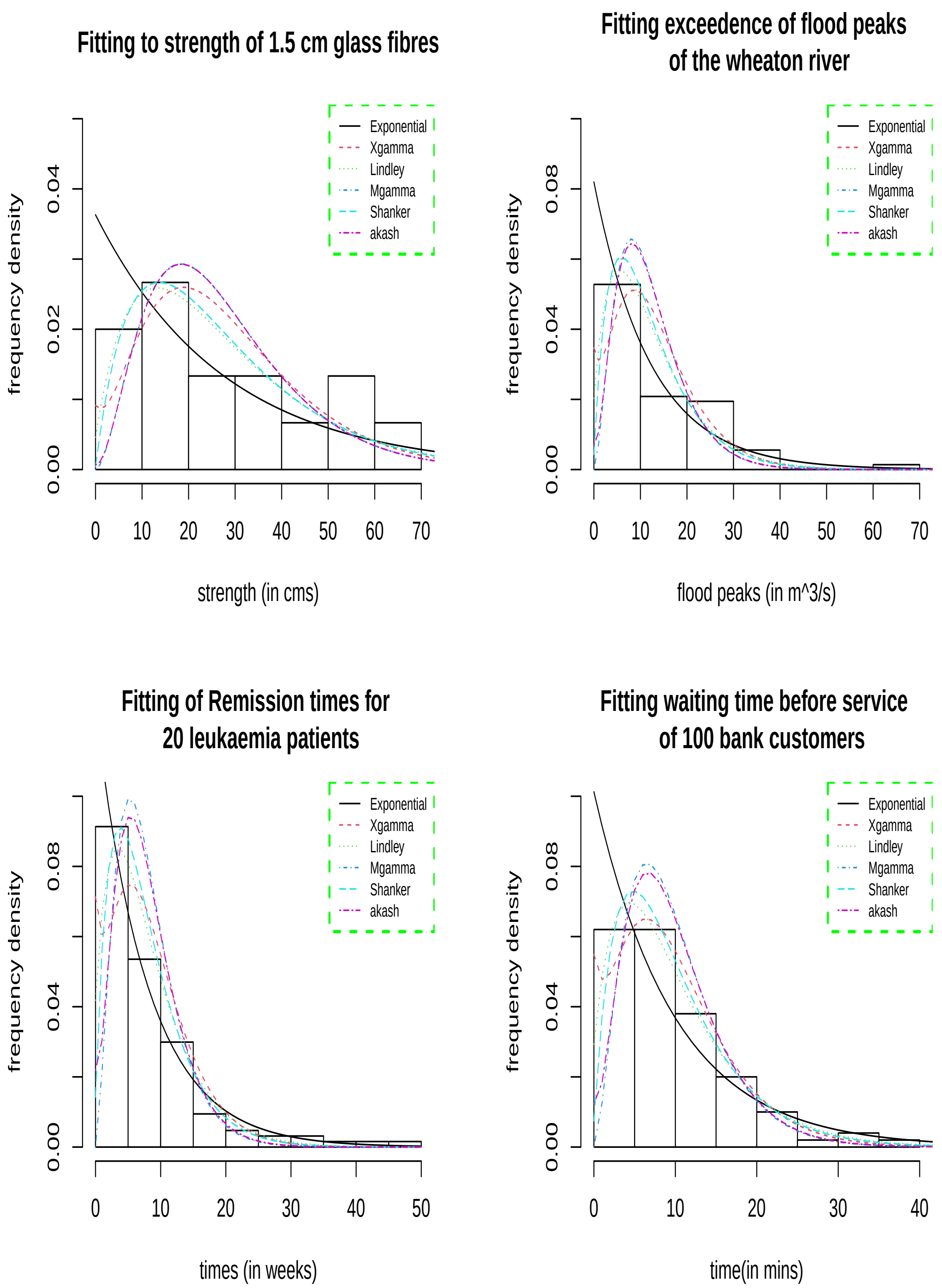

Figure 9: Histogram and Fitted densities for all the four datasets.

