STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–11.
Published online in International Academic Press (www.IAPress.org)

A Note on Prior Selection in Bayesian Estimation

Muhammad M. Seliem [1,*], Amr R. Kamel [1], Ibrahim M. Taha [1], Mona M. Abo Elnasr [2]

[1] *Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, Egypt*
[2] *Department of Applied Statistics, Faculty of Commerce, Mansoura University, Mansoura, Egypt*

Abstract The parameter according to the Bayesian approach is handled as a variable that is random and with a probability distribution, rather than an unknown and fixed number. Statistical inference is dependent on the posterior distribution of the parameter rather than only the likelihood function. Choosing the prior distribution is a fundamental step in determining the posterior distribution, and it can be done objectively or subjectively. When the subjective technique is utilized, the prior distribution reflects the prior information the researcher had before coming into contact with the data. When utilizing the objective method, the prior distribution can be chosen in such a way that it has the least influence on the prior distribution. In this paper, a new way of selecting the prior distribution in Bayesian analysis was proposed. According to this strategy, for a given distribution, the prior distribution should be comparable to or the same as the data distributed. The performance of this method was compared to other estimation methods and found to have significantly better performance compared to other estimators. This result was confirmed though a Monte Carlo simulation experiments, with some selected performance criteria mainly, the root mean squared error (RMSE) and the Bias.

Keywords Exponential Distribution, Gamma Distribution, MLE, Prior and Posterior distributions, RMSE

Mathematics Subject Classification: 62E10, 60K10, 60N05

DOI: 10.19139/soic-2310-5070-1752

1. INTRODUCTION

Bayesian estimation is currently a widespread technique in estimation, and the resulting Bayes estimators have promising theoretical features Kaplan (2023). The Bayesian approach has a significant advantage over the traditional frequentist approach in that the posterior distribution is very easy to derive, and then credible intervals, which are the Bayesian analogues of the classical confidence interval, can be easily obtained either theoretically or through Monte Carlo methods Tanner (1993). Based on Bayes' theorem, Modarres (2006) defines Bayesian estimation as the fundamental method for combining prior judgment with experimental data. Modarres (2006) offers the Bayes' theorem as a result of the conditional probability concept. For discrete variables, the generalized form of Bayes' theorem is

$$P(A_{\mathrm{j}}|B) = \frac{P(A_{\mathrm{j}})P(B|A_{\mathrm{j}})}{\sum_i^n P(A_{\mathrm{i}})P(B|A_{\mathrm{i}})} \tag{1}$$

The prior probability $P(A_{\mathrm{j}})$, is represented on the right side of Bayes' equation, and the remaining element, i.e. the relative likelihood, is based on evidentiary observations. The posterior probability of event A_{j} given event B is $P(A_{\mathrm{j}}|B)$, which is the updated probability of event A_{j}. According to the aforementioned equation, probability data may be updated by combining the prior probability (from previous information) and the relative likelihood.

*Correspondence to: Muhammad M. Seliem (Email: Seliem.m.m@hims.edu.eg). Department of Business Information System, Higher Institute of Management Sciences, Ministry of Higher Education, New Cairo, Egypt

ISSN 2310-5070 (online) ISSN 2311-004X (print)

The rest of the paper is organized as follows. In the next **Section**, introduces the estimation methods. While in **Section 3**, discusses the proposed methodology. In **Section 4**, we illustrate the Monte Carlo simulation study and concluding remarks are included in Section

2. Estimation Methods

Estimation techniques are essential in Bayesian analysis since they enable researchers to deduce model parameters from observable data. Two often used estimating methods are maximum likelihood (ML) estimation and the method of moments (MOM). In recent times, Bayesian analysis has become a potent framework for estimating parameters, providing several benefits compared to old frequentist approaches Held and Bové (2020). ML estimation is to identify the parameter values that optimize the probability of witnessing the provided data. This strategy is often used because of its straightforwardness and favorable statistical characteristics, such as reliability and optimal performance in the long run Gould et al. (2006). The method of moments involves aligning the sample moments of the data with the appropriate theoretical moments of the model, and then determining the parameter values Hall (2003). Although ML and MOM methods have been widely used for parameter estimate for many years, Bayesian analysis has recently attracted considerable interest. Bayesian approaches provide a logical structure for integrating previous knowledge about the parameters, which may be particularly valuable when there is a scarcity of data or when meaningful prior distributions are accessible Zyphur and Oswald (2015). Bayesian analysis utilizes the combination of prior knowledge and the likelihood function to generate a posterior distribution, which reflects the revised views on the parameter values based on the observed data. Bayesian estimation offers many benefits, such as the capability to measure uncertainty using credible intervals, the adaptability to deal with intricate models and hierarchical structures, and the possibility to enhance estimate accuracy with the use of informative priors van de Schoot et al. (2021). However, Bayesian approaches also need the definition of proper prior distributions, which may be complex and may impact the subsequent judgments Gelman et al. (2017). This section will offer an overview of different estimating methodologies, examining their basic concepts, and strengths, as well as emphasizing current advancements and applications within the framework of Bayesian analysis.

