STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–21.
Published online in International Academic Press (www.IAPress.org)

Classical and Bayesian estimation of the Mgamma distribution parameter with an application to orthodontic tooth movement data

Hossein Pasha-Zanoosi*

Department of Basic Sciences and General Courses, Khorramshahr University of Marine Science and Technology, Iran

Abstract In this paper, a novel one-parameter continuous distribution known as the Mgamma distribution is investigated for modeling lifetime data. This distribution is constructed as a finite mixture of two Gamma distributions with shape parameters 2 and 3, which enables flexibility in capturing skewed data and non-monotone hazard rate patterns. Several estimation approaches for the parameter θ are examined, including the maximum likelihood estimator (MLE), the method of moments (MoM) estimator, the uniformly minimum variance unbiased estimator (UMVUE), and Bayesian procedures based on Lindley's approximation and Markov Chain Monte Carlo (MCMC) techniques. A comprehensive simulation study is carried out to assess the performance of the proposed estimators in terms of bias and mean squared error. For interval estimation, the asymptotic confidence interval (ACI) derived from the MLE and the highest posterior density (HPD) intervals obtained from MCMC samples are constructed and compared. Moreover, the practical applicability of the Mgamma distribution is demonstrated through an analysis of real orthodontic tooth movement data.

Keywords Mgamma distribution, Orthodontic tooth movement data, UMVUE, Lindley's approximation

AMS 2010 subject classifications 62F10, 62F12, 62N05

DOI: 10.19139/soic-2310-5070-4611

1. Introduction

The Lindley distribution was introduced by Lindley [9] in the context of Bayesian statistics. It is a one-parameter continuous distribution that can be expressed as a mixture of an exponential distribution and a Gamma distribution. Specifically, it is a two-component mixture of exponential(θ) and Gamma$(2,\theta)$ distributions with mixing proportions $\theta/(\theta+1)$ and $1/(\theta+1)$, respectively. This mixture representation provides greater flexibility than the standard exponential distribution, making it suitable for modeling lifetime data with monotone increasing hazard rates. Since its introduction, numerous distributions have been developed using a similar mixture approach to achieve greater flexibility in modeling various types of data, particularly in reliability and survival analysis.

Shanker [15] introduced the Akash distribution, which is a two-component mixture of exponential(θ) and Gamma$(3,\theta)$ distributions with mixing proportions $\theta^2/(\theta^2+2)$ and $2/(\theta^2+2)$, respectively. In the same year, Shanker [16] proposed the Shanker distribution, another two-component mixture of exponential(θ) and Gamma$(2,\theta)$ distributions with mixing proportions $\theta^2/(\theta^2+1)$ and $1/(\theta^2+1)$, respectively. Later, Shanker [17] introduced the Aradhana distribution, which extends the mixture approach to three components, combining exponential(θ), Gamma$(2,\theta)$, and Gamma$(3,\theta)$ distributions with mixing proportions $\theta^2/(\theta^2+2\theta+2)$, $2\theta/(\theta^2+2\theta+2)$, and $2/(\theta^2+2\theta+2)$, respectively. Shanker [18] also proposed the Sujatha distribution, another three-component mixture of exponential(θ), Gamma$(2,\theta)$, and Gamma$(3,\theta)$ distributions with mixing proportions $\theta^2/(\theta^2+\theta+2)$, $\theta/(\theta^2+\theta+2)$, and $2/(\theta^2+\theta+2)$, respectively. In the same year, Shanker [19] introduced

*Correspondence to: Hossein Pasha-Zanoosi (Email: pashazanoosi@yahoo.com). Department of Basic Sciences and General Courses, Faculty of Economics and Management, Khorramshahr University of Marine Science and Technology, Khorramshahr, Iran.

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 202x International Academic Press

the Devya distribution, which is a five-component mixture of exponential(θ), Gamma($2,\theta$), Gamma($3,\theta$), Gamma($4,\theta$), and Gamma($5,\theta$) distributions with mixing proportions $\theta^4/(\theta^4+\theta^3+2\theta^2+6\theta+24)$, $\theta^3/(\theta^4+\theta^3+2\theta^2+6\theta+24)$, $2\theta^2/(\theta^4+\theta^3+2\theta^2+6\theta+24)$, $6\theta/(\theta^4+\theta^3+2\theta^2+6\theta+24)$, and $24/(\theta^4+\theta^3+2\theta^2+6\theta+24)$, respectively. Additionally, Shanker [21] introduced the Suja distribution, which is a two-component mixture of exponential(θ) and Gamma($5,\theta$) distributions with mixing proportion $\theta^4/(\theta^4+24)$.

Sen et al. [14] introduced the xgamma distribution, which is a two-component mixture of exponential(θ) and Gamma($3,\theta$) distributions with mixing proportions $\theta/(1+\theta)$ and $1/(1+\theta)$, respectively. Chesneau et al. [4] introduced the 2S-Lindley distribution, which represents the sum of two independent and identically distributed Lindley random variables. This distribution can be expressed as a three-component mixture of Gamma($4,\theta$), Gamma($3,\theta$), and Gamma($2,\theta$) distributions with mixing proportions $1/(1+\theta)^2$, $2\theta/(1+\theta)^2$, and $\theta^2/(1+\theta)^2$, respectively. Chouia and Zeghdoudi [5] proposed the XLindley distribution, which is a two-component mixture of exponential(θ) and Lindley(θ) distributions with mixing proportions $1/(1+\theta)^2$ and $(\theta^2+2\theta)/(1+\theta)^2$, respectively. Ahsan-ul Haq [2] introduced the Haq distribution, which is a two-component mixture of exponential(θ) and xgamma(θ) distributions with mixing proportions $\theta/(1+\theta)$ and $1/(1+\theta)$, respectively. Saaidia and Zeghdoudi [12] introduced the ZLindley distribution, which is another two-component mixture of exponential(θ) and Gamma($2,\theta$) distributions with mixing proportions $1/(2(1+\theta))$ and $(1+2\theta)/(2(1+\theta))$, respectively. For further examples of such mixture-based distributions, the reader is referred to Messaadia and Zeghdoudi [11], Aijaz et al. [3], Radhakrishnan et al. [20], Elgarhy [6], and Saaidia et al. [13].

In line with this ongoing effort to develop more flexible one-parameter lifetime distributions, Khan and Ruidas [8] introduced the Mgamma distribution. This distribution is formulated as a finite mixture of two Gamma components: Gamma($2,\theta$) and Gamma($3,\theta$), with mixing probability $\theta/(1+\theta)$. The Mgamma distribution is a novel one-parameter continuous distribution that offers greater flexibility in modeling real-world data, particularly datasets that exhibit positive skewness and heavy tails. Its applicability has been demonstrated across various fields, including engineering, hydrology, medicine, and queueing systems. In the field of dentistry, for instance, such distributions are particularly useful for modeling the skewed nature of periodontal bone loss measurements or the survival times of dental implants.

The probability density function (PDF) of the Mgamma distribution is given by:

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

The corresponding cumulative distribution function (CDF) is expressed as:

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

One of the key advantages of the Mgamma distribution is its ability to exhibit non-monotone hazard rates, which are often observed in practice but cannot be captured by simpler one-parameter models such as the exponential or Lindley distributions. The hazard rate function (HRF) based on Equations (1) and (2) is defined as the ratio of the PDF to the survival function:

$$h(x;\theta)=\frac{f(x;\theta)}{1-F(x;\theta)}=\frac{\theta^3 x\left(1+\frac{x}{2}\right)}{1+\theta+\theta x\left(1+\theta+\frac{\theta x}{2}\right)}$$

Khan and Ruidas [8] investigated the structural properties of the Mgamma distribution and demonstrated that its PDF exhibits right-skewed behavior for all values of θ, with the distribution becoming more concentrated near zero as θ increases. They also showed that the HRF demonstrates remarkable flexibility: for small values of θ, it is increasing; for moderate values, it exhibits a bathtub-shaped pattern; and for large values, it becomes monotonically increasing, similar to the Lindley distribution. They further argued that this flexibility in hazard behavior makes the Mgamma distribution particularly attractive for reliability engineering, survival analysis, and other fields where non-monotone hazard rates are commonly encountered.

In this paper, we focus on the estimation of the parameter θ of the Mgamma distribution. We derive and compare three classical estimators: the MLE, the MoM estimator, and the UMVUE. We demonstrate that the MLE and MoM

estimators are equivalent and obtain their closed-form expressions. Additionally, we develop a comprehensive Bayesian framework for estimating θ using Lindley's approximation and MCMC techniques. The performance of all proposed estimators is evaluated through an extensive simulation study in terms of bias, mean squared error, and interval coverage. The practical applicability of the Mgamma distribution is illustrated using real orthodontic tooth movement data.

The remainder of this paper is organized as follows. Section 2 presents the derivation of the MLE and MoM estimators, establishes their equivalence, and provides the closed-form solution, along with the construction of the asymptotic confidence interval. Section 3 is dedicated to the derivation of the UMVUE using the Lehmann-Scheffé theorem, and also establishes the asymptotic (limiting) distributions of the proposed classical estimators. Section 4 introduces Bayesian estimation methods, including Lindley's approximation and MCMC techniques. Section 5 presents the simulation study comparing the performance of all estimators. Section 6 provides an application to real orthodontic tooth movement data. Finally, Section 7 concludes the paper with a summary of findings and suggestions for future research.

