RT&A, No 2 (84)
Volume 20, June 2025

COMPARATIVE ANALYSIS OF BAYESIAN ESTIMATION TECHNIQUES FOR GAMMA DISTRIBUTIONS

Vijay Kumar Lingutla[1], Nagamani Nadiminti[2], R. K. Davala[3]

•

[1,2,3]Department of Mathematics, School of Advanced Sciences, VIT-AP University, Inavolu, Beside AP Secretariat, Amaravati AP-522237, India

[1]vijay.lingutla@gmail.com, [2]maniinadiminti999@gmail.com, [3]davalaravikumar@gmail.com

Abstract

This paper explores the estimation of the common scale parameter for two Gamma populations using three Bayesian methods: Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH). Unlike traditional methods like Maximum Likelihood Estimation (MLE), which can be sensitive to model misspecification, these Bayesian methods are more flexible in handling parameter uncertainty and complex data. ABC bypasses explicit likelihood calculations, HMC improves sampling efficiency through gradient information, and MH offers simplicity and adaptability. A simulation study compares these methods based on bias, mean squared error (MSE), and computational efficiency. Additionally, we apply these methods to real glucose level data from male and female individuals with diabetes, highlighting their strengths and limitations. The results offer practical recommendations for real-world applications involving Gamma distributions, contributing to the advancement of Bayesian estimation techniques in complex parameter estimation problems.

Keywords: Gamma Distribution, Bayesian estimation, Approximate Bayesian Computation, Hamiltonian Monte Carlo, Metropolis-Hastings.

1. Introduction

The gamma distribution is widely used in reliability analysis, survival studies, and queuing theory due to its flexibility in modeling positive continuous data Lawless [1]. A key parameter in these applications is the scale parameter, which determines the distribution's spread. Estimating this parameter is relatively straightforward when the shape parameters of the gamma populations are known. However, when these shape parameters are unknown and may differ across populations, the estimation problem becomes significantly more complex.

Traditional methods, such as Maximum Likelihood Estimation (MLE), are commonly used for parameter estimation in gamma distributions. Although MLE is asymptotically efficient, it is highly sensitive to deviations from model assumptions, such as outliers or misestimated shape parameters Huber [2]. These limitations highlight the need for more flexible and robust estimation techniques, especially in real-world applications.

Computational Bayesian methods provide a promising alternative. This study explores three advanced Bayesian techniques: Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and the Metropolis-Hastings (MH) algorithm. ABC is particularly useful when likelihood functions are intractable, as it relies on summary statistics for inference Beaumont [3]. HMC efficiently samples high-dimensional spaces by leveraging gradient information, leading to faster convergence and lower autocorrelation compared to traditional Markov Chain Monte Carlo

(MCMC) methods Neal [4]. Robert and Casella [5] further examined its efficiency in improving Markov chain mixing. Betancourt [6] provided a detailed explanation of why HMC excels in high-dimensional settings, reinforcing its inclusion in this study. Meanwhile, MH remains widely used for posterior sampling due to its simplicity and flexibility Hastings [7]. Chib and Greenberg [8] later explored its practical applications in Bayesian inference, demonstrating its adaptability across statistical models.

A common assumption in analyzing multiple gamma populations is that their scale parameters are equal, simplifying model formulation and improving statistical efficiency Gelman et al. [9]. Rubin [10] further explored the advantages of this assumption, particularly in Bayesian inference, where hierarchical structures enable borrowing strength across related populations. Ibrahim, Chen, and Sinha [11] provided a comprehensive discussion on Bayesian methods for survival data, highlighting their relevance to this study. Recent advancements in adaptive MCMC techniques have further improved estimation robustness Andrieu et al. [12]. Additionally, Andrieu et al. [13] introduced particle MCMC methods, which integrate sequential Monte Carlo with MCMC strategies to handle complex posterior distributions.

While Bayesian methods have been studied in various contexts, few have directly compared their performance in estimating the common scale parameter of two gamma populations with unknown shape parameters. This research addresses that gap by systematically evaluating ABC, HMC, and MH based on bias, mean squared error (MSE), and computational efficiency. The findings will highlight each method's strengths and limitations, offering practical guidance for their application.