2.1. Method of Moments

Consider a population pdf, $f(x;\theta_1,...,\theta_n)$, depending on one or more parameters. The moments about the origin,μ_j' were defined. Generally, these will depend on the parameters, say

$$\mu_j'(\theta_1,...,\theta_n) = E(X^j), \quad j = 1,2,...,k. \tag{2}$$

It is possible to define estimators of these distribution moments. If $X_1,...,X_n$ is a random sample from $f(x;\theta_1,...,\theta_k)$, the first k sample moments are given by

$$m_j' = \frac{\sum_{i=1}^n X_i^j}{n}, \quad j = 1,2,...,k.$$

More generally, the method of moments principle is to choose as estimators of the parameters $\theta_1,...,\theta_k$ the values $\hat{\theta}_1,...,\hat{\theta}_k$ that render the population moments equal to the sample moments. In other words, $\hat{\theta}_1,...,\hat{\theta}_k$ are solutions of the equations

$$m_j' = \mu_j'(\theta_1,...,\theta_k), \quad j = 1,2,...,k. \tag{3}$$

The characteristics of MOM method can be summarized as; (a) Matches the sample moments (e.g., mean, variance, skewness) of the data with the corresponding theoretical moments of the model, (b) Solves the resultant system of equations to acquire the parameter estimations, (c) Does not need describing the whole probability distribution of the data, (d) Tends to be less efficient than ML estimation, particularly for small samples, (e) Can be prone to bias if the sample moments do not precisely represent the genuine population moments, (f) Relatively straightforward to implement, particularly for linear models, (g) Useful when the likelihood function is difficult to describe or optimize, (h) Provides an easy way to estimating, but may not completely exploit all the information in the data Hall (2003).

2.2. *Method of Maximum Likelihood*

Let $L(\theta) = f(x;\theta), \theta \in \Omega$ be the joint pdf of $X_1, ..., X_n$. For a given set of observations $x_1, ..., x_n$ a value $\hat{\theta}$ in Ω at which $L(\theta)$ is called a maximum likelihood estimate (MLE) of θ. That is $\hat{\theta}$ is a value of θ that satisfies

$$f(x_1, ..., x_n; \hat{\theta}) = \max_{\theta \in \Theta} f(x_1, ..., x_n; \theta) \tag{4}$$

The characteristics of ML method can be summarized in the following points; (a) Seeks to identify the parameter values that optimize the chance of witnessing the provided data, (b) Relies on the likelihood function, which indicates the probability of the data given the model parameters, (c) Maximizes the likelihood function to acquire the parameter estimations, (d) Produces parameter estimates that are consistent, efficient, and asymptotically normal under specific circumstances, (e) Does not need knowledge of the underlying parameter distributions, (f) Provides a framework for hypothesis testing and calculating confidence intervals, and (g) Can be computationally demanding for complex models Gould et al. (2006).

2.3. *Bayesian Estimation*

This approach takes into account any prior knowledge of the experiment that the statistician has and it is one application of a principle of statistical inference that may be called Bayesian statistics.

2.3.1. ***Overview:*** In some kinds of problems, it is reasonable to assume that the parameter varies for different cases, and it may be proper to treat θ as a random variable In other cases, $\pi(\theta)$ may reflect prior information or belief as to what the true value of the parameter may be. In either case, introduction of the pdf $\pi(\theta)$, which usually is called a prior density for the parameter θ, constitutes an additional assumption that may be helpful or harmful depending on its correctness. In any event, averaging the risk relative to a pdf $\pi(\theta)$ is a procedure that provides a possible way to discriminate between two estimators when neither of their risk functions is uniformly smaller than the other for all θ. Posterior distribution is a conditional density function of θ if known as the value of T observation, on the Bayesian method, inferences are based on the posterior distribution Box and Tiao (2011). So the posterior distribution can be defined as

$$\pi_{\theta|x}(\theta) \propto f(x_1, \ldots, x_n; \theta) \times \pi(\theta) \tag{5}$$