2. MLE and MoM estimator

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n drawn from the Mgamma distribution. Then the likelihood function according to Equations (1) and (2) is:

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

The log-likelihood function (ignoring additive constants) is:

$$\ell(\theta) = 3n\log(\theta) - n\log(1+\theta) - \theta\sum_{i=1}^{n} x_i.$$

Differentiating with respect to θ and setting to zero gives the score equation:

$$\frac{d\ell(\theta)}{d\theta} = \frac{3n}{\theta} - \frac{n}{1+\theta} - \sum_{i=1}^{n} x_i = 0.$$

By letting $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ yields the MLE estimating equation:

$$\frac{3}{\theta} - \frac{1}{1+\theta} = \bar{x}. \tag{4}$$

Simultaneously, the first population moment of the Mgamma distribution is:

$$E[X] = \frac{2\theta+3}{\theta(1+\theta)}.$$

Equating this to the sample mean $\bar{x}$ gives the MoM estimating equation:

$$\frac{2\theta+3}{\theta(1+\theta)} = \bar{x}, \tag{5}$$

which is exactly the left-hand side of Equation (4). Therefore, both the MLE and MoM estimating equations reduce to the same quadratic equation:

$$\bar{x}\,\theta^2 + (\bar{x}-2)\theta - 3 = 0. \tag{6}$$

The closed-form explicit estimator, valid for both MLE and MoM, is the positive root of Equation (6):

$$\hat{\theta}_{MLE} = \hat{\theta}_{MoM} = \frac{2 - \bar{x} + \sqrt{\bar{x}^2 + 8\bar{x} + 4}}{2\bar{x}}, \quad \bar{x} > 0. \tag{7}$$

Thus, for the Mgamma distribution, the MLE and the MoM estimator are identical.

2.1. Remark on the original paper

Although a closed-form explicit solution is readily available from Equation (7), it is important to note that in the original paper by Khan and Ruidas [8], the authors did not utilize this explicit expression to estimate the parameter. Instead, they solved the non-linear score equation $\frac{d\ell(\theta)}{d\theta} = 0$ using the Newton-Raphson iterative procedure. The Newton-Raphson iteration is given by:

$$\theta^{(k+1)} = \theta^{(k)} - \frac{\ell'(\theta^{(k)})}{\ell''(\theta^{(k)})},$$

with an appropriate initial value (e.g., $\theta^{(0)} = 2/\bar{x}$). The iteration continues until convergence is achieved.

2.2. ACI

The asymptotic distribution of the MLE is normal:

$$\hat{\theta} \sim \mathcal{N}\left(\theta, I_n^{-1}(\theta)\right),$$

where $I_n(\theta)$ is the observed Fisher information for the entire sample of size n, given by:

$$I_n(\theta) = -\frac{d^2\ell(\theta)}{d\theta^2} = \frac{3n}{\theta^2} - \frac{n}{(1+\theta)^2} = n \cdot I(\theta), \tag{8}$$

where $\ell(\theta)$ is the log-likelihood function and $I(\theta)$ denotes the Fisher information per observation (see Section 3.1). Consequently, an asymptotic $100(1-\alpha)\%$ confidence interval for θ is:

$$\hat{\theta} \pm z_{\alpha/2}\sqrt{I_n^{-1}(\hat{\theta})}.$$

3. UMVUE

To derive the UMVUE for θ, we utilize the Lehmann-Scheffé theorem. The first step is to find an unbiased estimator based on a single observation. Consider the function:

$$W(X) = \frac{2(X+1)}{X(X+2)}. \tag{9}$$

Its expectation is:

$$E[W(X)] = \int_0^\infty \frac{2(x+1)}{x(x+2)} \cdot \frac{\theta^3}{1+\theta}\frac{x(x+2)}{2} e^{-\theta x} dx = \frac{\theta^3}{1+\theta}\int_0^\infty (x+1)e^{-\theta x} dx = \theta.$$

Thus, $W(X)$ is an unbiased estimator for θ.

The sum $T = \sum_{i=1}^n X_i$ is a complete and sufficient statistic for θ. By the Lehmann-Scheffé theorem, the UMVUE is the conditional expectation:

$$\hat{\theta}_{UMVUE}(t) = E[W(X_1)|T = t].$$

The PDF of T is:

$$f_T(t) = \frac{\theta^{3n}}{(1+\theta)^n} e^{-\theta t} \sum_{k=0}^{n} \binom{n}{k} \frac{t^{3n-k-1}}{\Gamma(3n-k)}, \quad t > 0.$$

The PDF of the sum of the remaining $n-1$ variables is:

$$f_{T_{n-1}}(u) = \frac{\theta^{3n-3}}{(1+\theta)^{n-1}} e^{-\theta u} \sum_{m=0}^{n-1} \binom{n-1}{m} \frac{u^{3n-m-4}}{\Gamma(3n-m-3)}.$$

Using the definition of conditional expectation:

$$E[W(X_1)|T=t] = \frac{\int_0^t W(u) f_{X_1}(u) f_{T_{n-1}}(t-u) du}{f_T(t)}.$$

Substituting the expressions and simplifying, we obtain:

$$\int_0^t W(u) f_{X_1}(u) f_{T_{n-1}}(t-u) du = \frac{\theta^{3n}}{(1+\theta)^n} e^{-\theta t} \sum_{m=0}^{n-1} \binom{n-1}{m} \frac{1}{\Gamma(3n-m-3)} \times \int_0^t (u+1)(t-u)^{3n-m-4} du.$$

Evaluating the integral:

$$\int_0^t (u+1)(t-u)^{3n-m-4} du = \frac{t^{3n-m-2}\Gamma(3n-m-3)}{\Gamma(3n-m-1)} + \frac{t^{3n-m-3}}{3n-m-3}.$$

Therefore, the UMVUE is obtained in closed form as:

$$\hat{\theta}_{UMVUE}(t) = \frac{\sum_{m=0}^{n-1} \binom{n-1}{m} \left[\frac{t^{3n-m-2}}{\Gamma(3n-m-1)} + \frac{t^{3n-m-3}}{(3n-m-3)\Gamma(3n-m-3)} \right]}{\sum_{k=0}^{n} \binom{n}{k} \frac{t^{3n-k-1}}{\Gamma(3n-k)}}, \tag{10}$$

where $t = \sum_{i=1}^{n} x_i > 0$, $\Gamma(\cdot)$ is the Gamma function, and the term $\frac{t^{3n-m-3}}{(3n-m-3)\Gamma(3n-m-3)}$ in Equation (10) is defined only for $m \leq 3n-4$.

For $n = 1$, Equation (10) reduces to:

$$\hat{\theta}_{UMVUE}(t) = \frac{2(1+t)}{t(t+2)}, \tag{11}$$

which matches the unbiased estimator $W(X)$ given in Equation (9), as expected.

3.1. Asymptotic distribution of the estimators

In this subsection, we establish the limiting (asymptotic) distributions of the proposed classical estimators for the parameter θ of the Mgamma distribution.

Theorem 1 (Asymptotic normality of the MLE). Let $X_1, \ldots, X_n$ be a random sample from the Mgamma distribution with PDF given in Equation (1). Under the standard regularity conditions (see, e.g., Casella and Berger [22]), the MLE $\hat{\theta}_{MLE}$ satisfies:

$$\sqrt{n}\left(\hat{\theta}_{MLE} - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, I^{-1}(\theta)\right), \quad \text{as } n \to \infty,$$

where $I(\theta)$ is the Fisher information per observation, defined in Equation (8) in terms of the observed Fisher information $I_n(\theta)$ as $I(\theta) = I_n(\theta)/n$.

Proof. The result follows directly from the standard asymptotic theory of MLEs (see, e.g., Casella and Berger [22]). Under the regularity conditions—namely, that the support of the distribution does not depend on θ, that the log-likelihood is twice differentiable with respect to θ, and that the Fisher information $I(\theta)$ is finite and positive—the MLE is consistent and asymptotically normal. A Taylor expansion of the score function $\ell'(\hat{\theta}_{MLE}) = 0$ around the true value θ gives:

$$0 = \ell'(\theta) + (\hat{\theta}_{MLE} - \theta)\ell''(\theta) + O_p(1),$$

which, after rearrangement and application of the law of large numbers and the central limit theorem, yields the desired result.

Corollary 1 (Asymptotic normality of the MoM estimator). Since the MLE and the MoM estimator coincide for the Mgamma distribution (see Equation (7)), the MoM estimator shares the same limiting distribution:

$$\sqrt{n}\left(\hat{\theta}_{MoM} - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, I^{-1}(\theta)\right), \quad \text{as } n \to \infty.$$

Theorem 2 (Asymptotic normality of the UMVUE). Under the same regularity conditions, the UMVUE $\hat{\theta}_{UMVUE}$ derived in Section 3 is asymptotically efficient and satisfies:

$$\sqrt{n}\left(\hat{\theta}_{UMVUE} - \theta\right) \xrightarrow{d} \mathcal{N}\left(0, I^{-1}(\theta)\right), \quad \text{as } n \to \infty.$$

Proof. By the Lehmann-Scheffé theorem, the UMVUE is the unique unbiased estimator that is a function of the complete sufficient statistic $T = \sum_{i=1}^{n} X_i$. Since the UMVUE attains the Cramér-Rao lower bound asymptotically (see Lehmann and Casella [23]), it is asymptotically efficient and thus shares the same limiting distribution as the MLE. More formally, the asymptotic efficiency of the UMVUE implies that:

$$\sqrt{n}\left(\hat{\theta}_{UMVUE} - \theta\right) = \sqrt{n}\left(\hat{\theta}_{MLE} - \theta\right) + o_p(1),$$

and the result follows from Theorem 1 and Slutsky's theorem.

These asymptotic results justify the construction of the ACI for θ presented in Section 2 and provide the theoretical basis for comparing the proposed estimators in the simulation study.