This study builds on the work of Nagamani and Tripathy [14], who estimated the common scale parameter for two gamma populations using simulation-based methods. We extend their approach by incorporating advanced Bayesian techniques for a more comprehensive evaluation. Through a simulation-based study, we contribute to the growing literature on Bayesian estimation Gilks et al. [15]. Spiegelhalter et al. [16] refined these methods, particularly in model checking and predictive inference. Christensen et al. [17] provided insights into Bayesian estimation for various statistical models, including gamma distributions, reinforcing the theoretical foundation of this study. Prior research has also examined Bayesian inference for complex distributions [18], with Liu et al. [19] exploring Bayesian approaches for intricate distributional structures. Kamal et al. [20] further contributed by studying Bayesian estimation for parametric models with unknown distributions, highlighting the need for robust Bayesian methods. Additionally, Sisson, Fan, and Beaumont [21] provided an authoritative discussion on ABC methods, detailing their advantages and limitations, further justifying their use in complex estimation scenarios.

The paper is structured as follows: Section 2 discusses the methodology, covering the gamma distribution model, simulation setup, and parameter estimation methods—ABC, HMC, and MH—along with their assumptions and hyperparameters. Section 3 presents the simulation results, evaluating and comparing the methods based on key performance metrics. Section 4 applies these methods to real-life glucose level data from male and female individuals with diabetes, demonstrating the practical use of Bayesian estimation techniques in analyzing real-world complex data. Finally, Section 5 concludes the paper by summarizing the findings, offering recommendations for choosing the appropriate estimation method, and suggesting directions for future research in Bayesian parameter estimation.

2. Methodology

Consider two independent random samples drawn from Gamma distributions: $X = (x_1, x_2, \ldots, x_m)$ and $Y = (y_1, y_2, \ldots, y_n)$ with sizes m and n, respectively. These samples are assumed to follow Gamma distributions $G(\mu_1, \beta)$ and $G(\mu_2, \beta)$ with a common scale parameter β and different shape parameters μ_1 and μ_2. The probability density functions (PDFs) for these Gamma distributions are:

$$f(x; \mu_1, \beta) = \frac{x^{\mu_1 - 1} e^{-x/\beta}}{\Gamma(\mu_1)\beta^{\mu_1}}, \quad x > 0,\ \mu_1 > 0,\ \beta > 0, \tag{1}$$

$$f(y;\mu_2,\beta)=\frac{y^{\mu_2-1}e^{-y/\beta}}{\Gamma(\mu_2)\beta^{\mu_2}},\quad y>0,\ \mu_2>0,\ \beta>0. \tag{2}$$

The Bayesian framework provides a robust approach to parameter estimation, leveraging prior knowledge to improve inference, particularly in complex models such as Gamma distributions. In this study, we consider Bayesian estimation of the common scale parameter β and the shape parameters μ_1 and μ_2 for two Gamma populations. Let $\pi_1(\beta)$, $\pi_2(\mu_1)$, and $\pi_3(\mu_2)$ represent the prior density functions for β, μ_1, and μ_2, respectively.

The likelihood function of (β,μ_1,μ_2) for the observed data (X,Y) is expressed as:

$$L(\beta,\mu_1,\mu_2|X,Y)=\prod_{i=1}^{m}\frac{1}{\Gamma(\mu_1)\beta^{\mu_1}}x_i^{\mu_1-1}e^{-x_i/\beta}\prod_{j=1}^{n}\frac{1}{\Gamma(\mu_2)\beta^{\mu_2}}y_j^{\mu_2-1}e^{-y_j/\beta},$$

where $X=\{x_1,x_2,\ldots,x_m\}$ and $Y=\{y_1,y_2,\ldots,y_n\}$ are samples from the two Gamma populations. The posterior density function of (β,μ_1,μ_2) given (X,Y) is:

$$f(\beta,\mu_1,\mu_2|X,Y)=\frac{f(\beta,\mu_1,\mu_2,X,Y)}{\int_0^\infty\int_0^\infty\int_0^\infty f(\beta,\mu_1,\mu_2,X,Y)\,d\beta\,d\mu_1\,d\mu_2}.$$

The posterior expectation of any function $g(\beta,\mu_1,\mu_2)$ is:

$$E[g(\beta,\mu_1,\mu_2)|X,Y]=\frac{\int_0^\infty\int_0^\infty\int_0^\infty g(\beta,\mu_1,\mu_2)f(\beta,\mu_1,\mu_2,X,Y)\,d\beta\,d\mu_1\,d\mu_2}{\int_0^\infty\int_0^\infty\int_0^\infty f(\beta,\mu_1,\mu_2,X,Y)\,d\beta\,d\mu_1\,d\mu_2}. \tag{3}$$

The integral ratio in equation (3) is analytically intractable due to the complexity of the posterior distribution. Therefore, numerical techniques such as Hamiltonian Monte Carlo (HMC), Metropolis-Hastings (MH), or Approximate Bayesian Computation (ABC) are employed to approximate the posterior distributions. These methods enable efficient computation of Bayesian estimators for the parameters under different prior distributions and loss functions.