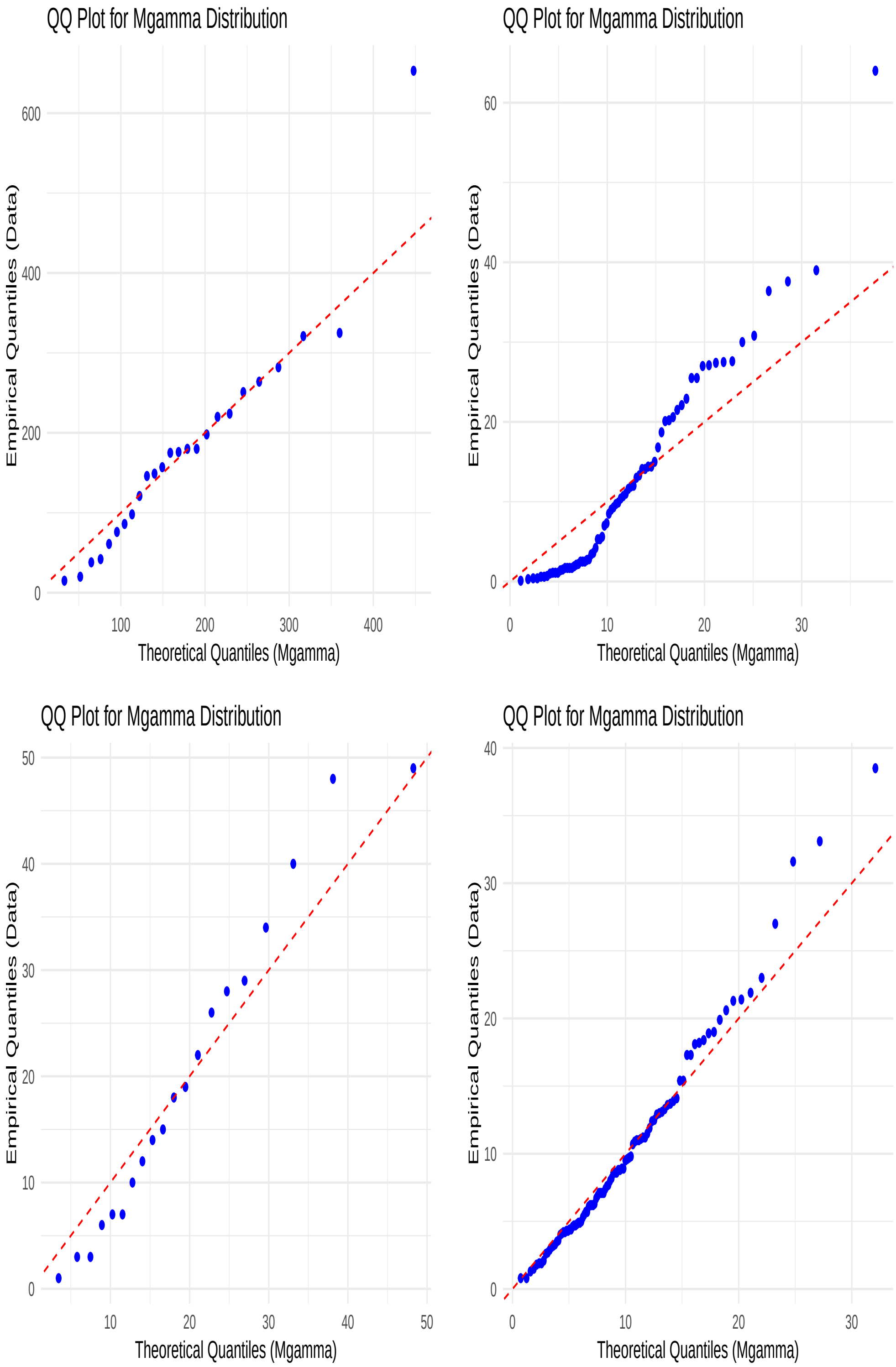

Figure 10: Q-Q plot for all the four datasets using Mgamma distribution accordingly.

Table 1: Model comparison for Dataset I: Strength of 1.5 cm glass fibres

Criteria	Exponential	Lindley	Xgamma	Shanker	Akash	Mgamma
MLE	0.6636	0.9961	1.3376	0.9562	1.3554	1.8354
AIC	88.83	81.27	85.80	81.14	81.86	**58.66**
KS Statistic	0.4186	0.3862	0.4196	0.3661	0.3705	**0.3361**
Cramér–von Mises	3.862	3.332	3.881	3.113	3.192	**2.496**
Anderson–Darling	18.426	16.425	18.468	15.406	15.736	**12.590**

Table 2: Model comparison for Dataset II: Exceedances of flood peaks (Wheaton River)

