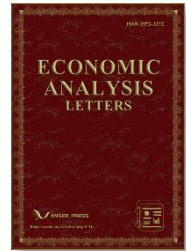




Economic Analysis Letters

Homepage: <https://www.anserpress.org/journal/eal>



A simulation study on the insurance claims distribution using Weibull distribution

Hamza Abubakar ^{a,*} and Muhammad Lawal Danrimi ^b

^a Department of Mathematics, Isa Kaita College of Education, Katsina State, Nigeria

^b Department of Accounting, Umaru Musa Yar'adua University, Katsina State, Nigeria

ABSTRACT

The Weibull distribution is extensively useful in the field of finance, insurance and natural disasters. Recently, It has been considered as one of the most frequently used statistical distributions in modelling and analyzing stock pricing movement and uncertain prediction in financial and investment data sets, such as insurance claims distribution. It is well known that the Bayes estimators of the two-parameter Weibull distribution do not have a compact form and the closed-form expression of the Bayes estimators cannot be obtained. In this paper and the Bayesian setting, it is assumed that the scale parameter of the Weibull model has a gamma prior under the assumption that its shape parameter is known. A simulation study is performed using random claims amount to compare the performance of the Bayesian approach with traditional maximum likelihood estimators in terms of Root Mean Square Errors (RMSE) and Mean Absolute Error (MAE) for different sample sizes, with specific values of the scale parameter and shape parameters. The results have been compared with the estimated result via the maximum likelihood method. The result revealed that the Bayesian approach behaves similarly to the maximum likelihood method when the sample size is small. Nevertheless, in all cases for both methods, the RMSE and MAE decrease as the sample size increases. Finally, applications of the proposed model to the insurance claim data set have been presented.

KEYWORDS

Weibull distribution; Bayesian estimator; maximum likelihood estimator; Insurance claim

* Corresponding author: Hamza Abubakar
E-mail address: zeeham4u2c@yahoo.com

ISSN 2972-3272

doi: 10.58567/eal02030008

This is an open-access article distributed under a CC BY license
(Creative Commons Attribution 4.0 International License)



Received 26 February 2023; Accepted 8 July 2023; Available online 17 July 2023

1. Introduction

The Weibull distribution is extensively useful in modelling and analysis in the field of finance, insurance and natural disasters. Data modelling in applied sciences particularly in survival and reliability engineering is a key tool in accelerating their prediction and unlocking the value of their data. It plays a vital role in the modelling reliability of data that are useful for further decision-making processes. Various statistical distribution models have been used in describing the probabilistic behaviour of random phenomena and play a crucial role in the decision-making process in Science, engineering, business and finance. The distributions are found useful in all branches of economic and scientific knowledge. For example, the classical distributions such as normal distribution, Pareto distribution, gamma distribution, Maxwell distribution and Poisson distribution, have received great attention among researchers and attracted very useful applications in every branch of science, business, engineering, demography and finance (Sharma et al. 2020). There has been a great deal of interest in the actuarial literature on the newly proposed models in several insurance application to survival modeling and analysis and other related contexts. In the insurance analysis in capitalizing on the availability of computational resources. Among these, the applications of the statistical models in insurance analysis have received much attention (Ahmad et al., 2020).

In non-life insurance, the claims process is divided into two parts: claim frequency and claim severity, with the pure or risk premium determined by the product of the underlying predicted claim rate and expected claim severity. Particular emphasis is placed on the probabilistic modeling of various features of a single batch of claims, with the aggregate claims accruing over a set timeframe, often one year, under a variety of conditions imposed on the claim frequency and severity mechanisms. For more details on insurance claims distribution refer to (Garrido, et al. 2016). Bayesian modelling plays an important role in different areas of applied statistics. In recent years, there has been a great deal of interest in actuarial modelling using various statistical models. The Bayesian approach is one of the most powerful rules of probability and statistics. It is applied in describing the probability of an event, based on prior knowledge of conditions that might be related to the event. In Statistics and modelling, Bayesian approach have had a recent resurgence with the global rise of artificial intelligence (Chen et al., 2017) and data-driven machine learning systems (Ching and Phoon, 2019) in all aspects of business (Charles et al., 2018), computational science (Vasishth et al., 2018) and technology (Park et al., 2021), economics (Kowal et al., 2017) and insurance (Zhang, 2017). Bayesian inference is being applied to many fields where data mining tools and predictive analytics are needed for knowledge discovery (Abubakar and Muhammad Sabri, 2021; Abubakar and Sabri, 2022, 2023).

In this study, the importance of statistical models as a comprehensive modelling tool for the analysis of the claims process has been reemphasized. Various studies conducted on the application of the Bayesian method in the estimation of parameters of various statistical distributions include the work of (Kaminskiy and Krivtsov, 2005) who proposes the application of interval assessment for reliability expression (rather than the Weibull parameters), which is often easier to get. Zaidi et al., (2012) in the reliability modelling, a Bayesian network approach was devised. The study subjected as systems with a Weibull failure distribution to Bayesian reliability models. A study by Sultan et al., (2014) under universal entropy and squared error loss functions, offers Bayes estimators of parameters of the inverse Weibull distribution for complete, type I, and type II censored samples. based on their predicted risks, the proposed estimators were compared (average loss over sample space). Yanuar et al., (2019) offer a Bayesian inference for Weibull distribution scale parameters. The prior distribution used in this work is Jeffreys' prior, which is an inverse gamma and non-informative prior conjugate. Their work also seeks to investigate numerous theoretical aspects of posterior distributions dependent on priors utilized, then apply them to generated data and compare both Bayes estimators. Köksal Babacan and Kaya, (2019) study was based on the assumptions of Gamma priors in investigating the fitness of the Bayes estimators of the scale and shape parameters of the Weibull distribution. The study found that continuous conjugate joint prior distribution does not exist when computing Bayes estimates for a Weibull distribution. In a recent study by Almetwally and Almongy, (2021), a Bayesian