This Bayesian framework provides a comprehensive basis for comparing computational methods in terms of their accuracy, efficiency, and applicability to real-world data involving Gamma populations.

2.1. Approximate Bayesian Computation (ABC)

Approximate Bayesian Computation (ABC) is a powerful Bayesian inference method used when the likelihood function is difficult or impossible to compute directly. Instead of using the likelihood function, ABC uses a set of summary statistics to approximate the posterior distribution [3]. The ABC method is particularly useful when the model is complex and simulating data from the model is easier than evaluating the likelihood function.

ABC Procedure

- Two populations are modeled as Gamma distributions, with the following relationships for the survival times:
$$X\sim\text{Gamma}(\mu_1,\beta),\quad Y\sim\text{Gamma}(\mu_2,\beta).$$
Here, X and Y represent the survival times for the two populations, with β being the common scale parameter, and μ_1 and μ_2 as the shape parameters.

- The prior distributions for the parameters are chosen as:
$$\beta\sim\text{Gamma}(b_1,c_1),\quad \mu_1\sim\text{Gamma}(b_2,c_2),\quad \mu_2\sim\text{Gamma}(b_3,c_3).$$

- Synthetic data are generated by drawing samples from the prior distributions and simulating data based on the Gamma model for both populations.

- The summary statistics (such as the mean and variance) of both the observed data (S_{obs}) and the simulated data (S_{sim}) are computed. A distance metric, typically the Euclidean distance, is then used to quantify the difference:

$$d = \|S_{\text{obs}} - S_{\text{sim}}\|.$$

- A tolerance level ϵ is defined, and parameter sets are accepted if the distance satisfies $d \leq \epsilon$.
- The procedure is repeated iteratively, generating synthetic data and comparing it to the observed data. Accepted parameter sets help build an approximate posterior distribution.
- From the approximate posterior distribution, point estimates such as the posterior mean or median are derived, along with credible intervals for the parameters.

This method provides a flexible and robust framework for Bayesian inference under complex data scenarios by approximating posterior distributions without requiring explicit likelihood functions. It efficiently handles parameter uncertainty and supports reliability analysis for two gamma populations.

2.2. Hamiltonian Monte Carlo (HMC)

Hamiltonian Monte Carlo (HMC) is a Markov Chain Monte Carlo (MCMC) method that utilizes Hamiltonian dynamics to propose new states in the parameter space. This method is known for its efficiency in exploring high-dimensional parameter spaces by using gradient information from the posterior distribution [4]. Unlike traditional random walk-based MCMC algorithms, HMC uses the gradient of the target distribution to guide the chain's movements, leading to faster mixing and reduced autocorrelation between samples. In our study, HMC is applied to sample from the posterior distribution of the common scale parameter, allowing for efficient estimation even in higher-dimensional settings.

HMC Procedure

- Two populations are modeled as Gamma distributions with common scale parameter β and shape parameters μ_1 and μ_2:

$$X \sim \text{Gamma}(\mu_1, \beta), \quad Y \sim \text{Gamma}(\mu_2, \beta).$$

- The prior distributions for the parameters are chosen as:

$$\beta \sim \text{Gamma}(b_1, c_1), \quad \mu_1 \sim \text{Gamma}(b_2, c_2), \quad \mu_2 \sim \text{Gamma}(b_3, c_3).$$

- The Hamiltonian Monte Carlo method involves simulating the joint distribution of parameters, leveraging their gradients to propose new values for the parameters. This is done by introducing auxiliary momentum variables and using Hamiltonian dynamics to simulate the movement of the system.
- The trajectory of the parameters is simulated using the Hamiltonian dynamics, and the positions are updated using the leapfrog integration method to minimize the energy (or maximize the posterior distribution).
- A Metropolis acceptance step is applied to the proposed new position, where the new state is accepted with a probability given by the ratio of the posterior densities at the new and old positions.
- The process is repeated for multiple iterations, and a sequence of samples is generated from the posterior distribution of the parameters.

- From the posterior samples, point estimates such as the posterior mean or median are derived, along with credible intervals for the parameters.