Criteria	Exponential	Lindley	Xgamma	Shanker	Akash	Mgamma
MLE	0.8194	0.1530	0.2044	0.1646	0.2412	0.2422
AIC	252.120	264.228	266.470	279.790	291.070	**251.648**
KS Statistic	0.1421	0.2408	0.2562	0.2272	0.3030	**0.1325**
Cramér–von Mises	0.2319	0.8198	1.0660	0.7775	1.5190	**0.1489**
Anderson–Darling	1.4680	7.4400	8.8136	6.9887	19.100	**1.4231**

Table 3: Model comparison for Dataset III: Remission times of 20 leukaemia patients

Criteria	Exponential	Lindley	Xgamma	Shanker	Akash	Mgamma
MLE	0.0511	0.0977	0.1383	0.1019	0.1522	0.1521
AIC	79.459	78.872	79.398	79.244	81.790	**78.746**
KS Statistic	0.1143	0.1189	0.1589	0.1095	0.2007	**0.0104**
Cramér–von Mises	0.0660	0.0344	0.0725	0.0424	0.1366	**0.0181**
Anderson–Darling	0.410	0.2960	0.5644	0.3351	1.4800	**0.2704**

Table 4: Model comparison for Dataset IV: Waiting times of 100 bank customers

Criteria	Exponential	Lindley	Xgamma	Shanker	Akash	Mgamma
MLE	0.1012	0.1865	0.2634	0.1983	0.2952	0.2989
AIC	329.020	319.037	321.020	317.629	320.629	**317.062**
KS Statistic	0.1730	0.0676	0.0624	0.0881	0.1002	**0.0601**
Cramér–von Mises	0.7153	0.0581	0.0839	0.1280	0.2170	**0.0535**
Anderson–Darling	4.2280	0.4863	0.6447	0.9044	1.3797	**0.0326**

From the table 1 - 4, we find that Mgamma provides the best fit, capturing the skewness and heavy tails of the data. Across all four datasets i.e. spanning engineering, hydrology, medical survival, and queueing contexts the Mgamma distribution consistently outperformed the classical one-parameter models. The mixture structure of Mgamma allows it to model skewness and heavy-tailed behaviour more effectively. Also the model accommodates non-monotone hazard rates (e.g., bathtub-shaped or increasing-decreasing), which are common in practice but not captured by simpler distributions. Even with small samples (as in the remission time data), Mgamma provides superior fit. In every dataset, Mgamma yielded the lowest AIC and other criteria, reinforcing its statistical adequacy. Thus, the Mgamma distribution can be considered a highly competitive and flexible alternative for real-life data modelling, outperforming several classical and recently proposed lifetime distributions. It is well known that distributions with more parameters usually provide greater flexibility for data modeling. However, the presence of multiple parameters often makes estimation more complex and computationally demanding. In contrast, the proposed Mgamma distribution has the advantage of involving only a single parameter, which is straightforward to estimate, while still offering substantial flexibility. Notably, it provides a superior fit compared to several existing multi-parameter models.

12. Conclusion

In this article, we introduced a new one-parameter continuous distribution, the so called Mgamma distribution, by a finite mixture approach of two Gamma distributions with different shape parameters. We studied several of its statistical properties including moments, moment and cumulant generating function, order statistics, mean deviation, incomplete moments and inequality curves. We generated random variables and studied their mean square error and bias and found that both decrease as the sample size increases. The estimation of model parameters is approached by the method of maximum likelihood and method of moments estimation. Two real-life data sets are analyzed using various model selection criteria. It is found that the new distribution provides a consistently better fit than the other probability models available in the literature. For future studies, this new distribution can be modified using a different technique to make it more flexible in data modelling, such as developing its pseudo form, considering the ratio of two random variables, introducing transmuted versions with an additional parameter, or exploring Bayesian estimation approaches.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of interest

There is no conflict of interest.

Acknowledgements

The authors are grateful to the Deanship of Graduate Studies and Scientific Research, Islamic University of Madinah, Saudi Arabia, for the support provided to the Post-Publication Program (4). The authors also thank the unknown referees and editors Eyup Cetin and Baris Kiremitci for their valuable comments and suggestions to improve the quality of the article.

References

[1] Geoffrey J McLachlan and Kaye E Basford. *Mixture models: Inference and applications to clustering*, volume 38. M. Dekker New York, 1988.