The symbol of ($\propto$) states that the posterior distribution is proportional or comparable to the likelihood function if multiplied by the prior distribution. Posterior Distribution The conditional density of θ given the sample observations $x = x_1,\ ..., x_n$ ís called the posterior density or posterior pdf, and is given by Hasan and Baizid (2017)

$$\pi_{\theta|x}(\theta) = \frac{f(x_1, ..., x_n|\theta) \times \pi(\theta)}{\int_{R(\theta)} f(x_1, ..., x_n|\theta) \times \pi(\theta)\, d(\theta)} \tag{6}$$

2.3.2. ***Types of Loss Function:*** In Bayesian estimation, the aim is to discover the parameter values that minimize the anticipated loss, which is the weighted average of the losses incurred by various parameter estimations. The loss function reflects the ”cost” or ”penalty” associated with the difference between the real parameter values and the estimated values. The choice of the loss function is critical, since it determines the properties of the resulting Bayesian estimator Rosasco et al. (2004). Some typical loss functions used in Bayesian estimation include: Squared error loss function (SELF), Quadratic loss function (QLF) and Absolute value error loss function (AVELF). Loss Function If $\hat{\theta}$ is an estimator of $\tau(\theta)$, then a loss function is any real-valued function, $L(\hat{\theta};\theta)$, such that

$$L(\hat{\theta};\theta) \geq 0, \text{ for every t}$$

and

$$L(t;\theta) = 0, \text{ when t = } \tau(\theta)$$

- SELF

Where $\hat{\theta}$ is the Bayesian SELF estimator for the θ parameter. The Bayesian estimate of θ on the any distribution is achieved by minimizing the expectations of the loss function obtained as follows Triana and Purwadi (2019).

$$L(\hat{\theta};\theta) = [\hat{\theta} - \theta]^2 \tag{7}$$

In the case of SELF, the Bayes estimate is the mean of the posterior Distribution then,

$$\hat{\theta} = E(\theta)$$

- QFL

Where $\hat{\theta}$ is the Bayesian SELF estimator for the θ parameter. The Bayesian estimate of θ on the any distribution is achieved by minimizing the expectations of the loss function obtained as follows Triana and Purwadi (2019).

$$L(\hat{\theta};\theta) = [\frac{\hat{\theta} - \theta}{\theta}]^2$$

Under this loss function the Bayes estimator of θ is obtained by solving the equation

$$\frac{\partial E[L(\hat{\theta};\theta)]}{\partial\theta} = 0$$

- AVELF

$$L(\hat{\theta};\theta) = |\hat{\theta} - \theta|$$

Other loss functions include the entropy loss, Linex loss, Generalized entropy loss, and Weighted squared error loss can be found in (Atanasiu, 2008; Mahdavi et al., 2022; Faladiba and Ahdika, 2024; Abd and Rasheed, 2024; GENCER and SARAÇOĞLU, 2024; Maswadah and Mohamed, 2024; Saini, 2024; Arqand et al., 2024)

3. Proposed Methodology

The idea of this paper in choosing prior distribution is to assume it came from the data distribution, given in mind the boundaries of the integration or summation. Utilizing this method would grantee a better performance for the Bayesian estimator. We have developed a general algorithm that can be applied to a wide range of distributions, providing a practical and flexible tool for Bayesian analysis. The steps involved in our algorithm are as follows:

1. Consider the unknown parameter θ as a random variable.
2. Use a probability distribution (prior) to describe the uncertainty about the unknown parameter.
3. Update the parameter distribution using the Bayes theorem:
4. $\pi(\theta|x_1,\ldots,x_n) \propto f(x_1,\ldots,x_n;\theta) \times \pi(\theta)$
5. The Bayes estimator of y is set to be the expected value of the posterior distribution $\pi(\theta|x_1,\ldots,x_n)$ under quadratic loss function.

Our proposed method assumes that the prior distribution arises from the data distribution. This approach offers several advantages:

- **Data-Driven Approach:** By basing the prior distribution on the data itself, our method avoids relying on subjective or arbitrary prior specifications. This is particularly beneficial in scenarios where prior information is limited or uncertain.

- **Flexibility:** Our method allows for a more flexible and adaptive Bayesian analysis, as the prior distribution can be tailored to the specific characteristics of the data. This adaptability is crucial in handling complex and diverse datasets.