estimation approach with squares error for the loss function was used in producing generalized power Weibull parameters. The optimum censoring strategy was chosen based on two distinct optimality criteria (mean squared error, bias, and relative efficiency). Tung et al., (2021), study on the Bayesian analysis and effectiveness of Gibbs sampling are described, and Raftery–Lewis, Geweke, and Gelman–Rubin metrics are used to assess the algorithm's convergence. Shakhatreh et al., (2021) focuses on estimating the Weibull distribution's differential entropy using multiple non-informative prior distributions. For the differential entropy, a reference prior and a probability matching prior have been devised. Although there is a large literature on frequentist estimate of Weibull distribution parameters, there hasn't been much work on Bayesian inference of Weibull distribution parameter(s) under the assumption of gamma prior to application to claims modelling.

The goal of this study is to investigate probability models based on the Weibull distribution for modelling a variety of Insurance Claims and estimating the expected claim of Insurance Plans. The main objective of the study is to explore the Bayesian approach to Weibull distribution on the assumption of gamma prior to the shape parameter. The purpose is to model adequately the number of claims occurring under insurance policies to estimate the expected number of claims. To the best of the author's knowledge, the insurance claim data has not been modelled on the assumption of Weibull distribution with scale parameters following gamma prior distribution. The specific objectives of this study are to (i) construct an insurance claims model on the assumption of Weibull distribution with shape parameter following a Gamma distribution; (ii) apply the Bayesian model framework to estimate the parameter of the model developed in (i); (iii) compare the performance of Bayesian estimator with maximum likelihood method based on the number of insurance claims data.

The rest of this paper is structured as follows. In Section 2, materials and methods including the Weibull distribution model, likelihood function, Bayesian estimation method, prior assumptions on insurance claim amount data, and method of estimating expected future insurance claims amounts from the posterior distribution are presented. In Section 3, a simulation study on the Weibull distribution as claims amount assuming a gamma prior for the scale parameter has been presented. In Section 4, performance evaluation metrics of the Bayesian approach and the maximum likelihood via simulation study are conducted. The result and discussion are presented in Section 5. This paper has finally concluded in Section 6.

2. Materials and methods

The researcher measured the expected claims data using the Bayesian approach. To determine the posterior distribution, we incorporate the likelihood function of the best-fit model of the claims return amounts with the gamma prior distribution. The expected insurance claims amount will be determined by the posterior distribution's expectation.

2.1. Weibull Distribution as a claims model

Let $X_t, t = 1, 2, 3, \dots, n$ be the random claims amounts (Garrido, et al. 2016). If a random variable X based on the assumption of Weibull distribution denoted as $X \sim \text{Weibull}(\alpha, \beta)$. The probability density function (PDF) is defined as

$$f_{\text{Weibull}}(x_t | \alpha, \beta) = \begin{cases} \beta \alpha x_t^{(\beta-1)} e^{-(\alpha x_t)^\beta}, & x_t > 0 \\ 0, & x_t \leq 0 \end{cases} \quad (1)$$

The Cumulative distribution function (CDF) of Equation (1) can be derived as follows.

$$F_{\text{Weibull}}(x_t, \alpha, \beta) = P(X_t \leq x_t) = \begin{cases} 1 - e^{-(\alpha x_t)^\beta}, & x_t \geq 0 \\ 0, & x_t < 0 \end{cases} \quad (2)$$

where β is defined as the shape representing the slope of the curve and α is the Weibull distribution scale parameter representing the characteristic life.

In our study, x_t is the Insurance claims data over investment period t ; $\beta > 0$ is the shape parameter (slope/threshold) that determines the basic shape of the Weibull distribution PDF; $\alpha > 0$ is the scale parameter (characteristic life) of the Weibull distribution (WD) showing claims distribution claims data.

The expected value and the variance of the Insurance claims data following Weibull distribution are given

$$E(X) = \alpha \Gamma\left(1 + \frac{1}{\beta}\right) \quad (3)$$

$$Var(X) = \alpha^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right] \quad (4)$$

where β is the shape, α is the scale parameter and Γ is the gamma function.