HMC provides an efficient method for exploring complex high-dimensional posterior distributions by leveraging gradient information, which improves convergence and reduces autocorrelation compared to traditional MCMC methods.

2.3. Metropolis-Hastings (MH)

Metropolis-Hastings (MH) is one of the most widely used MCMC algorithms for sampling from posterior distributions. It relies on a proposal distribution to generate candidate samples and accepts or rejects these candidates based on a specified acceptance probability [7]. The MH algorithm is a versatile method that can be adapted to a wide variety of problems. The general procedure for MH involves:

MH Procedure

- Two populations are modeled as Gamma distributions with the following relationships for the survival times:
$$X \sim \text{Gamma}(\mu_1, \beta), \quad Y \sim \text{Gamma}(\mu_2, \beta).$$
Here, X and Y represent the survival times for the two populations, with β being the common scale parameter, and μ_1 and μ_2 as the shape parameters.
- The prior distributions for the parameters are chosen as:
$$\beta \sim \text{Gamma}(b_1, c_1), \quad \mu_1 \sim \text{Gamma}(b_2, c_2), \quad \mu_2 \sim \text{Gamma}(b_3, c_3).$$
- The Metropolis-Hastings algorithm starts with an initial value for the parameters and proposes new values based on a proposal distribution. The proposal distribution is typically chosen to be symmetric, such as a normal distribution.
- The ratio of the posterior densities at the new and old parameter values is computed, and the new parameters are accepted with a probability given by the Metropolis criterion:
$$\mu = \min\left(1, \frac{p(\theta_{\text{new}}|X, Y)}{p(\theta_{\text{old}}|X, Y)}\right).$$
If the new parameters are accepted, they replace the old ones; otherwise, the old parameters are retained.
- This process is repeated iteratively for a set number of steps or until convergence is reached, generating a sequence of samples from the posterior distribution of the parameters.
- From the posterior samples, point estimates such as the posterior mean or median are derived, along with credible intervals for the parameters.

The Metropolis-Hastings algorithm is widely used for sampling from complex posterior distributions due to its simplicity and flexibility. It is particularly effective in situations where the likelihood is difficult to compute directly.

3. Simulation Study

To assess the performance of the proposed methodology for estimating the parameters of two Gamma populations, we conducted an extensive simulation study. The study aimed to evaluate the accuracy, robustness, and efficiency of the Metropolis-Hastings (MH) algorithm in recovering the true parameter values across various experimental conditions, while also including the

Hamiltonian Monte Carlo (HMC) and Approximate Bayesian Computation (ABC) methods for comparison.

In this section, we present experimental results evaluating the performance of different estimation methods across varying sample sizes. We estimate the common scale parameter β and the potentially distinct shape parameters μ_1 and μ_2 using the ABC, HMC, and MH algorithms.

Two sets of parameter values are considered:

- Set 1: $(\mu_1, \mu_2, \beta) = (1, 2, 1.5)$
- Set 2: $(\mu_1, \mu_2, \beta) = (1, 2, 2.5)$

We evaluate the performance of the estimators using two metrics:

- Average Value (Estimate)
- Mean Square Error (MSE)

To facilitate comparison, we generate 10,000 random samples from the two Gamma populations across various sample sizes and parameter combinations. Specific hyperparameters ($b_1 = b_2 = b_3 = 1$ and $c_1 = c_2 = c_3 = 2$) are employed to calculate biases and MSE for all estimators, with results summarized in the tables below.

Table 1 & 2 presents the Estimate and MSE of MLE and Bayesian estimators under a SEL (squared error loss) function. The first column indicates sample sizes and the second column details the parameters μ_1, μ_2, and β. Columns 3 and 4 show the Estimate and MSE of the ABC estimates, while subsequent columns reflect the estimates of HMC and MH algorithm.

Tables 3 & 4 display the 95% HPD intervals derived from HMC and MH for the parameters. The first two columns indicate sample sizes and parameters, the third column shows asymptotic confidence intervals calculated using the HMC, and the fourth column presents the HPD intervals estimated through MH algorithm.