To demonstrate the effectiveness of our proposed Methodology, we applied it to the exponential and gamma distributions. A continuous random variable X is said to have one parameter exponential distribution with parameter θ if its probability density function (pdf) is given by Ijaz et al. (2019):

$$f(x_1,...,x_n;\theta)=\theta e^{-\theta x};\quad \theta>0 \text{ and } 0\leq x\leq\infty$$

For Bayesian estimation, we need to specify a prior distribution for the parameter. Consider a gamma prior for the parameter θ having density function Hasan and Baizid (2017):

$$\pi_1(\theta)=\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta};\quad \alpha,\beta \text{ and } \theta>0$$

And in this paper we proposed exponential prior for the parameter θ having density function

$$\pi_2(\theta)=\lambda e^{-\lambda\theta};\quad \lambda>0 \text{ and } 0\leq\theta\leq\infty$$

- **Bayesian Gamma (B.G) Prior:** Then the posterior density function of θ for the given random sample X is given by Kundu and Gupta (2008):

$$\begin{aligned}\pi_{1|x}(\theta)&\propto L(x_1,...,x_n;\theta)\times\pi_1(\theta)\\&\propto\theta^n e^{-\theta\sum_{i=1}^n x_i}\cdot\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}\\&\propto\theta^{\alpha+n-1}e^{-\theta(\beta+\sum_{i=1}^n x_i)}\\&=\frac{1}{K}\theta^{\alpha+n-1}e^{-\theta(\beta+\sum_{i=1}^n x_i)}\end{aligned}$$

where

$$\begin{aligned}K&=\int_0^\infty\theta^{\alpha+n-1}e^{-\theta(\beta+\sum_{i=1}^n x_i)}d(\theta)\\&=\frac{\Gamma(\alpha+n)}{(\beta+\sum_{i=1}^n x_i)^{\alpha+n}}\end{aligned}$$

where $\widetilde{\beta}=\beta+\sum_{i=1}^n x_i$ and $\widetilde{\alpha}=\alpha+n$, then $\pi_{1|x}(\theta)\sim Gamma(\widetilde{\alpha},\widetilde{\beta})$) as follows:

$$\pi_{1|x}(\theta)=\frac{(\tilde{\beta})^{\tilde{\alpha}}}{\Gamma(\tilde{\alpha})}\theta^{\tilde{\alpha}-1}e^{-\tilde{\beta}\theta}$$

The Bayes estimate

$$\hat{\theta}_{B.G}=\frac{\tilde{\alpha}}{\tilde{\beta}}$$

- **Bayesian Exponential (B.E) Prior:** Then the posterior density function of $\theta$for the given random sample X is given by

$$
\begin{aligned}
\pi_{2|x}(\theta) &\propto L(x_1,...,x_n;\theta) \times \pi_2(\theta) \\
&\propto \theta^n e^{-\theta \sum_{i=1}^n x_i} \cdot \lambda e^{-\lambda\theta} \\
&\propto \theta^n e^{-\theta(\lambda+\sum_{i=1}^n x_i)} \\
&= \frac{1}{K}\theta^n e^{-\theta(\lambda+\sum_{i=1}^n x_i)}
\end{aligned}
$$

where

$$
\begin{aligned}
K &= \int_0^\infty \theta^n e^{-\theta(\lambda+\sum_{i=1}^n x_i)} d\theta \\
&= \frac{\Gamma(n+1)}{(\lambda+\sum_{i=1}^n x_i)^{n+1}}
\end{aligned}
$$

where $\breve{\beta} = \beta + \sum_{i=1}^n x_i$, and $\breve{\alpha} = \alpha + n$, then $\pi_{2|x}(\theta) \sim Gamma(\breve{\alpha}, \breve{\beta})$) as follows:

$$\pi_{2|x}(\theta) = \frac{(\breve{\beta})^{\breve{\alpha}}}{\Gamma(\breve{\alpha})}\theta^{\breve{\alpha}-1}e^{-\breve{\beta}\theta}$$

The Bayes estimate

$$\hat{\theta}_{B.E} = \frac{\breve{\alpha}}{\breve{\beta}}$$

4. Monte Carlo Simulation Study

In this section, we conduct a Monte Carlo simulation study to estimate the parameters based on complete sample by using MLE, Moments method and Bayesian estimation using Gamma Prior and Exponential Prior under SEFL. R software ”version 4.1.2” is used to perform our Monte Carlo simulation study. Monte Carlo experiments were carried out based on 1000 random sample for following data generated form Exponential distribution by using numerical analysis, wherex_i is distributed as Exponential distribution for parameter (θ) with different actual values of parameter and for different samples sizes $n = 10, 20, 50, 100, 250$ and 500. We compare the performances of the MLE, Moments methods and Bayesian estimation based on the Bias and root mean squared errors (RMSE) for different sample sizes. Therefore, we report all the results up to three decimal places. Remember that Bias estimator is

$$Bias(\hat{\theta}) = \frac{1}{L}\sum_{i=1}^{L}(\hat{\theta}_i - \theta),$$

and the RMSE of the estimator is:

$$RMSE(\hat{\theta}) = \sqrt{\frac{1}{L}\sum_{i=1}^{L}(\hat{\theta}_i - \theta)^2}$$