4. Bayesian estimation

In this section, we consider Bayesian estimation of the parameter θ for the Mgamma distribution. Bayesian inference combines prior information about the parameter with the observed data through the likelihood function to obtain the posterior distribution. We employ a gamma prior distribution for θ, which is a natural conjugate prior for many lifetime distributions. The gamma prior has the following density:

$$\pi(\theta) = \frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta}, \quad \theta > 0,\ a, b > 0, \tag{12}$$

where a and b are the hyperparameters representing the shape and rate parameters of the prior distribution, respectively. This prior is chosen for its flexibility and its ability to represent various prior beliefs about θ by adjusting the hyperparameters. In the absence of prior information, non-informative priors such as $a = b = 0$ (Jeffreys prior) or $a = b = 0.001$ can be used.

Combining the gamma prior in Equation (12) with the likelihood function in Equation (3), the posterior distribution of θ given the data $x = (x_1, \ldots, x_n)$ is proportional to:

$$\pi(\theta|x) \propto L(\theta|x)\,\pi(\theta) = \frac{\theta^{3n+a-1}}{(1+\theta)^n}\exp\left[-\theta\left(b + \sum_{i=1}^{n} x_i\right)\right]\prod_{i=1}^{n} x_i\left(1 + \frac{x_i}{2}\right).$$

Thus, the posterior distribution can be expressed as:

$$\pi(\theta|x) = \frac{1}{C}\frac{\theta^{3n+a-1}}{(1+\theta)^n}\exp\left[-\theta\left(b+\sum_{i=1}^{n}x_i\right)\right], \tag{13}$$

where C is the normalizing constant given by:

$$C = \int_0^\infty \frac{\theta^{3n+a-1}}{(1+\theta)^n}\exp\left[-\theta\left(b+\sum_{i=1}^{n}x_i\right)\right]d\theta.$$

The posterior distribution in Equation (13) does not belong to any standard family of distributions, making direct analytical inference difficult. Therefore, we employ approximation methods such as Lindley's approximation and MCMC techniques for point and interval estimation.

4.1. Lindley's approximation

Lindley's approximation (Lindley, [10]) is a widely used method for computing posterior expectations and obtaining Bayesian estimates under a squared error loss function. For a parameter θ with posterior density $\pi(\theta|x)$, the posterior expectation of a function $g(\theta)$ is:

$$E[g(\theta)|x] = \frac{\int g(\theta)\exp[\ell(\theta)+\rho(\theta)]d\theta}{\int \exp[\ell(\theta)+\rho(\theta)]d\theta},$$

where $\ell(\theta) = \log L(\theta|x)$ is the log-likelihood function and $\rho(\theta) = \log\pi(\theta)$ is the log-prior density. Under the squared error loss function, the Bayesian estimator of θ is the posterior mean, i.e., $g(\theta) = \theta$.

Lindley's approximation for the posterior mean up to order $O(n^{-2})$ is given by:

$$\hat{\theta}_B \approx \theta_0 + \frac{1}{2}\sigma^2(3\ell_3+\ell_1) + \sigma^2(\ell_2\rho_1+\rho_2) + \frac{1}{2}\sigma^4(\ell_1\ell_3-\ell_2^2), \tag{14}$$

where θ_0 is the MLE of θ, and

$$\ell_i = \left.\frac{\partial^i\ell(\theta)}{\partial\theta^i}\right|_{\theta=\theta_0}, \quad \rho_i = \left.\frac{\partial^i\rho(\theta)}{\partial\theta^i}\right|_{\theta=\theta_0}, \quad \sigma^2 = -\frac{1}{\ell_2}.$$

Since $\ell_1 = 0$ at the MLE, the approximation simplifies to:

$$\hat{\theta}_B \approx \theta_0 - \frac{3\ell_3}{2\ell_2^2} - \frac{\rho_1}{\ell_2} - \frac{\rho_2}{\ell_2^2}. \tag{15}$$

For the Mgamma distribution, the required derivatives of the log-likelihood function are:

$$\ell_1 = \frac{3n}{\theta_0} - \frac{n}{1+\theta_0} - \sum_{i=1}^{n}x_i = 0,$$

$$\ell_2 = -\frac{3n}{\theta_0^2} + \frac{n}{(1+\theta_0)^2} = -I(\theta_0),$$

$$\ell_3 = \frac{6n}{\theta_0^3} - \frac{2n}{(1+\theta_0)^3}.$$

For the gamma prior $\pi(\theta) \propto \theta^{a-1}e^{-b\theta}$ given in Equation (12), the log-prior density and its derivatives are:

$$\rho(\theta) = (a-1)\log\theta - b\theta,$$

$$\rho_1 = \frac{a-1}{\theta_0} - b,$$

$$\rho_2 = -\frac{a-1}{\theta_0^2}.$$

Substituting ℓ_2, ℓ_3, ρ_1, and ρ_2 into Equation (15) yields the explicit Lindley estimator for the Mgamma distribution:

$$\hat{\theta}_{Lindley} = \theta_0 - \frac{3\left(\frac{6n}{\theta_0^3} - \frac{2n}{(1+\theta_0)^3}\right)}{2\left(-\frac{3n}{\theta_0^2} + \frac{n}{(1+\theta_0)^2}\right)^2} - \frac{\frac{a-1}{\theta_0} - b}{-\frac{3n}{\theta_0^2} + \frac{n}{(1+\theta_0)^2}} - \frac{-\frac{a-1}{\theta_0^2}}{\left(-\frac{3n}{\theta_0^2} + \frac{n}{(1+\theta_0)^2}\right)^2}. \tag{16}$$

This closed-form expression provides a computationally efficient Bayesian estimator that incorporates both the likelihood information (through ℓ_2 and ℓ_3) and the prior information (through ρ_1 and ρ_2).

It is important to note that Equation (16) contains all necessary terms for accurate Bayesian estimation. Unlike some simplified versions found in the literature that omit the ρ_2 term or incorrectly combine the ℓ_3 and ρ_1 terms, Equation (16) maintains the correct algebraic structure derived from the full Lindley formula in Equation (14). Omitting the ρ_2 term can lead to biased estimates, particularly when the prior is informative and the sample size is small.

The approximate posterior variance can be obtained from the inverse of the observed Fisher information adjusted by the prior:

$$Var(\theta|x) \approx -\frac{1}{\ell_2 + \rho_2}, \tag{17}$$

which provides a measure of uncertainty for the Bayesian point estimate. However, for more accurate interval estimation, we recommend the MCMC approach described in the following subsection, as it provides the full posterior distribution without relying on asymptotic approximations.

4.2. MCMC method

To obtain more accurate Bayesian inference, including HPD intervals, we employ MCMC methods, specifically the Metropolis-Hastings (M-H) algorithm. The M-H algorithm generates samples from the posterior distribution in Equation (13) without requiring the normalizing constant C.

The Metropolis-Hastings algorithm proceeds as follows:

1. Start with an initial value $\theta^{(0)}$, typically chosen as the MLE $\hat{\theta}_{MLE}$.
2. For $t = 1, 2, \ldots, N$:
   (a) Propose a candidate value θ^* from a proposal distribution $q(\theta^*|\theta^{(t-1)})$. A common choice is the normal distribution $\mathcal{N}(\theta^{(t-1)}, \tau^2)$, where τ is a tuning parameter chosen to achieve a reasonable acceptance rate.
   (b) Compute the acceptance probability:

$$\alpha = \min\left\{1, \frac{\pi(\theta^*|x)q(\theta^{(t-1)}|\theta^*)}{\pi(\theta^{(t-1)}|x)q(\theta^*|\theta^{(t-1)})}\right\}.$$

   (c) Generate $u \sim \mathcal{U}(0,1)$. If $u \leq \alpha$, accept θ^* and set $\theta^{(t)} = \theta^*$; otherwise, set $\theta^{(t)} = \theta^{(t-1)}$.
3. After a burn-in period of B iterations, the remaining $M = N - B$ samples $\{\theta^{(B+1)}, \ldots, \theta^{(N)}\}$ are used for posterior inference.

The Bayesian point estimates under the squared error loss function are obtained as the posterior mean:

$$\hat{\theta}_{MCMC} = \frac{1}{M} \sum_{t=B+1}^{N} \theta^{(t)}.$$

The posterior variance and other summary statistics can be computed directly from the MCMC samples. The posterior median and credible intervals are obtained from the empirical quantiles of the MCMC sample.

4.3. HPD interval

The HPD interval is the Bayesian analogue of the confidence interval and is defined as the shortest interval containing a specified posterior probability $(1-\alpha) \times 100\%$. For a unimodal posterior distribution, the HPD interval can be constructed as:

$$\{\theta : \pi(\theta|x) \geq k_\alpha\},$$

where k_α is the largest constant such that the posterior probability of the region is at least $1-\alpha$. In practice, the HPD interval is computed using the MCMC sample by finding the interval with the shortest length among all intervals that contain $100(1-\alpha)\%$ of the posterior mass.

The steps to compute the HPD interval from MCMC samples are:

1. Sort the MCMC samples in ascending order: $\theta_{(1)} < \theta_{(2)} < \cdots < \theta_{(M)}$.
2. For each $i = 1, 2, \ldots, M - M(1-\alpha)$, compute the interval $[\theta_{(i)}, \theta_{(i+M(1-\alpha))}]$ and its length.
3. The HPD interval is the interval with the smallest length among all these intervals.

5. Simulation

To evaluate the performance of the proposed estimation methods for the Mgamma distribution, an extensive simulation study is conducted. The simulation study compares the classical estimators (MLE and UMVUE), the Bayesian Lindley approximation, and the Bayesian MCMC approach under different prior specifications. The performance of the estimators is evaluated in terms of bias and mean squared error (MSE). For the interval estimation, the ACI based on the MLE and the HPD intervals from the MCMC approach are compared in terms of average length and coverage probability (CV).