The observations from our simulation study yield several insights into the performance of the estimators:

- Increasing sample sizes consistently reduce MSE across all estimators, underscoring the importance of data size for estimation accuracy.
- ABC, HMC and MH converge to stable estimates as sample sizes increase, ensuring consistency and reliability.
- ABC, HMC and MH algorithms performed well, but HMC consistently produced slightly more precise estimates with narrower posterior distributions. ABC & MH occasionally exhibited higher variability, especially for smaller sample sizes.
- The accuracy and precision of parameter estimates improved significantly as sample sizes increased. Smaller sample sizes resulted in wider 95% Highest Posterior Density (HPD) intervals, indicating greater uncertainty, while larger sample sizes yielded narrower intervals.
- Compared to traditional methods like Gibbs sampling and MLE, our study shows that HMC offers smoother sampling and faster convergence, while ABC demonstrated strong robustness, especially in noisy data scenarios. The Metropolis-Hastings algorithm performed competitively with Gibbs sampling for larger sample sizes but showed slightly higher MSE in smaller samples.

4. Real Data analysis

To validate our model, we analyze glucose levels among male and female individuals with diabetes who have never smoked. The objective is to identify potential gender-based differences in blood sugar regulation. The dataset, sourced from [22] and openly available, serves as the

Table 1: *For various sample sizes, we compare biases and mean square errors of multiple estimators under squared error loss for $\theta = (\mu_1, \mu_2, \beta)$.*

		ABC		HMC		MH	
(m, n)	Observed	Avg	Error	Avg	Error	Avg	Error
(10,10)	μ_1=1	1.7905	0.6628	1.1535	0.1273	1.3618	0.2541
	μ_2=2	1.6857	0.5596	1.2527	0.6027	1.8651	0.4216
	β=1.5	1.2389	0.3975	1.6196	0.0678	1.4224	0.1889
(20,10)	μ_1=1	1.6831	0.5541	0.8695	0.04	0.9881	0.2450
	μ_2=2	2.6094	0.4185	2.0003	0.1467	2.1232	0.3532
	β=1.5	1.04	0.2177	1.5748	0.0362	1.4375	0.1432
(10,20)	μ_1=1	1.7955	0.5234	1.1233	0.0563	1.4720	0.2424
	μ_2=2	2.5821	0.6474	1.9726	0.1209	2.2212	0.2847
	β=1.5	1.038	0.2586	1.4731	0.0301	1.2743	0.1611
(30,30)	μ_1=1	1.2239	0.1997	0.9195	0.0218	0.9970	0.0301
	μ_2=2	1.7548	0.2896	1.71	0.1552	1.7605	0.1887
	β=1.5	1.4231	0.1077	1.4259	0.0295	1.3532	0.0941
(50,50)	μ_1=1	1.1262	0.1623	0.9873	0.0201	0.9643	0.0191
	μ_2=2	1.9885	0.2107	2.0923	0.0886	2.3846	0.1020
	β=1.5	1.4474	0.0963	1.4652	0.0198	1.3364	0.0718

Table 2: *For various sample sizes, we compare biases and mean square errors of multiple estimators under squared error loss for $\theta = (\mu_1, \mu_2, \beta)$.*

		ABC		HMC		MH	
(m, n)	Observed	Avg	Error	Avg	Error	Avg	Error
(10,10)	μ_1=1	1.7383	0.9155	1.291	0.2019	1.4806	0.3656
	μ_2=2	1.4756	0.4434	1.4266	0.3709	2.0112	0.4304
	β=2.5	2.5085	0.714	2.1637	0.266	2.1064	0.4615
(20,10)	μ_1=1	1.0409	0.0017	0.9642	0.1341	1.0525	0.2491
	μ_2=2	2.5289	0.2797	2.238	0.2216	2.2882	0.3655
	β=2.5	2.3103	0.036	2.2247	0.1688	2.1455	0.3645
(10,20)	μ_1=1	1.5312	0.289	1.2376	0.1073	1.4720	0.2424
	μ_2=2	1.5578	0.2059	2.2075	0.1531	2.2212	0.2847
	β=2.5	2.2135	0.0906	2.1395	0.1662	1.2743	0.3611
(30,30)	μ_1=1	0.801	0.1975	0.9858	0.0195	1.0349	0.0307
	μ_2=2	1.4272	0.7557	1.8763	0.1037	1.8547	0.2535
	β=2.5	2.6424	0.6259	2.1338	0.1073	2.1037	0.2996
(50,50)	μ_1=1	1.7408	0.5488	0.9148	0.0155	0.9937	0.0194
	μ_2=2	2.2353	0.0554	2.1885	0.0943	2.4751	0.1760
	β=2.5	1.7755	0.5248	2.3097	0.0947	2.1290	0.2338

Table 3: *95% HPD intervals using HMC and MH algorithams for the parameters (μ_1, μ_2, β)=(1, 1.5, 2) at various sample sizes*