Where $\hat{\theta}_i$ is the estimated of θ, at i^{th} experiment of L=1000 Monte Carlo experiments. The results of simulation were reported in Tables (1 to 6) with graphical representations in Figures (1 to 3). The conclusions from the Monte Carlo simulation were summarized in the following points:

- The Bias and RMSE values of the parameter θ for all estimation methods decreases as the sample size (n) increases.
- The Bayesian estimation method has more relative efficiency than MLEs and Moments methods for most parameter of Exponential distribution in all Tables.
- We can analyze that by increasing θ, the RMSE and Bias for the parameter θ decrease in most cases.

- In the Bayesian methods, when sample size is larger, the Bias and the RMSE are reduced to as little as the sample size is 250.
- Bayesian estimation using Exponential Prior, has achieved the lowest value Bias and RMSE of the classical estimation methods. This indicates that the duration data of linear accelerator is more consistent with the Exponential distribution when estimating the parameters of this distribution in Bayes.
- Figures (1 to 3) shows a visual representation of the RMSE results of the used methods estimation presented in details in Tables (1 to 3), which concludes that the **Exponential Prior** estimation outperformed the other methods estimation.

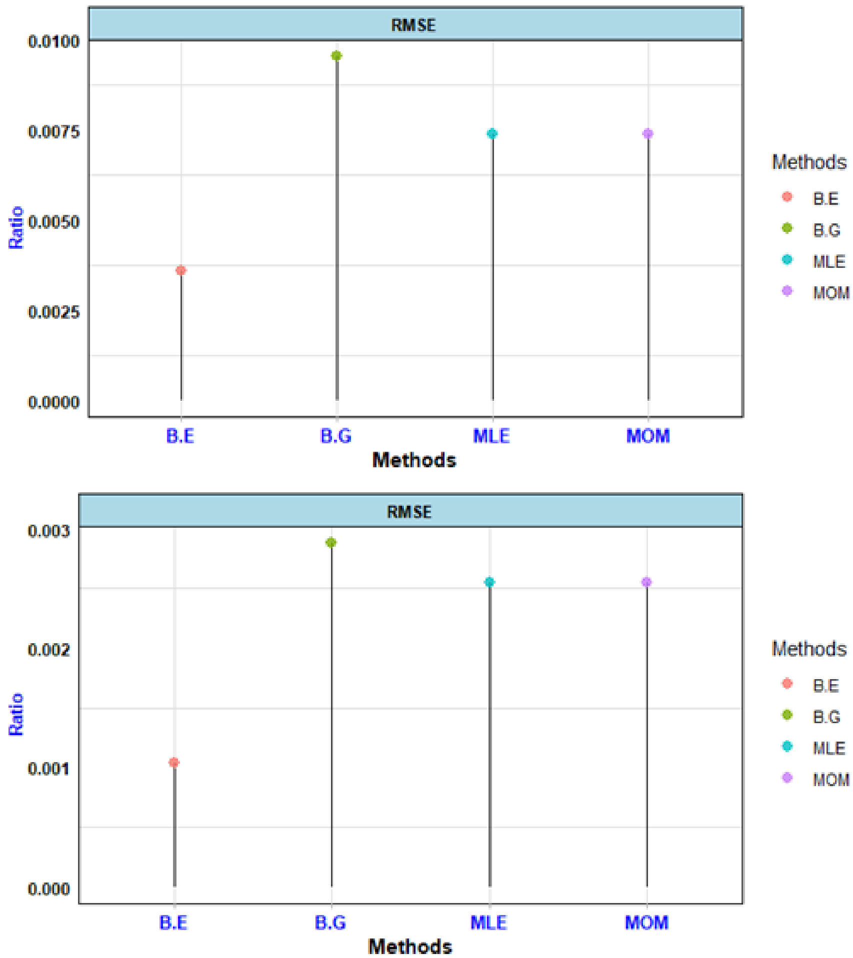

Figure 2. Performance Evaluation of Case 3: $\alpha = \theta = 1.5, \beta = 2.5,\ \lambda = 1 and n = 20$.