The simulation setup considers four different values of the true parameter: $\theta = 0.7$, 2.0, 5.0, and 8.0, and four sample sizes: $n = 25$, 50, 75, and 100. For the Bayesian methods, three prior distributions are considered: (P1) non-informative Gamma$(0.001, 0.001)$, (P2) informative Gamma$(\theta, 1)$ with mean equal to the true parameter, and (P3) informative Gamma$(2\theta, 2)$ with mean equal to the true parameter and smaller variance. The number of replications for the classical and Lindley methods is 50,000. For the MCMC method, due to the high computational cost of running 50,000 MCMC chains (each with 50,000 iterations), the number of replications was set to 399, while each replication consists of 50,000 MCMC iterations with a burn-in period of 10,000. This configuration ensures reliable posterior inference while keeping the total computational time feasible.

The simulation results for the MLE and UMVUE are presented in Table 1, while Tables 2 and 3 show the Bayesian estimation results using Lindley's approximation and MCMC, respectively. Table 4 presents the HPD interval results for the MCMC method. For a comprehensive graphical comparison of the proposed methods, Figures 1 and 2 display heatmap representations of the MSE and average interval lengths, respectively, across all simulation scenarios. The results are discussed in detail below.

It should be noted that the column labeled "Length" in Tables 1 and 4 reports the average length of the confidence intervals (ACI and HPD, respectively), while the column labeled "CV" reports the corresponding coverage probabilities.

Table 1. Simulation results for the MLE and UMVUE of θ under the Mgamma distribution with different sample sizes and true parameter values.

		UMVUE		MLE		ACI	
θ	n	Bias	MSE	Bias	MSE	Length	CV (%)
0.7	25	-0.0002	0.0072	0.0102	0.0075	0.3313	95.15
	50	0.0000	0.0036	0.0052	0.0036	0.2325	94.99
	75	-0.0003	0.0023	0.0032	0.0024	0.1892	95.12
	100	-0.0000	0.0017	0.0026	0.0018	0.1637	95.05
2.0	25	0.0011	0.0646	0.0349	0.0683	0.9995	95.13
	50	-0.0005	0.0315	0.0162	0.0324	0.6997	95.10
	75	-0.0003	0.0211	0.0108	0.0215	0.5696	94.98
	100	-0.0006	0.0158	0.0077	0.0160	0.4925	94.94
5.0	25	0.0013	0.4512	0.0943	0.4785	2.6335	95.14
	50	-0.0005	0.2241	0.0456	0.2307	1.8432	94.72
	75	-0.0032	0.1467	0.0273	0.1494	1.4992	95.04
	100	0.0007	0.1100	0.0236	0.1116	1.2973	95.14
8.0	25	-0.0021	1.2152	0.1515	1.2880	4.3028	94.96
	50	0.0005	0.5911	0.0765	0.6089	3.0132	94.99
	75	0.0018	0.3926	0.0523	0.4005	2.4526	94.97
	100	0.0001	0.2953	0.0379	0.2997	2.1200	94.90

Table 2. Bayesian estimation results using Lindley's approximation for the Mgamma distribution under three different priors: P1: Gamma(0.001, 0.001), P2: Gamma(θ, 1), and P3: Gamma(2θ, 2).

		P1		P2		P3	
θ	n	Bias	MSE	Bias	MSE	Bias	MSE
0.7	25	-0.0051	0.0070	-0.0055	0.0071	-0.0055	0.0069
	50	-0.0024	0.0036	-0.0026	0.0035	-0.0021	0.0035
	75	-0.0019	0.0024	-0.0018	0.0023	-0.0016	0.0023
	100	-0.0009	0.0018	-0.0017	0.0017	-0.0013	0.0018
2.0	25	-0.0156	0.0652	-0.0177	0.0621	-0.0200	0.0591
	50	-0.0046	0.0320	-0.0084	0.0313	-0.0090	0.0301
	75	-0.0032	0.0210	-0.0065	0.0210	-0.0039	0.0203
	100	-0.0035	0.0160	-0.0027	0.0155	-0.0060	0.0151
5.0	25	-0.0342	0.4583	-0.0617	0.4027	-0.0640	0.3625
	50	-0.0164	0.2214	-0.0214	0.2105	-0.0268	0.1960
	75	-0.0113	0.1458	-0.0229	0.1380	-0.0187	0.1335
	100	-0.0099	0.1109	-0.0085	0.1088	-0.0107	0.1044
8.0	25	-0.0876	1.1676	-0.1174	1.0244	-0.1276	0.8507
	50	-0.0418	0.5859	-0.0445	0.5458	-0.0495	0.4984
	75	-0.0148	0.3984	-0.0350	0.3745	-0.0309	0.3438
	100	-0.0208	0.2906	-0.0170	0.2857	-0.0205	0.2641

5.1. Comparison of classical estimators

The simulation results for the MLE and UMVUE reveal several important patterns regarding their performance characteristics. Both estimators exhibit the expected asymptotic behavior, with bias and MSE decreasing as the sample size increases, confirming their consistency properties. The UMVUE demonstrates superior performance

Table 3. Bayesian estimation results using MCMC for the Mgamma distribution under three different priors: P1: Gamma(0.001, 0.001), P2: Gamma(θ, 1), and P3: Gamma(2θ, 2).

		P1		P2		P3	
θ	n	Bias	MSE	Bias	MSE	Bias	MSE
0.7	25	0.0088	0.0073	0.0153	0.0078	0.0098	0.0065
	50	0.0049	0.0033	0.0017	0.0044	0.0028	0.0034
	75	0.0056	0.0027	0.0045	0.0020	0.0027	0.0023
	100	0.0045	0.0017	0.0006	0.0019	0.0036	0.0018
2.0	25	0.0435	0.0643	0.0335	0.0699	0.0289	0.0661
	50	0.0181	0.0350	0.0224	0.0294	0.0224	0.0337
	75	0.0220	0.0222	0.0146	0.0219	0.0087	0.0225
	100	0.0115	0.0163	0.0146	0.0147	-0.0047	0.0166
5.0	25	0.0876	0.4425	0.0935	0.3888	0.0586	0.3375
	50	0.0281	0.2047	0.0381	0.2149	0.0310	0.1971
	75	0.0201	0.1372	0.0372	0.1357	0.0804	0.1689
	100	0.0164	0.0993	0.0285	0.0905	0.0009	0.0993
8.0	25	0.2212	1.4127	0.1231	0.9991	0.0623	0.7269
	50	-0.0003	0.6025	0.0405	0.5655	0.0415	0.4418
	75	0.0793	0.4102	0.0427	0.3674	0.0567	0.3372
	100	0.0038	0.2727	0.0065	0.2546	0.0263	0.2660

Table 4. Bayesian estimation results using MCMC for the Mgamma distribution under three different priors: P1: Gamma(0.001, 0.001), P2: Gamma(θ, 1), and P3: Gamma(2θ, 2).

		P1		P2		P3	
θ	n	Length	CV (%)	Length	CV (%)	Length	CV (%)
0.7	25	0.3274	94.99	0.3292	95.24	0.3249	95.74
	50	0.2309	95.49	0.2289	92.23	0.2289	95.24
	75	0.1886	92.48	0.1882	97.24	0.1875	96.24
	100	0.1632	94.99	0.1620	94.74	0.1628	94.99
2.0	25	0.9935	95.99	0.9733	95.24	0.9573	92.98
	50	0.6953	93.23	0.6913	97.74	0.6860	93.48
	75	0.5692	94.49	0.5637	94.24	0.5593	94.24
	100	0.4904	94.99	0.4893	95.24	0.4829	92.98
5.0	25	2.5987	95.99	2.4986	96.49	2.3919	95.49
	50	1.8226	96.24	1.7880	94.74	1.7492	96.24
	75	1.4889	96.74	1.4711	96.24	1.4646	94.74
	100	1.2883	96.49	1.2771	95.74	1.2562	95.99
8.0	25	4.2587	93.98	3.9447	95.24	3.7016	97.24
	50	2.9601	93.23	2.8739	95.24	2.7833	96.24
	75	2.4463	93.73	2.3798	93.73	2.3277	95.74
	100	2.0992	95.99	2.0650	94.99	2.0330	94.74

in terms of bias, remaining consistently closer to zero across all sample sizes and parameter values compared to the MLE, which is expected given its theoretical unbiasedness property.