(m,n)	Observed	HMC	MH
(10,10)	μ_1=1	[0.5781, 1.6895]	[0.7294, 2.0582]
	μ_2=2	[0.8307, 2.6554]	[0.7122, 2.8704]
	β=1.5	[1.1448, 2.087]	[0.857, 2.3381]
(20,10)	μ_1=1	[0.6122, 1.1922]	[0.6248, 1.4402]
	μ_2=2	[1.2255, 2.4953]	[0.8936, 3.4056]
	β=1.5	[1.2714, 1.9175]	[0.871, 2.2512]
(10,20)	μ_1=1	[0.7751, 1.5938]	[0.7907, 2.127]
	μ_2=2	[1.2922, 2.4633]	[1.1446, 3.1959]
	β=1.5	[1.1423, 1.8007]	[0.7812, 2.0972]
(30,30)	μ_1=1	[0.6976, 1.1776]	[0.6866, 1.3578]
	μ_2=2	[1.2081, 2.108]	[1.0913, 2.3701]
	β=1.5	[1.1481, 1.7068]	[0.9123, 1.9463]
(50,50)	μ_1=1	[0.7833, 1.2777]	[0.7247, 1.2439]
	μ_2=2	[1.5996, 2.5279]	[1.7072, 3.102]
	β=1.5	[1.1658 1.8268]	[0.9585, 1.781]

Table 4: *95% HPD intervals using HMC and MH algorithams for the parameters (μ_1, μ_2, β)=(1, 1.5, 2) at various sample sizes*

(m,n)	Observed	HMC	MH
(10,10)	μ_1=1	[0.6585, 1.8312]	[0.8111, 2.1997]
	μ_2=2	[0.9728, 2.7928]	[0.8452, 3.2114]
	β=2.5	[1.7432, 2.627]	[1.3433, 3.3201]
(20,10)	μ_1=1	[0.6758, 1.3285]	[0.6798, 1.481]
	μ_2=2	[1.4236, 2.7487]	[1.0481, 3.6334]
	β=2.5	[1.8816, 2.6857]	[1.4003, 3.1843]
(10,20)	μ_1=1	[0.8573, 1.7827]	[0.8826, 2.2122]
	μ_2=2	[1.5443, 2.6648]	[1.2453, 3.3589]
	β=2.5	[1.7822 2.9561]	[1.2655, 2.965]
(30,30)	μ_1=1	[0.735, 1.2751]	[0.7218, 1.3871]
	μ_2=2	[1.4345, 2.2498]	[1.22, 2.5561]
	β=2.5	[1.8064, 2.8706]	[1.4851, 2.8742]
(50,50)	μ_1=1	[0.7341, 1.0959]	[0.7384, 1.2774]
	μ_2=2	[1.7131, 2.6098]	[1.7623, 3.1623]
	β=2.5	[1.833, 2.7369]	[1.9046, 2.8637]

foundation for this study. The primary goal is to assess glucose level distributions across genders and evaluate whether they follow a Gamma distribution. Both visual and statistical techniques are employed to examine the goodness-of-fit of the Gamma model, ensuring a robust assessment of its applicability to the data.

A collection of 50 sample observations related to glucose levels among male and female individuals are as follows:
Glucose Levels Among Male: 124.3, 180.5, 72.9, 90.8, 196.2, 75.1, 131.5, 128.5, 117.8, 75, 118, 124, 184.7, 152.3, 149.9, 162.8, 70.7, 101.7, 99.4, 125.7, 120.3, 73.6, 161.1, 92.1, 119.7, 163.5, 166.8, 187.9, 103.1, 198.9, 89.3, 160.7, 169, 195.9, 91.2, 194.1, 109.4, 84.8, 122.9, 101, 162.7, 138.3, 106.9, 152.8, 161.2, 88.9, 80.7, 163.7, 104.6, 177.4.

Glucose Levels Among Females: 157.2, 125.9, 127.7, 147.9, 145.8, 190.2, 151.4, 105.7, 85.8, 179.4, 154.1, 86.8, 186.2, 94.6, 94.8, 170, 95.8, 108.1, 139.5, 180.4, 165.9, 165.7, 81.9, 173.1, 92.4, 118.9, 165.9, 137.1, 92.3, 75.8, 181.1, 70.5, 163.4, 156.3, 85.8, 81.6, 177.8, 104.1, 79.8, 144, 174.8, 94.3, 90.8, 72.8, 123.8, 190.8, 140.2, 134.9, 71, 100.3.