[2] Mohamed E Ghitany, Barbra Atieh, and Saralees Nadarajah. Lindley distribution and its application. *Mathematics and computers in simulation*, 78(4):493–506, 2008.

[3] Sudhansu S Maiti, Molay Kumar Ruidas, and Sumanta Adhya. Natural discrete one parameter polynomial exponential family of distributions and the application. *Annals of Data Science*, 11(3):1051–1076, 2024.

[4] Subhradev Sen, Sudhansu S Maiti, and N Chandra. The xgamma distribution: statistical properties and application. *Journal of Modern Applied Statistical Methods*, 15(1):38, 2016.

[5] Dennis V Lindley. Fiducial distributions and bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 102–107, 1958.

[6] R Shanker. Shanker distribution and its applications. *International journal of statistics and Applications*, 5(6):338–348, 2015.

[7] Halim Zeghdoudi and Sihem Nedjar. Gamma lindley distribution and its application. *Journal of Applied Probability and Statistics*, 11(1):129–138, 2016.

[8] Norman Lloyd Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. Continuous univariate distributions. 1994.

[9] Hassan S Bakouch, Bander M Al-Zahrani, Ali A Al-Shomrani, Vitor AA Marchi, and Francisco Louzada. An extended lindley distribution. *Journal of the Korean Statistical Society*, 41(1):75–85, 2012.

[10] Haitham M Yousof, Mustafa Ç Korkmaz, and Subhradev Sen. A new two-parameter lifetime model. *Annals of Data Science*, 8:91–106, 2021.

[11] AAM Mahmoud, Rana H Khashab, Zakiah I Kalantan, Sulafah M Saleh Binhimd, Amani S Alghamdi, and Said G Nassr. On bivariate compound exponentiated survival function of the beta distribution: Estimation and prediction. *Journal of Radiation Research and Applied Sciences*, 17(2):100886, 2024.

[12] Xiangming Tang, Jin-Taek Seong, Randa Alharbi, Aned Al Mutairi, and Said G Nasr. A new probabilistic model: Theory, simulation and applications to sports and failure times data. *Heliyon*, 10(4), 2024.

[13] Mustafa Kamal, Meshayil M Alsolmi, Aned Al Mutairi, Eslam Hussam, Manahil SidAhmed Mustafa, Said G Nassr, et al. A new distributional approach: estima-

tion, monte carlo simulation and applications to the biomedical data sets. *Networks & Heterogeneous Media*, 18(4), 2023.

[14] Mahendra Saha, Abhimanyu Singh Yadav, Arvind Pandey, Shivanshi Shukla, and Sudhansu S Maiti. The extended xgamma distribution. *arXiv preprint arXiv:1909.01103*, 2019.

[15] Hamouda Messaadia and Halim Zeghdoudi. Zeghdoudi distribution and its applications. *International Journal of Computing Science and Mathematics*, 9(1):58–65, 2018.

[16] Lahsen Bouchahed and Halim Zeghdoudi. A new and unified approach in generalizing the lindley's distribution with applications. *Statistics*, 61, 2018.

[17] Moshe Shaked, J George Shanthikumar, and YL Tong. Stochastic orders and their applications. *SIAM Review*, 37(3):477–478, 1995.

[18] Subhradev Sen, N Chandra, and Sudhansu S Maiti. On properties and applications of a two-parameter xgamma distribution. *Journal of Statistical Theory and Applications*, 17(4):674–685, 2018.

[19] R Core Team. *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.* Vienna, Austria, 2025.

[20] Ross Ihaka and Robert Gentleman. R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, 5(3):299–314, 1996.

[21] Richard L Smith and J. C. Naylor. A comparison of maximum likelihood and bayesian estimators for the three-parameter weibull distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 36(3):358–369, 1987.

[22] Hassan S Bakouch, Abdus Saboor, and Muhammad Nauman Khan. Modified beta linear exponential distribution with hydrologic applications. *Annals of Data Science*, 8(1):131–157, 2021.

[23] Abdulhakim A Al-Babtain, Abdul Hadi N Ahmed, and Ahmed Z Afify. A new discrete analog of the continuous lindley distribution, with reliability applications. *Entropy*, 22(6):603, 2020.

[24] Vikas Kumar Sharm, Sanjay Kumar Singh, and Umesh Singh. Classical and bayesian methods of estimation for power lindley distribution with application to waiting time data. *Communications for statistical applications and methods*, 24(3):193–209, 2017.