In terms of MSE, the UMVUE and MLE show similar performance, as also illustrated in the heatmap of Figure 1, with the difference being relatively small, particularly at smaller sample sizes. For $n = 25$, the MSE values of both estimators are close to each other across all parameter values, with the UMVUE showing a slight advantage. As the sample size increases, the MSE of both estimators decreases and they converge toward each other, indicating

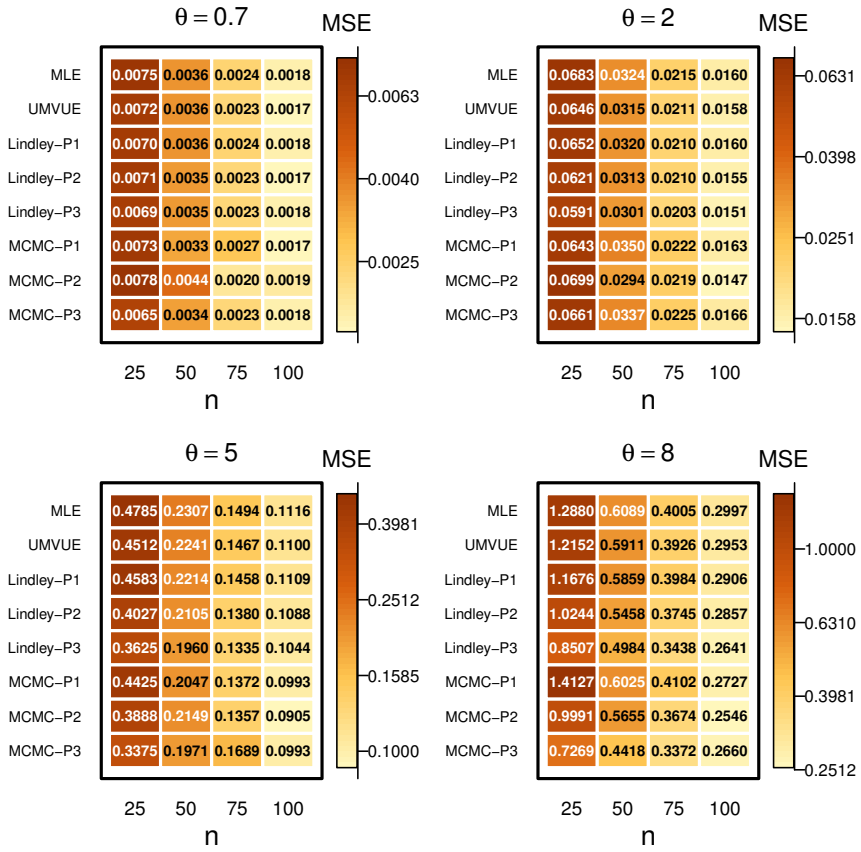


Figure 1. Heatmap comparison of the MSE for all proposed estimation methods across different sample sizes and true parameter values.

that both estimators are asymptotically efficient. This convergence is expected as the MLE's bias diminishes with increasing sample size, making its MSE increasingly comparable to that of the UMVUE.

5.2. Comparison of Bayesian estimators

The Bayesian estimation results, obtained through both Lindley's approximation and the MCMC approach, show that the choice of prior has a meaningful impact on estimation performance. Among the three priors considered, the P3 prior (Gamma(2θ, 2)) consistently provides the best performance in terms of MSE for both Lindley and MCMC methods, as can be clearly observed in the heatmap of Figure 1, where the cells corresponding to the P3 prior generally exhibit lighter colors compared to those of the P1 and P2 priors. This superiority can be attributed to the smaller variance of the P3 prior, which makes it more informative while maintaining the same mean as the P2 prior.

The Lindley approximation, despite its computational simplicity, provides reasonable estimates that are comparable to the MCMC results, particularly for moderate to large sample sizes, as also visible in Figure 1. However, the MCMC approach generally yields superior performance, especially in challenging scenarios with larger parameter values. The improvement of MCMC over Lindley is most evident when examining the MSE values, where MCMC consistently provides smaller or comparable MSE across all scenarios.

Regarding the prior comparison, the P1 prior (non-informative) yields the largest MSE among the three priors, which is expected given that it provides no prior information to guide the estimation and is clearly reflected by the darker cells in Figure 1. The P2 and P3 priors, both centered at the true parameter value, provide

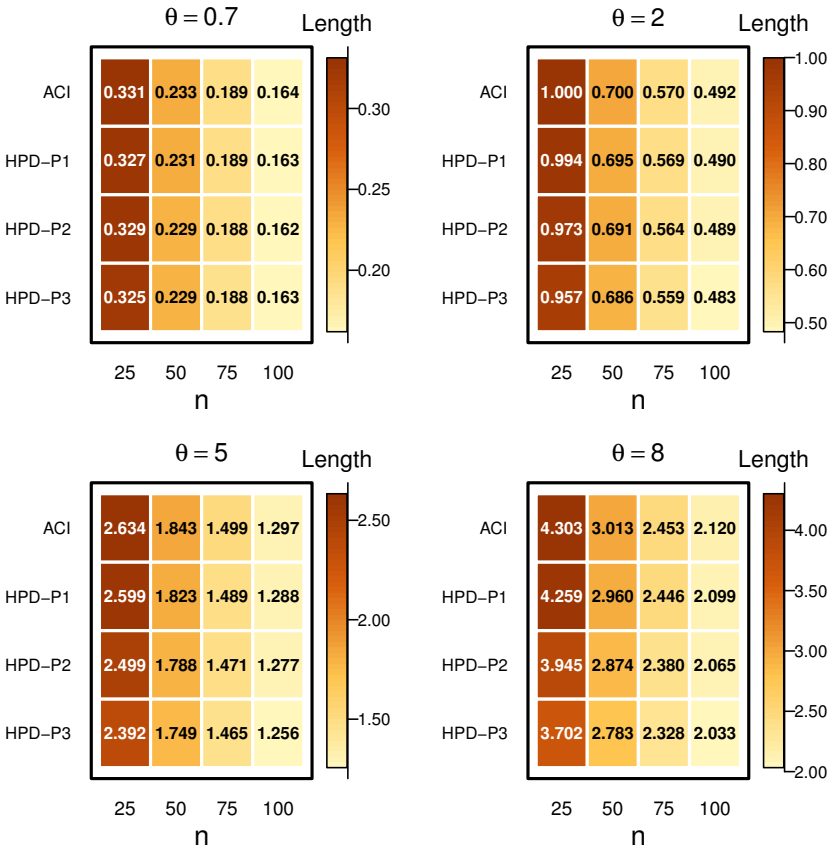


Figure 2. Heatmap comparison of the average interval lengths for the ACI and the HPD intervals under three prior specifications.

improved estimation efficiency. The P3 prior, with its smaller variance, offers the best performance, suggesting that incorporating stronger prior information when available can substantially improve estimation precision.

5.3. Overall comparison

Comparing all estimation methods collectively, several important conclusions can be drawn regarding their relative performance. The UMVUE stands out as providing the smallest bias among all methods, confirming its theoretical unbiasedness property, as also visible in the heatmap of Figure 1. The Bayesian MCMC method with the P3 prior generally provides the smallest MSE among all methods, indicating that incorporating informative prior information can lead to substantial improvements in estimation efficiency.

The trade-off between bias and MSE is evident when comparing the UMVUE and the MCMC estimators. While the UMVUE maintains near-zero bias, the MCMC estimators, particularly those with informative priors, provide smaller MSE at the cost of introducing some bias. This trade-off is particularly beneficial for larger parameter values where estimation is more challenging and the reduction in MSE outweighs the slight increase in bias.

Regarding interval estimation, the ACIs based on the MLE demonstrate excellent coverage properties, with all coverage probabilities remaining very close to the nominal 95% level across all simulation scenarios. This confirms that the asymptotic normal approximation works effectively for the Mgamma distribution. The average length of the ACI decreases appropriately with increasing sample size, reflecting the improved precision of the estimator. The HPD intervals from the MCMC approach offer competitive performance and, in most cases, provide shorter intervals than the ACI, as clearly illustrated in the heatmap of Figure 2. For all parameter values considered,

the HPD intervals under the non-informative P1 prior are consistently shorter than the corresponding ACI, with the difference becoming more pronounced as the true parameter value increases. This pattern indicates that the Bayesian approach yields more precise interval estimates, particularly for larger parameter values where the posterior distribution captures the skewness of the underlying model more effectively than the asymptotic normal approximation. Both methods, however, maintain coverage probabilities close to the nominal 95% level across all simulation scenarios, confirming the validity of both interval estimation approaches.

The computational efficiency of the different methods also merits consideration. The classical methods and the Lindley approximation are computationally efficient, making them suitable for routine applications. The MCMC approach, while computationally more demanding, provides the most comprehensive inference through full posterior sampling and is recommended when high precision is required and computational resources permit.

6. Application to orthodontic tooth movement data

Orthodontic tooth movement is a biological process in which the application of a controlled force induces bone resorption on the pressure side and bone apposition on the tension side, resulting in the gradual displacement of teeth through the alveolar bone. The rate of tooth movement varies among individuals and is influenced by numerous factors, including the magnitude and duration of the applied force, biological variability, and patient-specific characteristics. Consequently, the distribution of tooth movement rates often exhibits positive skewness and substantial variability, making it an appropriate candidate for modeling with flexible probability distributions.

In this section, we utilize a real dataset to illustrate the practical utility of the Mgamma distribution in modeling orthodontic tooth movement data. The dataset, originally reported by Abd El-Monsef and Abd El-Raouf [1], consists of measurements of tooth movement per week (in millimeters) resulting from orthodontic force applied using aligners. These data were collected from 100 orthodontic patients, with each observation representing the weekly rate of tooth displacement. The complete dataset is presented in Table 5.

Table 5. Weekly tooth movement data (in millimeters) for 100 orthodontic patients.

Tooth movement (mm)											
0.01	0.01	0.03	0.06	0.07	0.08	0.09	0.09	0.10	0.12	0.13	0.13
0.13	0.13	0.14	0.15	0.15	0.15	0.17	0.17	0.17	0.18	0.19	0.22
0.23	0.23	0.23	0.24	0.25	0.25	0.26	0.27	0.29	0.29	0.31	0.32
0.33	0.34	0.34	0.35	0.36	0.36	0.37	0.37	0.38	0.42	0.43	0.43
0.44	0.44	0.44	0.45	0.47	0.47	0.49	0.50	0.53	0.53	0.54	0.54
0.54	0.54	0.55	0.56	0.57	0.58	0.59	0.63	0.64	0.65	0.67	0.69
0.69	0.71	0.71	0.72	0.73	0.77	0.78	0.79	0.79	0.80	0.83	0.84
0.88	0.88	0.92	0.93	0.94	0.95	0.99	1.02	1.07	1.15	1.16	1.17
1.22	1.36	1.37	1.43								

The data presented in Table 5 range from 0.01 to 1.43 millimeters and exhibit a positively skewed distribution, with a concentration of observations at lower values and a gradually decreasing frequency toward higher values, which is characteristic of many biological and clinical measurements. This dataset is well-suited for evaluating the performance of the Mgamma distribution in capturing the skewness and variability commonly observed in orthodontic tooth movement studies. The application of the proposed estimation methods to this real dataset demonstrates the practical relevance and flexibility of the Mgamma distribution in clinical and biomedical research contexts.