In this study, we employ both visual and statistical techniques to assess the goodness-of-fit of the Gamma distribution for glucose levels across genders.

The Q-Q plot shown in Figure 1 demonstrates that the Gamma distribution provides a strong fit for the observed glucose levels in both male and female groups. The majority of data points align well with the reference line, particularly in the middle range, indicating that the Gamma model effectively captures the central trend of glucose levels. While some deviations are observed in the upper tail, suggesting slight underestimation of higher glucose values, the overall fit remains robust. This suggests that the Gamma distribution is a suitable choice for modeling glucose levels, as it accurately represents the underlying data structure while capturing key distributional characteristics.

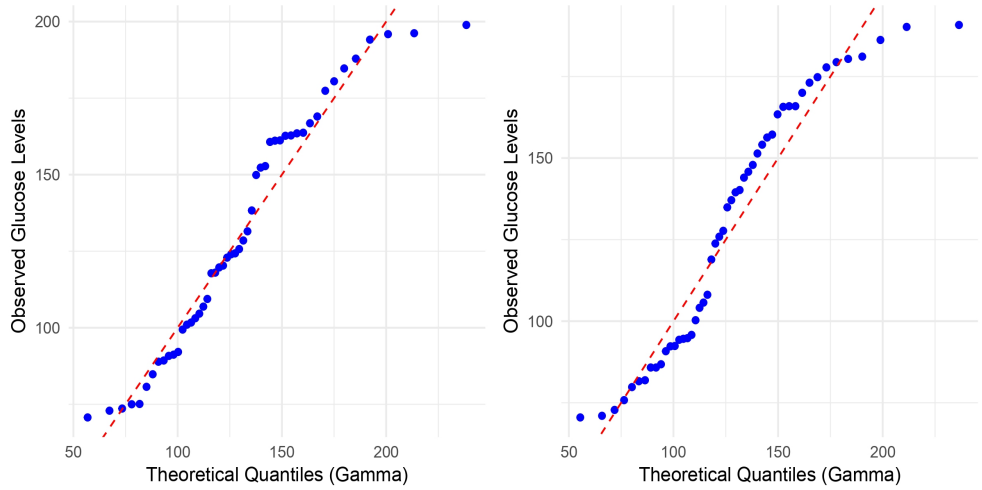

Figure 1: *Q-Q Plot curve compared with the theoretical Gamma model.*

To statistically test whether the scale parameters of the two populations are equal, we performed the Wald test. The null hypothesis for the Wald test is that the scale parameters of the two populations are equal ($H_0 : \beta_1 = \beta_2 = \beta$). The test results are presented in Table 5, with the computed Wald statistic of 0.000832 and a corresponding p-value of 0.97698. Since the p-value is greater than the standard significance level ($\alpha = 0.05$), we fail to reject the null hypothesis. This implies that there is insufficient evidence to conclude that the scale parameters are significantly different between the two populations, suggesting that the scale parameters may be equal.

The results of parameter estimation for the Weibull distribution, using Bayesian estimation methods, are summarized in Tables 6 and 7 . Table 6 presents the point estimates for the scale

Table 5: *Wald Test for Equality of Shape Parameters*

Data	$\hat{\beta}$	**V(β)**	**Statistic**	**P-value**
Set 1	11.82508	2.378837	0.000832	0.05726
Set 2	11.92254	2.398886		

Table 6: *Maximum Likelihood estimators and Bayes estimators of the combined model*

Method	μ_1	μ_2	β
ABC	9.1273	9.4199	12.3821
HMC	9.4844	9.9056	12.7042
MH	9.3260	9.1939	12.8135

parameter (β) and the two shape parameters (μ_1 and μ_2) obtained through various Bayesian estimation techniques: Approximate Bayesian Computation(ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH).

The point estimates for the Weibull distribution parameters, including the shape parameters (μ_1, μ_2) and the scale parameter (β), were obtained using three different Bayesian estimation methods: Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH). Among these, HMC provided the most accurate and consistent estimates, with values closest to each other for all parameters. ABC produced estimates that were slightly lower than those from HMC, especially for the shape parameters, while MH showed slightly more variability in its estimates. Overall, HMC yielded the best point estimates, indicating its efficiency in parameter estimation compared to the other methods.