2.2. Determining Likelihood function

The Maximum likelihood function is regarded as one of the most common methods for parameter estimation. Suppose that Let $X_t, t = 1, 2, 3, \dots, n$ is the amounts of the claims, which are independent and identically distributed $W(\beta, \alpha)$ random variables, where the distribution parameters are assumed to be unknown. The distribution parameters can be estimated via the maximum likelihood method. The likelihood function of the observed data x_1, x_2, \dots, x_n can be built from Equation (1) as

$$L(\alpha, \beta | x_1, x_2, \dots, x_n) = \prod_{t=1}^n f(x_t | \alpha, \beta) \prod_{t=1}^n x_t^{(\beta-1)} e^{-\alpha \sum_{t=1}^n x_t^\beta} \quad (5)$$

The likelihood function in Equation (5) can be simplified as follows

$$L(\alpha, \beta | x_1, x_2, \dots, x_n) = \beta^n \alpha^n \prod_{t=1}^n x_t^{(\beta-1)} e^{-\alpha \sum_{t=1}^n x_t^\beta} \quad (6)$$

The log-likelihood function can be obtained by taking the natural logarithm of Equation (6) written as

$$\log(L) = n \log(\beta) + n \log(\alpha) + (\beta - 1) \sum_{t=1}^n \log(x_t) - \alpha \sum_{t=1}^n x_t^\beta \quad (7)$$

We obtained the following equation by differentiating Equation (7) according to appropriate scale and shape parameters and equating the resulting expression to zero as follows,

$$\frac{\partial \log(L)}{\partial \beta} = \frac{n}{\alpha} - \sum_{t=1}^n x_t^\beta = 0 \quad (8)$$

$$\frac{\partial \log(L)}{\partial \alpha} = \left(\frac{n}{\alpha}\right) + \sum_{t=1}^n (\log x_t) - \alpha \sum_{t=1}^n (x_t^\beta) = 0 \quad (9)$$

The MLE of parameters is determined by solving the nonlinear systems of linear equations mentioned above.

To numerically optimize (maximization) the log-likelihood function in Equation (7), it is considerably simpler to apply nonlinear optimization such as computer software package or novel metaheuristics algorithms. The estimator of scale parameters from Equation (8) is as follows

$$\hat{\alpha} = \frac{n}{\sum_{t=1}^n x_t^{\hat{\beta}}} \quad (10)$$

The estimator of the shape parameter can be obtained by substitution of Equation (9) in Equation (10) as follows,

$$n + \hat{\beta} \sum_{t=1}^n \log x_t = \frac{n \hat{\beta} \sum_{t=1}^n x_t^{\hat{\beta}} \log x_t}{\sum_{t=1}^n x_t^{\hat{\beta}}} \quad (11)$$

Equation (11) can be solved numerically or using spreadsheet package to obtain the estimate $\hat{\beta}$. The value of $\hat{\alpha}$ can be easily obtained if $\hat{\beta}$ is obtained.

2.3. Bayesian Estimation Method

Recently, the Bayesian estimation approach has attracted a lot of attention from various disciplines for analyzing various lifetime data, and it has mostly been recommended as an alternative to the traditional maximum likelihood methods. Thomas Bayes proposed the Bayesian approach, which is based on Bayes' theorem. The Bayesian approach offers a simple rule for adjusting probabilities when new information becomes available. The new information is regarded as observed data in the Bayesian modelling framework, allowing us to update our previous assumptions about parameters of interest, which are assumed to be random variables. It is preferable in the Bayesian method to estimate the parameter using the x_1, x_2, \dots, x_n data for the statistical model defined by the probability (density) function $P(x_t|\xi)$. The parameter is considered a random variable in the Bayesian approach and thus has its distribution. If prior knowledge about the parameter is not available, it is possible to make use of a non-informative prior distribution in Bayesian analysis. In this study, the Gamma prior for shape parameters are considered and no specific prior on the scale parameter is assumed.

2.4. Prior Assumptions of the observed data

It is worth noting that if the shape parameter of the distribution is known, the scale parameter will have a conjugate prior distribution, a gamma prior is assumed in this case. When both parameters of a given distribution are unknown, it is obvious that they lack conjugate priors. In this study, we first consider the known shape parameter of the Weibull distribution (Nassar et al., 2018). So we consider the following priors on λ and ψ . $\pi(\lambda)$ has a gamma prior with the scale parameter and shape parameter, $Gamma(\psi, \lambda)$, i.e. it has the PDF in Equation (12). The posterior distribution was determined by multiplying the likelihood function of the Insurance claims amounts x_1, x_2, \dots, x_n by the prior distribution under the Bayesian approach. In this case, the prior distribution of the scale parameter is assumed to come from Gamma distribution with PDF as follows

$$\pi(\alpha|\psi, \lambda) = \frac{\lambda^\psi}{\Gamma(\psi)} x_t^{\psi-1} e^{-\lambda x_t}, \lambda > 0 \quad (12)$$

The hyper-parameters ψ and λ are assumed to be known real numbers. If hyper-parameters of independent Gamma distribution priors, then an estimate of hyper-parameters of Equation (12) is presented in the next subsection.