6.1. Descriptive summary of the data

To gain a better understanding of the dataset, we provide some basic descriptive statistics. The sample consists of $n = 100$ observations, with a mean of $\bar{x} = 0.5017$ and a median of 0.4400. The sample standard deviation is 0.3413, and the variance is 0.1165. The data range from 0.0100 to 1.4300, with an interquartile range (IQR)

of 0.4825 (Q1 = 0.2300, Q3 = 0.7125). The skewness is 0.7209, indicating a positive skew, while the kurtosis is −0.1457, suggesting a slightly platykurtic distribution. The MLE of the Mgamma distribution parameter is $\hat{\theta}_{MLE} = 4.358$. These summary statistics confirm the presence of positive skewness and substantial variability in the tooth movement data, which justifies the use of flexible distributions such as the Mgamma distribution. Such heterogeneity in tooth movement rates is well-documented in orthodontic literature and underscores the need for flexible probability models capable of accommodating skewed and heavy-tailed distributions.

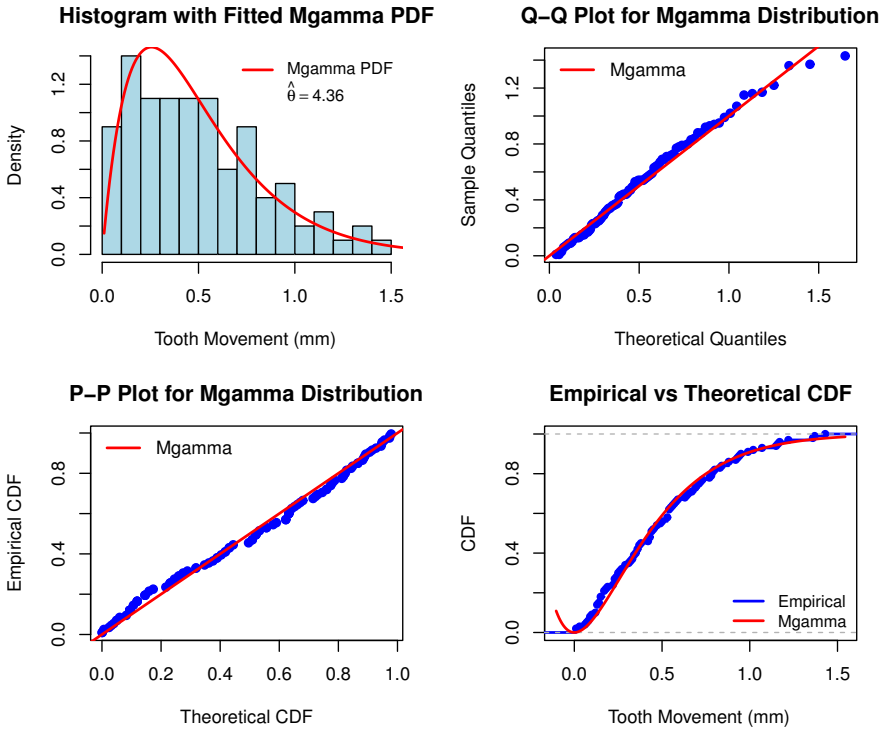


Figure 3. Diagnostic plots for the Mgamma distribution fitted to the orthodontic tooth movement data.

The diagnostic plots presented in Figure 3 provide a visual assessment of the goodness-of-fit for the Mgamma distribution. The histogram with the fitted Mgamma density curve demonstrates that the distribution adequately captures the shape of the data, particularly the concentration of observations at lower values and the decreasing frequency toward higher values. The Q-Q plot shows that the sample quantiles closely follow the theoretical quantiles, with most points lying near the reference line, indicating a reasonable fit. The P-P plot further supports this conclusion, as the empirical cumulative probabilities align well with the theoretical probabilities. Finally, the empirical cumulative distribution function closely matches the theoretical CDF of the Mgamma distribution, confirming that the Mgamma distribution provides a suitable model for the orthodontic tooth movement data. These diagnostic tools collectively validate the use of the Mgamma distribution for analyzing tooth movement rates and support the findings of the estimation procedures discussed in the preceding sections.

6.2. Goodness-of-fit analysis

To assess the adequacy of the Mgamma distribution compared to other competing distributions, we conducted goodness-of-fit tests including the Anderson-Darling (A-D), Cramér-von Mises (C-vM), and Kolmogorov-Smirnov (K-S) tests. The competing distributions considered in this analysis are the 2S-Lindley, Aradhana, Shanker, Akash, Devya, Sujatha, Suja, Lindley, Zlindley, Exponential, Xgamma, Haq, Burr-Hatke (Isaic-Maniu and Voda, [7]), and XLindley distributions. The PDFs and CDFs for these distributions (all parameterized by $\theta > 0$) are as follows:

- **2S-Lindley:**

$$f(x;\theta) = \frac{\theta^4}{(1+\theta)^2}\left(\frac{x^3}{6} + x^2 + x\right)e^{-\theta x}, \quad x > 0,$$
$$F(x;\theta) = 1 - \frac{1}{6(1+\theta)^2}\Big[\theta^3x^3 + (6\theta^3 + 3\theta^2)x^2 + (6\theta^3 + 12\theta^2 + 6\theta)x + (6\theta^2 + 12\theta + 6)\Big]e^{-\theta x}, \quad x > 0.$$

- **Aradhana:**

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

- **Shanker:**

$$f(x;\theta) = \frac{\theta^2}{\theta^2 + 1}(\theta + x)e^{-\theta x}, \quad x > 0,$$
$$F(x;\theta) = 1 - \frac{\theta^2 + \theta^2x + 1}{\theta^2 + 1}e^{-\theta x}, \quad x > 0.$$

- **Akash:**

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

- **Devya:**

$$f(x;\theta) = \frac{\theta^5}{\theta^4 + \theta^3 + 2\theta^2 + 6\theta + 24}(1 + x + x^2 + x^3 + x^4)e^{-\theta x}, \quad x > 0,$$
$$F(x;\theta) = 1 - \left(1 + \frac{\theta^4(x^4 + x^3 + x^2 + x) + \theta^3(4x^3 + 3x^2 + 2x)}{\theta^4 + \theta^3 + 2\theta^2 + 6\theta + 24} + \frac{6\theta^2(2x^2 + x) + 24\theta x}{\theta^4 + \theta^3 + 2\theta^2 + 6\theta + 24}\right)e^{-\theta x}, \quad x > 0.$$

- **Sujatha:**

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

- **Suja:**

$$f(x;\theta) = \frac{\theta^5}{\theta^4 + 24}(1 + x^4)e^{-\theta x}, \quad x > 0,$$
$$F(x;\theta) = 1 - \left(1 + \frac{\theta^4x^4 + 4\theta^3x^3 + 12\theta^2x^2 + 24\theta x}{\theta^4 + 24}\right)e^{-\theta x}, \quad x > 0.$$

- **Lindley:**

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

- **Zlindley:**

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

- **Exponential:**

$$f(x;\theta) = \theta e^{-\theta x}, \quad x > 0,$$
$$F(x;\theta) = 1 - e^{-\theta x}, \quad x > 0.$$

- **Xgamma:**

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

- **Haq:**

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

- **Burr-Hatke:**

$$f(x;\theta) = \frac{1+\theta(x+1)}{(x+1)^2}e^{-\theta x}, \quad x > 0,$$
$$F(x;\theta) = 1 - \frac{e^{-\theta x}}{x+1}, \quad x > 0.$$

- **XLindley:**

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

The goodness-of-fit results for the orthodontic tooth movement data are presented in Table 6, along with the MLEs of the parameter $\hat{\theta}$, the negative log-likelihood values ($-\ell$), and the corresponding p-values for each test.

The results demonstrate that the Mgamma distribution provides the best fit to the orthodontic tooth movement data among all considered distributions. The Mgamma distribution achieves the lowest negative log-likelihood value, indicating superior likelihood-based performance. Furthermore, the goodness-of-fit test statistics and their corresponding p-values confirm the excellent fit of the Mgamma distribution. The A-D, C-vM, and K-S tests all yield high p-values, indicating that the null hypothesis that the data follow the Mgamma distribution cannot be rejected at any conventional significance level. These findings provide strong statistical evidence in favor of the Mgamma distribution and confirm its adequacy for modeling the tooth movement data.

Table 6. Goodness-of-fit results for different distributions to the orthodontic tooth movement data.