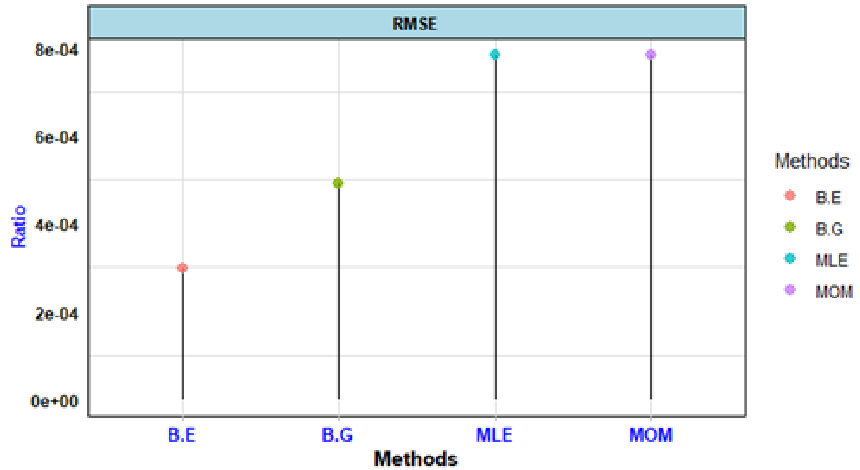

Figure 3. Performance Evaluation of Case 4: $\alpha = \theta = 1.5, \beta = 2.5,\ \lambda = 1 and n = 50$

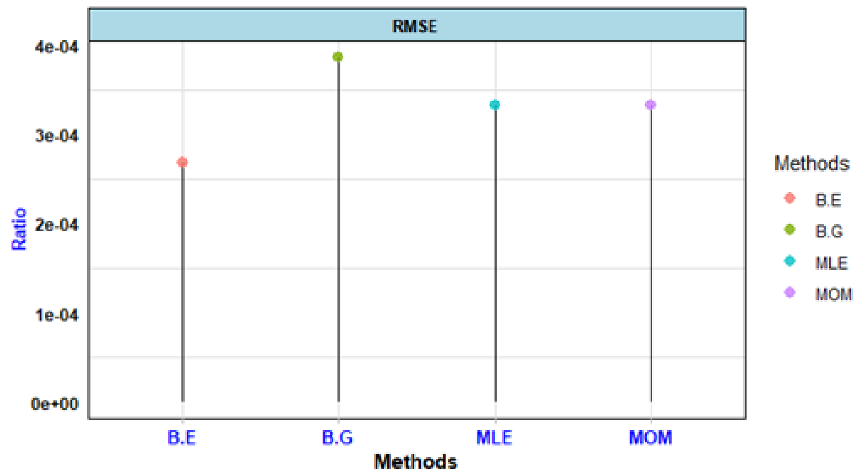

Figure 4. Performance Evaluation of Case 5: $\alpha = \theta = 1.5, \beta = 2.5,\ \lambda = 1 and n = 100$.

Table 1. Parameter estimation for Exponential distribution in case 1.

Case 1: $\alpha = \lambda = 1.25$ and $\beta = \theta = 2$												
	MLE			MOM			B.G			B.E		
n	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE
10	2.2333	0.2333	0.007378	2.2333	0.2333	0.007378	1.6986	-0.3014	0.009531	1.8861	-0.1139	0.003602
20	2.1077	0.1077	0.003406	2.1077	0.1077	0.003406	1.8361	-0.1639	0.005183	1.9452	-0.0548	0.001733
50	2.0397	0.0397	0.001255	2.0397	0.0397	0.001255	1.9303	-0.0697	0.002204	1.9777	-0.0223	0.000705
100	2.014	0.014	0.000443	2.014	0.014	0.000443	1.9595	-0.0405	0.001281	1.9837	-0.0163	0.000516
250	2.004	0.004	0.000127	2.004	0.004	0.000127	1.9822	-0.0178	0.000563	1.992	-0.008	0.000253
500	2.006	0.006	0.00019	2.006	0.006	0.00019	1.995	-0.005	0.000158	2	1E-08	1E-07

Table 2. Parameter estimation for Exponential distribution in case 2.

Case 2: $\alpha = \lambda = \beta = \theta = 1$												
	MLE			MOM			B.G			B.E		
n	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE
10	1.1166	0.1166	0.003687	1.1166	0.1166	0.003687	1.0941	0.0941	0.002976	1.0941	0.0941	0.002976
20	1.0539	0.0539	0.001705	1.0539	0.0539	0.001705	1.0486	0.0486	0.001537	1.0486	0.0486	0.001537
50	1.0199	0.0199	0.000629	1.0199	0.0199	0.000629	1.0191	0.0191	0.000604	1.0191	0.0191	0.000604
100	1.007	0.007	0.000221	1.007	0.007	0.000221	1.0068	0.0068	0.000215	1.0068	0.0068	0.000215
250	1.002	0.002	6.32E-05	1.002	0.002	6.32E-05	1.002	0.002	6.32E-05	1.002	0.002	6.32E-05
500	1.003	0.003	9.49E-05	1.003	0.003	9.49E-05	1.003	0.003	9.49E-05	1.003	0.003	9.49E-05