2.5. Estimating expected future of the posterior distribution based on the observed data

If the likelihood is based on random claims amount x_1, x_2, \dots, x_n , to obtain the Bayes estimator, we multiply the prior distributions with likelihood function as follows,

$$P(x_t, \alpha, \beta) = L(f(x_t|\alpha, \beta) \times \pi(\beta|\psi) \tag{13}$$

Substituting Equations (6) and(12) into Equation (13) we have,

$$P(x_t, \alpha, \beta) = \beta^n \alpha^n \left(\prod_{t=1}^n x_t^{(\beta-1)} \right) e^{-\alpha(\sum_{t=1}^n x_t^\beta)} \times \frac{\lambda^\psi}{\Gamma(\psi)} x_t^{\psi-1} e^{-\lambda\alpha} \tag{14}$$

Then, the marginal distribution of parameters given Insurance claim data is found by taking the integral of both parameters

$$P(x_t) = \int_0^\infty P(x_t, \beta, \alpha) = \frac{\lambda^\psi}{\Gamma(\psi)} \beta^n \left(\prod_{t=1}^n x_t^{(\beta-1)} \right) \int_0^\infty \alpha^{n+\psi-1} e^{-\alpha(\sum_{t=1}^n x_t^\beta + \lambda)} d\alpha \tag{15}$$

Equation (15) can be simplified as follows

$$= \frac{\lambda^\psi}{\Gamma(\psi)} \beta^n \left(\prod_{t=1}^n x_t^{(\beta-1)} \right) \int_0^\infty \left(\frac{\rho}{\sum_{t=1}^n x_t^\beta + \lambda} \right)^{n+\psi-1} e^{-\rho} \left(\frac{1}{\sum_{t=1}^n x_t^\beta + \lambda} \right) d\rho \tag{16}$$

where ρ is defined as follows,

$$\rho = \alpha \left(\sum_{t=1}^n x_t^\beta + \lambda \right) \rightarrow d\rho = d\alpha \left(\sum_{t=1}^n x_t^\beta + \lambda \right) \tag{17}$$

$$= \frac{\lambda^\psi}{\Gamma(\psi)} \beta^n \left(\prod_{t=1}^n x_t^{(\beta-1)} \right) \left(\frac{1}{\sum_{t=1}^n x_t^\beta + \lambda} \right)^{n+\psi} \int_0^\infty \rho^{n+\psi-1} e^{-\rho} d\rho \leftarrow \Gamma(n + \psi) \tag{18}$$

$$= \frac{\lambda^\psi}{\Gamma(\psi)} \beta^n \left(\prod_{t=1}^n x_t^{(\beta-1)} \right) \left(\frac{1}{\sum_{t=1}^n x_t^\beta + \lambda} \right)^{n+\psi} \Gamma(n + \psi) \tag{19}$$

Then the likelihood function is proportioned to the marginal function and the joint posterior distribution of the two parameters is obtained as follows,

$$P(\alpha|x_t) = \frac{P(x_t, \beta, \alpha)}{P(x_t)} = \frac{\left(\prod_{t=1}^n x_t^{(\beta-1)} \right) e^{-\alpha(\sum_{t=1}^n x_t^\beta)} \times \frac{\lambda^\psi}{\Gamma(\psi)} \beta^n \alpha^{n+\psi-1}}{\frac{\lambda^\psi}{\Gamma(\psi)} \beta^n \left(\prod_{t=1}^n x_t^{(\beta-1)} \right) \left(\frac{1}{\sum_{t=1}^n x_t^\beta + \lambda} \right)^{n+\psi} \Gamma(n + \psi)} \tag{20}$$

$$\frac{\alpha^{n+\psi-1}}{\Gamma(n + \psi)} e^{-\alpha(\sum_{t=1}^n x_t^\beta + \lambda)} \left(\sum_{t=1}^n x_t^\beta + \lambda \right)^{n+\psi} \sim Gamma \left(n + \psi, \sum_{t=1}^n x_t^\beta + \lambda \right) \tag{21}$$

Equation (20) is recognized as Gamma. Therefore, the Bayes estimate of under the squared error loss function becomes.

$$\hat{\alpha} = E(\alpha) = \frac{n + \psi}{\sum_{t=1}^n x_t^\beta + \lambda} \quad (22)$$

The distributions obtained in both parameters are not similar to the known distributions and their closed-form cannot be obtained. The estimations of parameters under the quadratic loss function are the expected values of these distributions. There exist many techniques to estimate approximations. This study focuses on Bayes estimators in terms of RMSE and MAE as measuring criteria.

2.6. Simulation study on Weibull Distribution with shape parameter following a Gamma Distribution

In this section, we examine the behaviour of the Bayesian estimators for a finite sample of size n . The performances of the Bayesian estimators have been compared with maximum likelihood estimators of the shape and the scale parameters of Weibull distribution in terms of errors accumulation criteria. We assume that shape parameters follow Gamma priors. The simulation studies have been conducted on the Python programming language, developed by the author. In a simulation study based on random claim amount, for comparing the performances of the estimators, we have generated from Weibull distribution. The simulation has been carried out according to the following steps:

- Step1. sample of sizes $n = 10, 50, 80, 150, 200, \dots, 1000$ from the random claims distribution.
- Step2. Compute the maximum likelihood estimates (MLE) for the proposed model parameters.
- Step3. For each sample size, samples with $\beta = 1, \alpha = 1, 1.5, 2$ and 2.5 values estimate Weibull parameters
- Step4. Compute the RMSEs and MAE.

2.7. Performance evaluation metrics

The performance of the Bayesian approach has been evaluated and compared with the maximum likelihood methods in terms of error accumulation. A method with lower error accumulation is considered the best fit model to the claims distribution data set. The following formula have been used for RMSE and MAE computation.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\beta_t - \hat{\beta}_t)^2 + (\alpha_t - \hat{\alpha}_t)^2}{n}} \quad (23)$$

$$MAE = \frac{\sum_{t=1}^n |\beta_t - \hat{\beta}_t| + |\alpha_t - \hat{\alpha}_t|}{n} \quad (24)$$

where $\beta_t, \hat{\beta}_t, \alpha_t$ and $\hat{\alpha}_t$ are the exact value of the shape parameters, estimated values of the shape parameter, the exact value of the scale parameter and the estimated values of the scale parameter respectively.