Distribution	$\hat{\theta}$	$-\ell$	A-D	C-vM	K-S
Mgamma	4.358	24.218	0.618 (0.631)	0.076 (0.714)	0.064 (0.779)
2S-Lindley	4.688	24.230	0.651 (0.601)	0.079 (0.698)	0.067 (0.735)
Aradhana	3.127	27.615	2.084 (0.083)	0.342 (0.103)	0.104 (0.214)
Shanker	2.295	30.388	2.075 (0.084)	0.261 (0.175)	0.103 (0.221)
Akash	2.802	33.145	4.229 (0.007)	0.759 (0.009)	0.145 (0.027)
Devya	3.704	33.324	4.267 (0.006)	0.764 (0.009)	0.145 (0.027)
Sujatha	2.992	29.613	2.847 (0.033)	0.485 (0.043)	0.121 (0.101)
Suja	3.315	44.163	8.450 (<0.001)	1.657 (<0.001)	0.205 (<0.001)
Lindley	2.554	29.180	2.673 (0.040)	0.455 (0.052)	0.116 (0.128)
Zlindley	2.298	30.438	3.182 (0.022)	0.556 (0.029)	0.125 (0.082)
Exponential	1.993	31.025	3.402 (0.017)	0.600 (0.022)	0.128 (0.068)
Xgamma	3.027	32.613	4.225 (0.007)	0.764 (0.009)	0.146 (0.025)
Haq	2.342	32.873	4.096 (0.008)	0.739 (0.010)	0.141 (0.034)
Burr-Hatke	1.305	34.495	5.074 (0.003)	0.954 (0.003)	0.158 (0.012)
XLindley	2.186	30.772	3.276 (0.020)	0.573 (0.026)	0.126 (0.076)

6.3. Parameter estimation and inference

Having confirmed the adequacy of the Mgamma distribution for modeling the orthodontic tooth movement data through various goodness-of-fit analyses, we now proceed to estimate the parameter θ using both classical and Bayesian approaches. The classical estimation methods considered are the MLE and the UMVUE. For Bayesian estimation, we employ both Lindley's approximation and the Metropolis-Hastings algorithm within the MCMC framework under three different prior specifications: (i) non-informative Gamma(0.001, 0.001) prior, (ii) informative Gamma(4.35, 1) prior, and (iii) informative Gamma(8.7, 2) prior. The choice of the informative priors (ii) and (iii) is based on the MLE estimate $\hat{\theta}_{MLE} = 4.358$, which serves as the prior mean for both specifications. Specifically, the mean of the Gamma(4.35, 1) prior is 4.35, while the mean of the Gamma(8.7, 2) prior is 4.35, both being approximately equal to the MLE. However, the variances differ between the two informative priors: the variance of Gamma(4.35, 1) is 4.35, whereas the variance of Gamma(8.7, 2) is 2.175, which is smaller. This deliberate choice allows us to investigate the effect of prior informativeness on the posterior estimates, where the Gamma(8.7, 2) prior represents a more informative prior than the Gamma(4.35, 1) prior due to its smaller variance.

Table 7 presents the point estimates of θ obtained from all estimation methods, while Table 8 provides the interval estimates including the asymptotic 95% confidence interval based on the MLE and the 95% HPD intervals from the MCMC approach under each prior specification.

The point estimates presented in Table 7 demonstrate remarkable consistency across all estimation methods, with all estimates lying within a narrow range from 4.329 to 4.363. This agreement across classical methods, Lindley's approximation, and MCMC under different priors provides strong evidence for the robustness and reliability of the parameter estimates, suggesting that the data contain sufficient information to dominate the prior specifications.

The interval estimates in Table 8 further support the precision of the estimation. The HPD intervals obtained from the MCMC approach are comparable in width to the ACI based on the MLE. Notably, while the posterior means under the informative Gamma(4.35, 1) and Gamma(8.7, 2) priors are very close to the MLE, the HPD interval length under the Gamma(8.7, 2) prior (1.091) is shorter than that under the Gamma(4.35, 1) prior (1.097), indicating that the Gamma(8.7, 2) prior is more informative and yields more precise estimates compared to the Gamma(4.35, 1) prior. This suggests that the choice of an appropriately informative prior can improve estimation efficiency without introducing significant bias, as the posterior mean remains close to the MLE.

Table 7. Point estimates of θ for the orthodontic tooth movement data using classical and Bayesian methods.

Method	Prior	$\hat{\theta}$
MLE	—	4.358
UMVUE	—	4.339
Lindley	Gamma(0.001, 0.001)	4.329
Lindley	Gamma(4.35, 1)	4.329
Lindley	Gamma(8.7, 2)	4.329
MCMC	Gamma(0.001, 0.001)	4.361
MCMC	Gamma(4.35, 1)	4.358
MCMC	Gamma(8.7, 2)	4.363

Table 8. Interval estimates of θ for the orthodontic tooth movement data.

Method	Prior	95% Interval	Length
MLE (Asymptotic)	—	(3.800, 4.917)	1.117
MCMC	Gamma(0.001, 0.001)	(3.831, 4.958)	1.126
MCMC	Gamma(4.35, 1)	(3.803, 4.899)	1.097
MCMC	Gamma(8.7, 2)	(3.820, 4.912)	1.091

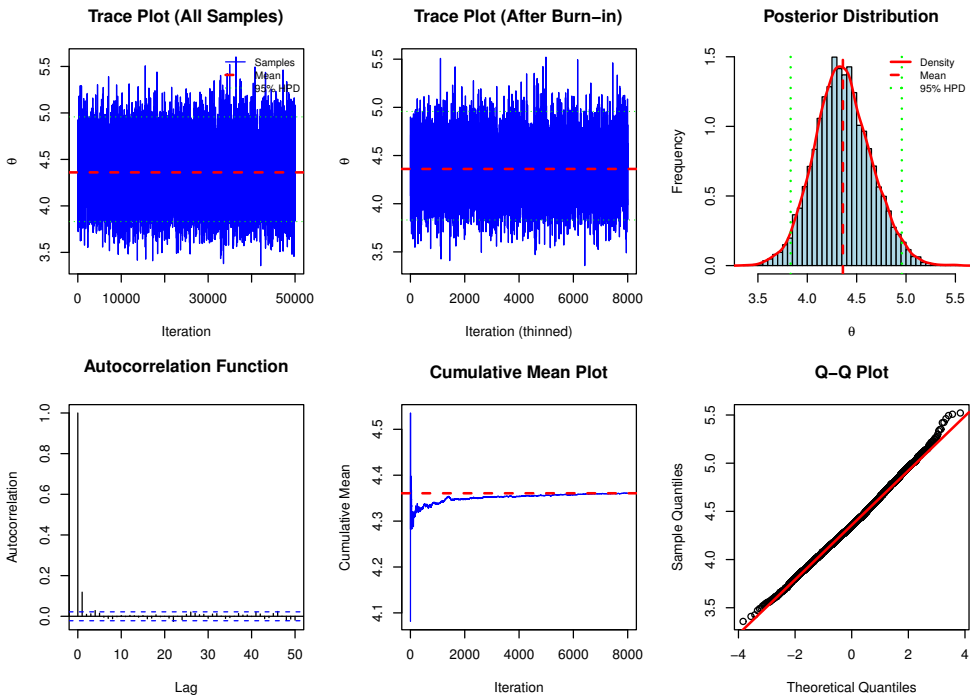


Figure 4. MCMC diagnostic plots for the Mgamma distribution under non-informative Gamma(0.001, 0.001) prior.

To further validate the Bayesian estimates obtained from the MCMC algorithm, Figures 4–6 present the corresponding diagnostic plots for the three prior specifications. Figure 4 shows the diagnostic plots under the non-informative Gamma(0.001, 0.001) prior, while Figures 5 and 6 display the corresponding plots under the informative Gamma(4.35, 1) and Gamma(8.7, 2) priors, respectively. For all three priors, the trace plots exhibit good mixing and stability around the posterior mean, indicating that the chains have converged to the target

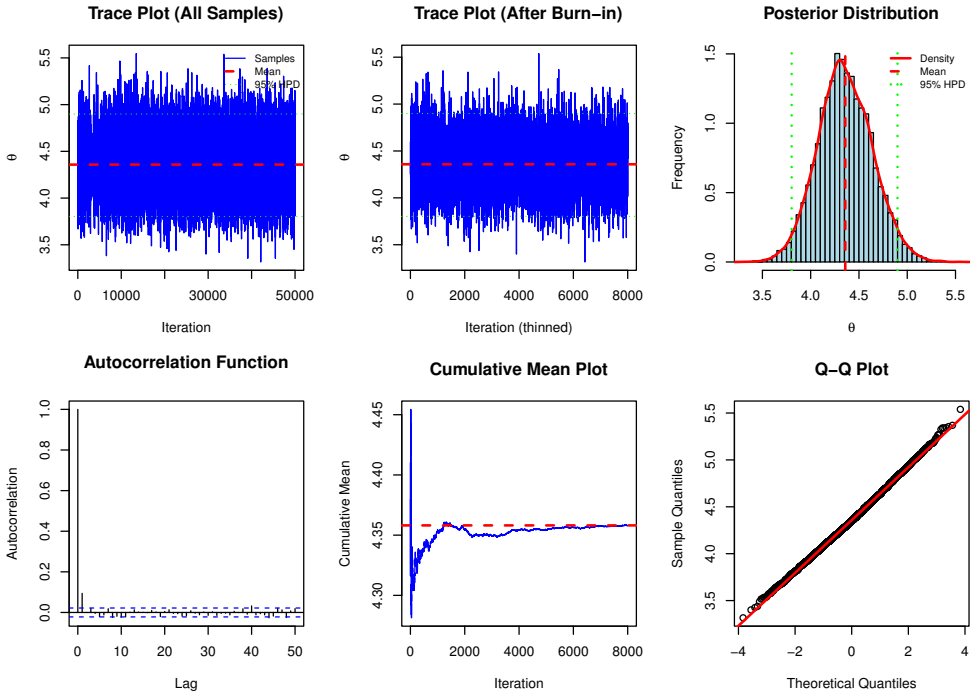


Figure 5. MCMC diagnostic plots for the Mgamma distribution under informative Gamma(4.35, 1) prior.