Table 3. Parameter estimation for Exponential distribution in case 3

Case 3: $\alpha = \theta = 1.5, \beta = 2.5$ and $\lambda = 1$												
	MLE			MOM			B.G			B.E		
n	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE
10	1.6749	0.1749	0.003331	1.6749	0.1749	0.003331	1.3286	-0.1714	0.0054	1.5373	0.0373	0.001812
20	1.5808	0.0808	0.002553	1.5808	0.0808	0.002553	1.4094	-0.0906	0.002878	1.5329	0.0329	0.0014
50	1.5298	0.0298	0.000942	1.5298	0.0298	0.000942	1.4618	-0.0382	0.001208	1.5132	0.0132	0.000417
100	1.5105	0.0105	0.000332	1.5105	0.0105	0.000332	1.4768	-0.0232	0.000734	1.5027	0.0027	8.54E-05
250	1.503	0.003	9.49E-05	1.503	0.003	9.49E-05	1.4896	-0.0104	0.000329	1.5	0.0000001	1E-07
500	1.5045	0.0045	0.000142	1.5045	0.0045	0.000142	1.4978	-0.0022	6.96E-05	1.503	0.003	9.49E-05

Table 4. Parameter estimation for Exponential distribution in case 4

Case 4: $\alpha = \lambda = \beta = \theta = 1.25$												
	MLE			MOM			B.G			B.E		
n	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE
10	1.3958	0.1458	0.004611	1.3958	0.1458	0.004611	1.3189	0.0689	0.002179	1.3396	0.0896	0.002832
20	1.3173	0.0673	0.002128	1.3173	0.0673	0.002128	1.2885	0.0385	0.001218	1.2733	0.0233	0.000737
50	1.2748	0.0248	0.000784	1.2748	0.0248	0.000784	1.2656	0.0156	0.000493	1.2594	0.0094	0.000297
100	1.2587	0.0087	0.000275	1.2587	0.0087	0.000275	1.2545	0.0045	0.000142	1.2514	0.0014	4.43E-05
250	1.2525	0.0025	7.91E-05	1.2525	0.0025	7.91E-05	1.2509	0.0009	2.85E-05	1.2497	-0.0003	9.5E-06
500	1.2538	0.0038	0.00012	1.2538	0.0038	0.00012	1.253	0.003	9.49E-05	1.2524	0.0024	7.59E-05

Table 5. Table 5: Parameter estimation for Exponential distribution in case 5.

Case 5: $\alpha = \lambda = \theta = 1.5$ and $\beta = 2$												
	MLE			MOM			B.G			B.E		
n	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE
10	1.6749	0.1749	0.005531	1.6749	0.1749	0.005531	1.4146	-0.0854	0.002701	1.4476	-0.0524	0.001657
20	1.5808	0.0808	0.002555	1.5808	0.0808	0.002555	1.4586	-0.0414	0.001309	1.4767	-0.0233	0.000737
50	1.5298	0.0298	0.000942	1.5298	0.0298	0.000942	1.4832	-0.0168	0.000531	1.4907	-0.0093	0.000294
100	1.5105	0.0105	0.000332	1.5105	0.0105	0.000332	1.4878	-0.0122	0.000386	1.4915	-0.0085	0.000269
250	1.503	0.003	9.49E-05	1.503	0.003	9.49E-05	1.494	-0.006	0.00019	1.4955	-0.0045	0.000142
500	1.5045	0.0045	0.000142	1.5045	0.0045	0.000142	1.5	0.0000001	1E-08	1.5008	0.0008	2.53E-05

Table 6. Table 6: Parameter estimation for Exponential distribution in case 6.

Case 6: $\alpha = \theta = 1.25$ and $\lambda = \beta = 1$												
	MLE			MOM			B.G			B.E		
n	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE	Estimate	Bias	RMSE
10	1.3958	0.1458	0.004611	1.3958	0.1458	0.004611	1.3619	0.1119	0.003539	1.3317	0.0817	0.002584
20	1.3173	0.0673	0.002128	1.3173	0.0673	0.002128	1.3092	0.0592	0.001872	1.2938	0.0438	0.001385
50	1.2748	0.0248	0.000784	1.2748	0.0248	0.000784	1.2736	0.0236	0.000746	1.2674	0.0174	0.0005
100	1.2587	0.0087	0.000275	1.2587	0.0087	0.000275	1.2585	0.0085	0.000269	1.2554	0.0054	0.000171
250	1.2525	0.0025	7.91E-05	1.2525	0.0025	7.91E-05	1.2525	0.0025	7.91E-05	1.2512	0.0012	3.79E-05
500	1.2538	0.0038	0.00012	1.2538	0.0038	0.00012	1.2538	0.0038	0.00012	1.2531	0.0031	0.000098