3. Result and discussion

Simulation results have been reported in Tables 1, Figures 1 until 4, displayed the RMSE of various estimation methods of Weibull distribution while Figures 6 until 8 displayed the MAE of various methods of Weibull distribution understudy for different values of shapes and scale parameters with different samples sizes $n = 10, 50, 80, 120, 150, 200$.

Table 1. The Various Estimators of $\beta = 1, \alpha = 1, 1.5, 2,$ and 2.5 parameters of Weibull distribution.

β	α	N	Bayesian estimator		maximum likelihood estimator	
			$\hat{\beta}_t$	$\hat{\alpha}_t$	$\hat{\beta}_t$	$\hat{\alpha}_t$
1	1	10	0.9571	1.1023	1.0101	1.0473
		50	1.0019	1.0535	1.0087	1.0235
		100	1.0035	1.0307	1.0034	1.0067
		150	1.0069	1.0102	1.0019	1.0029
		200	1.0101	1.0105	1.0011	1.0016
1	1.5	10	1.0072	1.5613	1.0301	1.5031
		50	1.0025	1.5351	1.0078	1.5015
		100	1.0014	1.5197	1.0054	1.5009
		150	1.0011	1.5103	1.0016	1.5005
		200	1.0002	1.5071	1.0010	1.0002
1	2	10	0.9961	2.1011	1.0501	2.0343
		50	1.0027	2.0651	1.0187	2.0105
		100	1.0012	2.0117	1.0033	2.0046
		150	1.0005	2.0123	1.0013	2.0053
		200	1.0001	2.0005	1.0007	2.0007
1	2.5	10	0.9917	2.5334	1.0108	2.5121
		50	0.9909	2.5150	1.0097	2.5095
		100	0.9935	2.5087	1.0061	2.5041
		150	1.0008	2.5012	1.0014	2.5014
		200	1.0002	2.5007	1.0012	2.5003

The simulation study results reported in Table 1 revealed that results obtained via the Bayesian methods are similar to Maximum likelihood estimator and agreed with each other when the sample size is small. With a small sample size, the estimates using Bayesian methods were reported to have a better fit than the Maximum likelihood estimator. With an increase in sample size, MLE reported having the best result with better fitting followed by the Bayesian Estimator. When the sample size increases the accumulation of the error (RMSE and MAE) decrease in all cases.

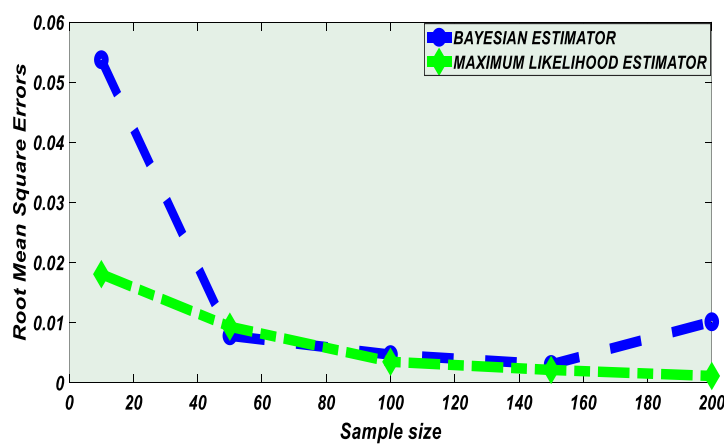


Figure 1. Comparing RMSE of Bayesian estimator with maximum likelihood estimator for $\alpha = 1$ and $\beta = 1$.

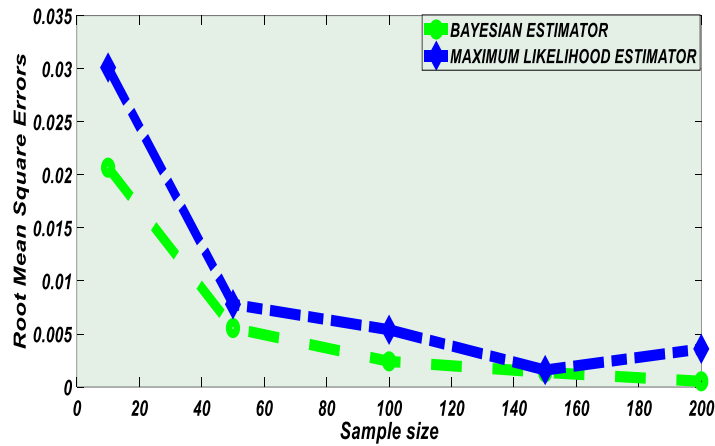


Figure 2. Comparing RMSE of Bayesian estimator with maximum likelihood estimator for $\alpha = 1.5$ and $\beta = 1$.