The 95% Highest Posterior Density (HPD) intervals for the parameters revealed the precision of each Bayesian method. HMC produced narrower and more precise HPD intervals, indicating higher confidence in the parameter estimates, with values tightly clustered around the point estimates. In contrast, MH generated wider HPD intervals, reflecting greater uncertainty in the parameter estimates. ABC's HPD intervals were not provided here, but based on the point estimates, it is expected that the intervals would be less precise than those from HMC. The narrow HPD intervals from HMC suggest it as the most reliable method for quantifying uncertainty in the parameters.

5. Conclusion

This study compared three Bayesian parameter estimation methods—Approximate Bayesian Computation (ABC), Hamiltonian Monte Carlo (HMC), and Metropolis-Hastings (MH)—for estimating the parameters of two Gamma populations in the context of analyzing glucose levels in male and female individuals with diabetes who have never smoked. The simulation results revealed that increasing the sample size consistently improved the accuracy of all methods. HMC generally provided the most precise estimates, with narrower posterior distributions and lower Mean Squared Error (MSE), particularly for larger sample sizes. ABC, while robust and effective in

Table 7: *95% HPD intervals using HMC and MH Algorithms for the parameters* (μ_1, μ_2, β)

Method	μ_1	μ_2	β
HMC	[8.6691, 10.0549]	[9.2609, 10.8336]	[12.2959, 13.0811]
MH	[4.9786, 11.4956]	[5.7222, 11.7434]	[10.3235, 15.5177]

handling noisy data, showed higher variability with smaller samples. MH produced results similar to Gibbs sampling, with competitive outcomes for larger sample sizes but slightly higher MSE for smaller ones. These findings suggest that HMC was the most efficient method, offering faster convergence and better mixing, making it particularly suitable for high-dimensional estimation tasks, such as analyzing complex glucose level data. On the other hand, ABC demonstrated robustness in more complex models, especially when the data showed significant noise. Overall, while HMC emerged as the preferred choice for accurate and efficient estimation, ABC and MH each showcased their strengths in particular contexts, emphasizing the trade-offs between precision and computational efficiency in real-world applications like diabetes research.

References

[1] Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. Wiley-Interscience.
[2] Huber, P. J. (2009). *Robust Statistics*. Wiley.
[3] Beaumont, M., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, **162**(4), 2025–2035.
[4] Neal, R. M. (2011). MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo* (pp. 113–162). CRC Press.
[5] Robert, C. P., Casella, G., & Chopin, N. (2013). *Monte Carlo Statistical Methods* (3rd ed.). Springer.
[6] Betancourt, M. (2017). A Conceptual Introduction to Hamiltonian Monte Carlo. *arXiv preprint arXiv:1701.02434*.
[7] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**(1), 97–109.
[8] Chib, S. (1995). Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association*, **90**(432), 1313–1321.
[9] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). CRC Press.
[10] Rubin, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics*, **12**(4), 1151–1172.
[11] Ibrahim, J. G., Chen, M. H., & Sinha, D. (2001). Bayesian Survival Analysis. *Springer*.
[12] Andrieu, C., Thoms, J., & Doucet, A. (2010). Adaptive MCMC: Theory and applications. *Springer*.
[13] Andrieu, C., Doucet, A., & Holenstein, R. (2008). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **70**(2), 269–292.
[14] Nagamani, N., & Tripathy, M. R. (2017). Estimating common scale parameter of two gamma populations: A simulation study. *American Journal of Mathematical and Management Sciences*, **36**(4), 346–362.
[15] Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1995). *Markov Chain Monte Carlo in Practice*. CRC Press.
[16] Spiegelhalter, D. J., Abrams, K. R., & Myles, J. P. (2002). *Bayesian Methods in Health Technology Assessment*. Wiley.
[17] Christensen, R., Johnson, W., Branscum, A., & Hanson, T. (2011). Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians. *CRC Press*.
[18] Bernardo, J. M. (2000). *Bayesian Theory*. Wiley.
[19] Liu, F., Li, H., & Wang, Y. (2023). Efficient parameter estimation for hierarchical models using ABC and Hamiltonian Monte Carlo. *Bayesian Analysis*, **18**(1), 1–21.
[20] Kamal, M. I., Singh, G., & Chatterjee, S. (2022). Bayesian analysis of hierarchical models using approximate Bayesian computation. *Statistical Modelling*, **22**(3), 221–246.
[21] Sisson, S. A., Fan, Y., & Beaumont, M. (2018). Handbook of Approximate Bayesian Computation. *CRC Press*.
[22] Ankit. (2025). Diabetes Prediction in India [Data set]. Kaggle. DOI: 10.34740/KAGGLE/DSV/10594461.