5. Conclusion

This paper presents a summary of popular estimation methods in mathematical statistics field, these include; moment method, maximum likelihood, and Bayesian estimation methods. As previously mentioned, the Bayesian estimation required the development of prior and posterior functions. The posterior function comes after selecting the prior function by multiplying the prior function times the likelihood function of the given distribution. The optimal way to select the prior function has been elusive in the past decades. A new method for picking the prior distribution in Bayesian analysis is proposed in this study. This method states that for a given distribution; the prior distribution would be better to have similar or the same distribution as the data distributed. This method's performance was compared to other estimating methods and found to be much better than other estimators. This finding was proven by a Monte Carlo simulation experiment, with certain specified performance criteria namely, RMSE and the Bias.

REFERENCES

Abd, M. N. and Rasheed, H. A. (2024). Bayesian estimation for the reliability function of two parameter exponential distribution under different loss functions. In *AIP Conference Proceedings*, volume 3036. AIP Publishing.

Arqand, A., Hahn, T. A., and Tan, E. Y.-Z. (2024). Generalized r\'enyi entropy accumulation theorem and generalized quantum probability estimation. *arXiv preprint arXiv:2405.05912*.

Atanasiu, V. (2008). Linearized credibility formula for exponentially weighted squared error loss function. *Revista Informatica Economică nr*, 2(46):12.

Box, G. E. and Tiao, G. C. (2011). *Bayesian inference in statistical analysis*. John Wiley & Sons.

Faladiba, M. N. and Ahdika, A. (2024). Bayes estimation of exponentially distributed survival data under symmetric and asymmetric loss functions. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 18(2):0973–0986.

Gelman, A., Simpson, D., and Betancourt, M. (2017). The prior can often only be understood in the context of the likelihood. *Entropy*, 19(10):555.

GENCER, G. and SARAÇOĞLU, B. (2024). Parameter estimations for exponential power distribution under loss functions: Maximum likelihood vs. bayes. *International Research in Math Sciences V*, page 29.

Gould, W., Pitblado, J., and Sribney, W. (2006). *Maximum likelihood estimation with Stata*. Stata press.

Hall, A. R. (2003). Generalized method of moments. *A companion to theoretical econometrics*, pages 230–255.

Hasan, M. and Baizid, A. (2017). Bayesian estimation under different loss functions using gamma prior for the case of exponential distribution. *Journal of Scientific Research*, 9(1):67–78.

Held, L. and Bové, D. S. (2020). Likelihood and bayesian inference. *Statistics for Biology and Health. Springer, Berlin, Heidelberg,*.

Ijaz, M., Asim, S. M., and Alamgir (2019). Lomax exponential distribution with an application to real-life data. *PloS one*, 14(12):e0225827.

Kaplan, D. (2023). *Bayesian statistics for the social sciences*. Guilford Publications.

Kundu, D. and Gupta, R. D. (2008). Generalized exponential distribution: Bayesian estimations. *Computational Statistics & Data Analysis*, 52(4):1873–1883.

Mahdavi, A., Kharazmi, O., and Contreras-Reyes, J. E. (2022). On the contaminated weighted exponential distribution: Applications to modeling insurance claim data. *Journal of Risk and Financial Management*, 15(11):500.

Maswadah, M. and Mohamed, S. (2024). Bayesian inference on the generalized exponential distribution based on the kernel prior. *Science*, 12(2):29–36.

Modarres, M. (2006). *Risk analysis in engineering: techniques, tools, and trends*. CRC press.

Rosasco, L., De Vito, E., Caponnetto, A., Piana, M., and Verri, A. (2004). Are loss functions all the same? *Neural computation*, 16(5):1063–1076.

Saini, S. (2024). Estimation of multi-stress strength reliability under progressive first failure censoring using generalized inverted exponential distribution. *Journal of Statistical Computation and Simulation*, pages 1–33.

Tanner, M. A. (1993). *Tools for statistical inference*, volume 3. Springer.

Triana, Y. and Purwadi, J. (2019). Exponential distribution parameter estimation with bayesian self method in survival analysis. In *Journal of Physics: Conference Series*, volume 1373. IOP Publishing.

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., et al. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, 1(1):1.

Zyphur, M. J. and Oswald, F. L. (2015). Bayesian estimation and inference: A user's guide. *Journal of Management*, 41(2):390–420.