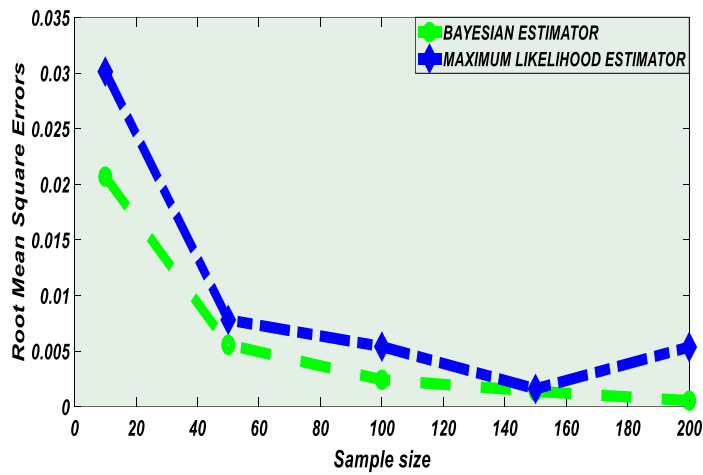


Figure 3. Comparing RMSE of Bayesian estimator with Maximum likelihood estimator for $\alpha = 2$ and $\beta = 1$.

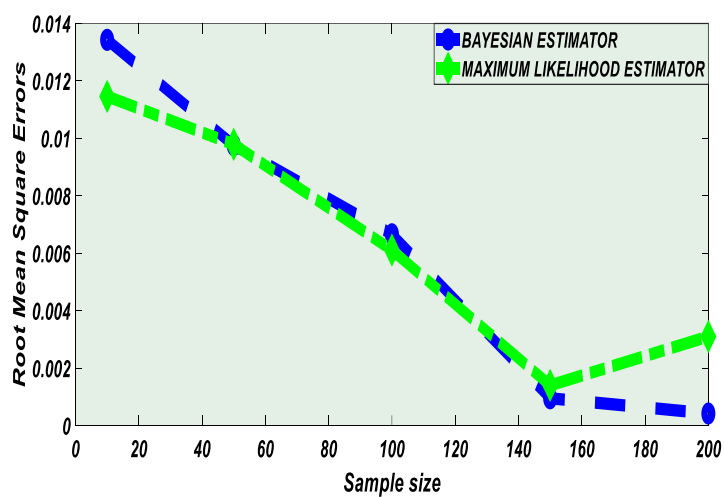


Figure 4. Comparing RMSE of Bayesian estimator with maximum likelihood estimator for $\alpha = 2.5$ and $\beta = 1$.

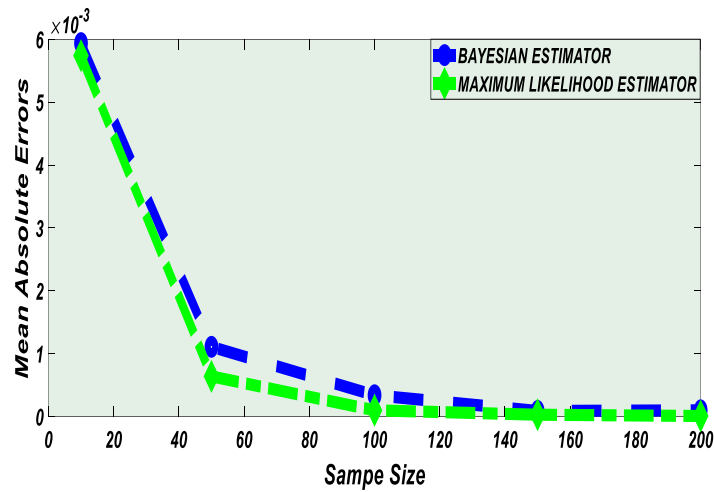


Figure 5. Comparing MAE of Bayesian estimator with maximum likelihood estimator for $\alpha = 1$ and $\beta = 1$.

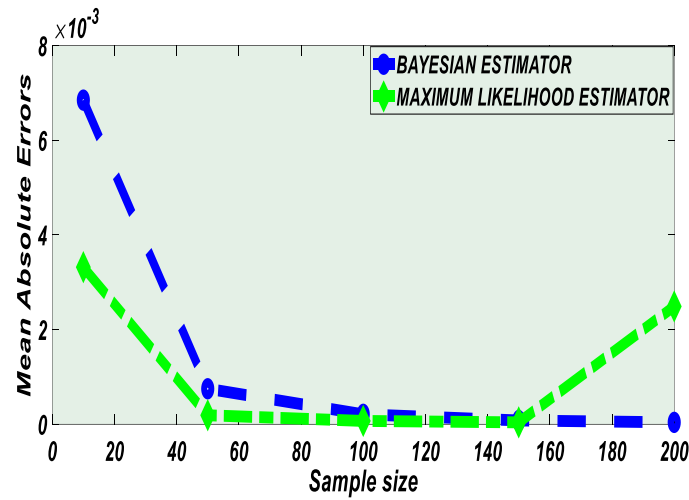


Figure 6. Comparing MAE of Bayesian estimator with maximum likelihood estimator for $\alpha = 1.5$ and $\beta = 1$.