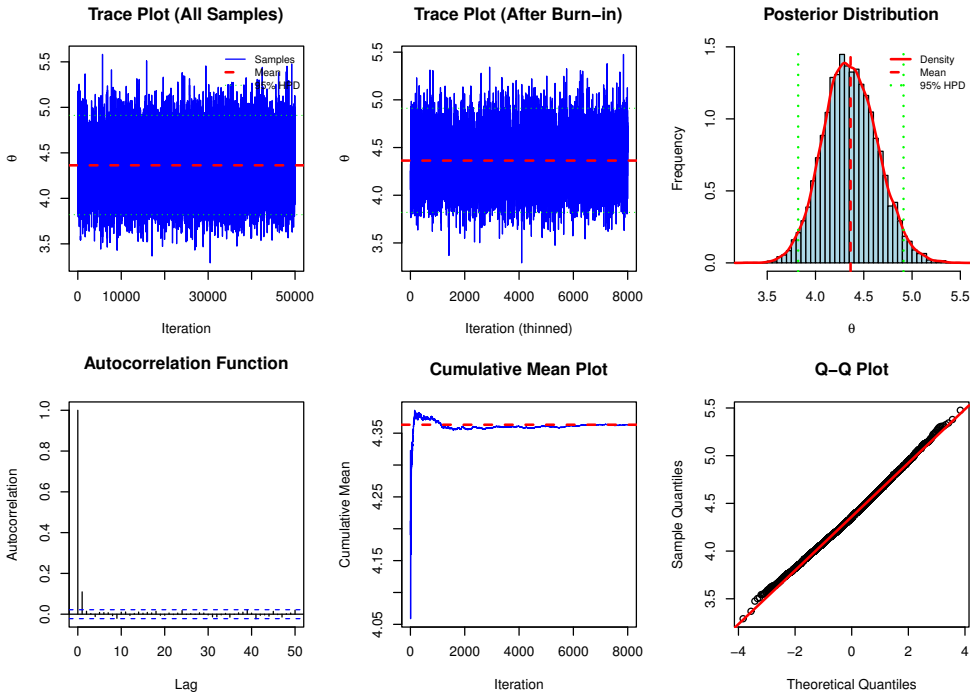


Figure 6. MCMC diagnostic plots for the Mgamma distribution under informative Gamma(8.7, 2) prior.

distribution. The autocorrelation functions decay rapidly to zero, confirming that the thinning procedure was effective in reducing sample dependence. The cumulative mean plots stabilize quickly, providing additional evidence of convergence. Furthermore, the Q-Q plots show that the posterior samples are approximately normally distributed.

7. Conclusion

In this paper, the Mgamma distribution has been introduced and investigated as a new one-parameter distribution for modeling lifetime data. The Mgamma distribution is formulated as a two-component mixture of Gamma distributions with shape parameters 2 and 3, providing high flexibility in modeling heterogeneous data. It was shown that the MLE and the MoM estimator are equivalent for this distribution and can be obtained in closed form. Additionally, the UMVUE was derived analytically using the Lehmann-Scheffé theorem. Bayesian estimators were computed using Lindley's approximation and the MCMC algorithm under three different prior specifications.

The simulation study revealed important patterns regarding the performance of the proposed estimators. The UMVUE exhibits the smallest bias among all methods, confirming its theoretical unbiasedness property. In terms of MSE, a direct comparison across all methods shows that the Bayesian MCMC approach with the informative P3 prior (Gamma(2θ, 2)) frequently provides the smallest MSE, particularly for larger parameter values ($\theta = 5.0$ and 8.0) and smaller sample sizes ($n = 25$ and 50). However, this advantage is not uniform across all scenarios; for smaller parameter values ($\theta = 0.7$ and 2.0) and larger sample sizes, the differences between UMVUE, Lindley, and MCMC become negligible, and all methods perform similarly. The Lindley approximation, despite its computational simplicity, provides competitive estimates, especially for moderate to large sample sizes, and often outperforms MCMC in terms of bias, though it generally yields slightly larger MSE than MCMC-P3.

Comparing the classical and Bayesian methods collectively, several practical recommendations emerge. If unbiasedness is the primary concern, the UMVUE is the preferred choice. If estimation efficiency (minimum MSE) is the priority, especially with small to moderate samples and large parameter values, the MCMC method with an informative prior (P3) is recommended, despite introducing some bias. For routine applications where computational speed is essential, the Lindley approximation with P3 offers a good balance between accuracy and simplicity. The MLE, while simple and asymptotically efficient, generally shows the largest MSE and bias among all methods and is therefore the least recommended when alternatives are available.

Regarding interval estimation, the ACI demonstrates excellent coverage properties, maintaining coverage probabilities very close to the nominal 95% level across all scenarios, confirming the validity of the asymptotic normal approximation for the Mgamma distribution. The HPD intervals from the MCMC approach also maintain coverage probabilities close to the nominal level and, in all simulation scenarios considered, provide shorter average lengths than the ACI. This advantage becomes more pronounced for larger parameter values ($\theta = 5.0$ and 8.0), where the posterior distribution effectively captures the skewness of the underlying model. Therefore, the HPD intervals are recommended when shorter and more precise interval estimates are desired, while the ACI remains a reliable and computationally simpler alternative.

The Mgamma distribution was fitted to orthodontic tooth movement data consisting of 100 observations. Goodness-of-fit test results confirmed the adequacy of the Mgamma distribution for modeling tooth movement data. The point estimates obtained from all methods showed remarkable consistency, and the interval estimates demonstrated high precision. The Mgamma distribution, with its ability to exhibit non-monotone hazard rates and excellent performance in fitting skewed data, is recommended as a powerful tool for modeling lifetime data across various fields, including engineering, biomedical sciences, medicine, and particularly orthodontic studies.

Extending the Mgamma distribution to two- and three-parameter versions, investigating additional theoretical properties such as stochastic orderings and various reliability measures, and applying it to censored and regression data are suggested for future work.

REFERENCES

1. M. M. E. Abd El-Monsef and M. M. Abd El-Raouf, *Rayleigh-half normal distribution for modeling tooth movement data*, Journal of Biopharmaceutical Statistics, vol. 30, no. 3, pp. 481–494, 2020.
2. Muhammad Ahsan-ul-Haq, *Statistical Analysis of Haq Distribution: Estimation and Applications*, Pak. J. Statist, vol. 38, no. 4, pp. 473–490, 2022.
3. Ahmad Aijaz, Muzamil Jallal, S. Qurat Ul Ain, and Rajnee Tripathi, *The Hamza distribution with statistical properties and applications*, Asian Journal of Probability and Statistics, vol. 8, no. 1, pp. 28–42, 2020.
4. Christophe Chesneau, Lishamol Tomy, and Jiju Gillariose, *On a sum and difference of two Lindley distributions: theory and applications*, REVSTAT-Statistical Journal, vol. 18, no. 5, pp. 673–695, 2020.
5. Sarra Chouia and Halim Zeghdoudi, *The XLindley distribution: properties and application*, Journal of Statistical Theory and Applications, vol. 20, no. 2, pp. 318–327, 2021.
6. Mohammed Elgarhy, *Garhy distribution with different estimation methods and applications to engineering and medical data*, Revista Internacional de Metodos Numericos para Calculo y Diseno en Ingeniería, vol. 61, 2026.
7. Alexandru Isaic-Maniu and Viorel Gh Voda, *Generalized Burr-Hatke Equation as Generator of a Homographic Failure Rate*, Journal of Applied Quantitative Methods, vol. 3, no. 3, 2008.
8. M. I. Khan and Molay Kumar Ruidas, *The Mgamma Distribution: Statistical Properties and Application to Real-Life Data*, European Journal of Pure and Applied Mathematics, vol. 18, no. 4, pp. 6653, 2025.
9. Dennis V Lindley, *Fiducial distributions and Bayes' theorem*, Journal of the Royal Statistical Society. Series B (Methodological), pp. 102–107, 1958.
10. Dennis V Lindley, *Approximate Bayesian methods*, Trabajos de estadística y de investigación operativa, vol. 31, no. 1, pp. 223–245, 1980.
11. Hamouda Messaadia and Halim Zeghdoudi, *Zeghdoudi distribution and its applications*, International Journal of Computing Science and Mathematics, vol. 9, no. 1, pp. 58–65, 2018.
12. Noureddine Saaidia, Thara Belhamra, and Halim Zeghdoudi, *On ZLindley distribution: statistical properties and applications*, Studies in Engineering and Exact Sciences, vol. 5, no. 1, pp. 3078–3097, 2024.
13. Noureddine Saaidia, Reza Pakyari, and Halim Zeghdoudi, *The Q-Lindley Distribution: Goodness-of-Fit Tests, Modeling, Inference, and Applications*, Journal of Statistical Theory and Applications, vol. 25, no. 1, pp. 6, 2026.
14. Subhradev Sen, Sudhansu S Maiti, and N Chandra, *The xgamma distribution: statistical properties and application*, Journal of Modern Applied Statistical Methods, vol. 15, no. 1, pp. 38, 2016.
15. Rama Shanker, *Akash distribution and its applications*, International Journal of Probability and Statistics, vol. 4, no. 3, pp. 65–75, 2015.
16. Rama Shanker, *Shanker distribution and its applications*, International Journal of Statistics and Applications, vol. 5, no. 6, pp. 338–348, 2015.
17. Rama Shanker, *Aradhana distribution and its applications*, International Journal of Statistics and Applications, vol. 6, no. 1, pp. 23–34, 2016.
18. Rama Shanker, *Sujatha distribution and its applications*, Statistics in Transition. New Series, vol. 17, no. 3, pp. 391–410, 2016.
19. Rama Shanker, *Devya distribution and its applications*, International Journal of Statistics and Applications, vol. 6, no. 4, pp. 189–202, 2016.
20. Rekha Radhakrishnan, D. Venkatesan, and C. B. Prasanth, *Sujit distribution and its applications*, Int. J. Agricult. Stat. Sci. Vol, vol. 21, no. 1, pp. 161–166, 2025.
21. R. Shanker, *Suja Distribution and Its Application*, Journal of Probability and Statistics, vol. 6, no. 2, pp. 11, 2017.
22. G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed., Duxbury Press, Pacific Grove, CA, 2002.
23. E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed., Springer, New York, NY, 1998.