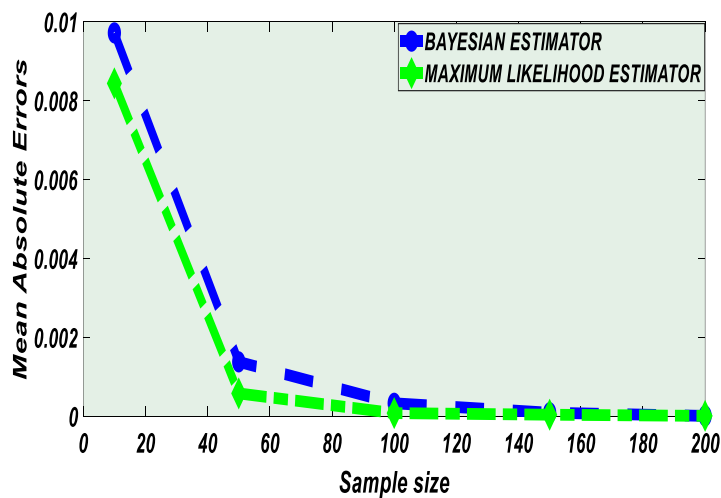


Figure 7. Comparing MAE of Bayesian estimator with maximum likelihood estimator for $\alpha = 2$ and $\beta = 1$.

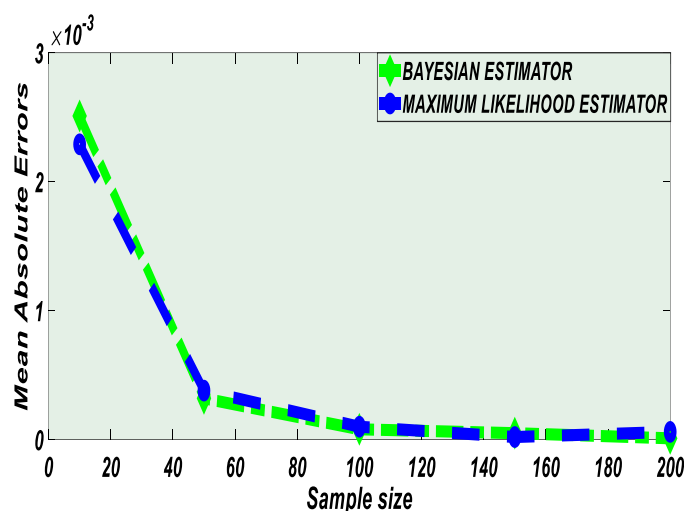


Figure 8. Comparing MAE of Bayesian estimator with Maximum likelihood estimator for $\alpha = 2.5$ and $\beta = 1$.

Figure 1 until Figure 8 displays the results of root mean square errors (RMSE) and mean absolute errors (MAE) of Bayesian estimates in comparison with maximum likelihood estimates for various sample sizes. It can be observed that the RMSE and MAE accumulation in the Bayesian methods are going hand in hand with Maximum likelihood when the sample size is small. The error accumulation in both estimates reduced drastically with the increase in sample size. However, when the sample size increases the error accumulation decrease in all cases. From Figure 1 to Figure 8, it can be seen that, for a small sample size, the estimates with Bayesian methods and Maximum likelihood are a better fit than Bayesian Estimator. This paper concluded that Bayesian methods agreed with the maximum likelihood method when the two parameters of the Weibull distribution are estimated. The study revealed that, as sample size increases, both Bayes estimation and maximum likelihood estimation have decreasing RMSE and MAE values. Especially, when the sample size is small, the Bayesian estimation approach can be used as an alternative.

4. Conclusion

Statistical distributions are important in financial sciences for data modelling and analysis. The purpose of this article was to model claim sizes using the Weibull statistical distribution, with the scale parameter assuming the Gamma distribution. The parameters of the Weibull distributions were estimated using two methods: the Bayesian approach, and the Maximum Likelihood estimator. A simulation study has been conducted to examine the performance of the Bayesian approach in comparison with the maximum likelihood estimator in estimating the parameters of the Weibull distribution for different sample sizes and parameter values.

Given the size of Insurance claim amounts and the general insurance industries, Weibull distributions can be used to model Insurance claim distributions. These are useful when analyzing claims rather than using a lengthy schedule of raw claims data. Analysis can take the form of estimating the likelihood of claims falling into a specific range, as well as reinsurance agreements in place and other mathematical analyses. It also demonstrates that the Bayesian approach and the maximum likelihood method are satisfactory and agreed at estimating the probabilities of lower claims. This is especially useful when setting up reserves. Interestingly, all estimation methods can be used concurrently; for example, when the organization is interested in the probabilities of low claims, it employs the Bayesian approach or the maximum likelihood distribution. Further research on a Weibull distribution can be conducted when both parameters follow specific distributions. According to the findings, the Weibull distribution with scale parameter following gamma distribution fits the insurance claims amount data better. This shows that

no method is superior to another; it all depends on the distributions used (Henclová, 2006; Hersch, 2019; Abubakar et al, 2022). Priority setting is primarily subjective and relies on a guide provided by the classical method. In conclusion, we presume that the estimation methods work in conjunction to ensure that good conclusions are reached in this type of claims data analysis.

Funding Statement

This research received no external funding.

Conflict of interest

The authors claim that the manuscript is completely original. The authors declare no conflict of interest.

References

- Abubakar, H., and Muhammad Sabri, S. R. (2021). A Simulation Study on Modified Weibull Distribution for Modelling of Investment Return. *Pertanika Journal of Science and Technology*, 29(4). <https://doi.org/10.47836/pjst.29.4.29>
- Abubakar, H., and Sabri, S. R. M. (2022). Simulated Annealing Algorithm as Heuristic Search Method in the Weibull Distribution for Investment Return Modelling. In *Digital Economy, Business Analytics, and Big Data Analytics Applications* (pp. 401–414). Springer. https://link.springer.com/chapter/10.1007/978-3-031-05258-3_32
- Abubakar, H., and Sabri, S. R. M. (2023). A Bayesian Approach to Weibull Distribution with Application to Insurance Claims Data. *Journal of Reliability and Statistical Studies*, 1–24. <https://doi.org/10.13052/jrss0974-8024.1611>
- Ahmad, Z., Mahmoudi, E., and Hamedani, G. (2020). A class of claim distributions: Properties, characterizations and applications to insurance claim data. *Communications in Statistics - Theory and Methods*. <https://doi.org/10.1080/03610926.2020.1772306>
- Almetwally, E. M., and Almongy, H. M. (2021). Maximum Product Spacing and Bayesian Method for Parameter Estimation for Generalized Power Weibull Distribution Under Censoring Scheme. *Journal of Data Science*. [https://doi.org/10.6339/jds.201904_17\(2\).0010](https://doi.org/10.6339/jds.201904_17(2).0010)
- Charles, V., Tsolas, I. E., and Gherman, T. (2018). Satisficing data envelopment analysis: A Bayesian approach for peer mining in the banking sector. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-017-2552-x>
- Chen, W., Shirzadi, A., Shahabi, H., Ahmad, B. Bin, Zhang, S., Hong, H., and Zhang, N. (2017). A novel hybrid artificial intelligence approach based on the rotation forest ensemble and naïve Bayes tree classifiers for a landslide susceptibility assessment in Langao County, China. *Geomatics, Natural Hazards and Risk*. <https://doi.org/10.1080/19475705.2017.1401560>
- Ching, J., and Phoon, K.-K. (2019). Constructing Site-Specific Multivariate Probability Distribution Model Using Bayesian Machine Learning. *Journal of Engineering Mechanics*. [https://doi.org/10.1061/\(asce\)em.1943-7889.0001537](https://doi.org/10.1061/(asce)em.1943-7889.0001537)
- Henclová, A. (2006). Notes on free lunch in the limit and pricing by conjugate duality theory. *Kybernetika*. <https://doi.org/10.18452/2963>
- Hersch, G. (2019). No Theory-Free Lunches in Well-Being Policy. *Review of Financial Studies*. <https://doi.org/10.1093/rfs/rfz029>
- Kaminskiy, M. P., and Krivtsov, V. V. (2005). A simple procedure for Bayesian estimation of the Weibull distribution. *IEEE Transactions on Reliability*. <https://doi.org/10.1109/TR.2005.858093>
- Köksal Babacan, E., and Kaya, S. (2019). A simulation study of the Bayes estimator for parameters in Weibull distribution. *Communications Faculty Of Science University of Ankara Series A1 Mathematics and Statistics*. <https://doi.org/10.31801/cfsuasmas.455276>
- Kowal, D. R., Matteson, D. S., and Ruppert, D. (2017). A Bayesian Multivariate Functional Dynamic Linear Model. *Journal of the American Statistical Association*. <https://doi.org/10.1080/01621459.2016.1165104>
- Nassar, M., Afify, A. Z., Dey, S., and Kumar, D. (2018). A new extension of Weibull distribution: Properties and different methods of estimation. *Journal of Computational and Applied Mathematics*. <https://doi.org/10.1016/j.cam.2017.12.001>
- Park, S., Choi, S., and Jun, S. (2021). Bayesian structure learning and visualization for technology analysis.

- Sustainability (Switzerland)*. <https://doi.org/10.3390/su13147917>
- Shakhatreh, M. K., Dey, S., and Alodat, M. T. (2021). Objective Bayesian analysis for the differential entropy of the Weibull distribution. *Applied Mathematical Modelling*. <https://doi.org/10.1016/j.apm.2020.07.016>
- Sultan, K. S., Alsadat, N. H., and Kundu, D. (2014). Bayesian and maximum likelihood estimations of the inverse Weibull parameters under progressive type-II censoring. *Journal of Statistical Computation and Simulation*. <https://doi.org/10.1080/00949655.2013.788652>
- Tung, Y. L., Ahmad, Z., Kharazmi, O., Ampadu, C. B., Hafez, E. H., and Mubarak, S. A. M. (2021). On a New Modification of the Weibull Model with Classical and Bayesian Analysis. *Complexity*. <https://doi.org/10.1155/2021/5574112>
- Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., and Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics*. <https://doi.org/10.1016/j.wocn.2018.07.008>
- Yanuar, F., Yozza, H., and Rescha, R. V. (2019). Comparison of Two Priors in Bayesian Estimation for Parameter of Weibull Distribution. *Science and Technology Indonesia*. <https://doi.org/10.26554/sti.2019.4.3.82-87>
- Zaidi, A., Bouamama, B. O., and Tagina, M. (2012). Bayesian reliability models of weibull systems: State of the art. *International Journal of Applied Mathematics and Computer Science*. <https://doi.org/10.2478/v10006-012-0045-2>
- Zhang, Y. (2017). Bayesian analysis of big data in insurance predictive modeling using distributed computing. *ASTIN Bulletin*. <https://doi.org/10.1017/asb.2017.15>