Comparison of Two Priors in Bayesian Estimation for Parameter of Weibull Distribution

Ferra Yanuar*1, Hazmira Yozza1, Ratna Vrima Rescha1

1Department of Mathematics, Faculty of Mathematics and Natural Science, Andalas University, Kampus Limau Manis, 25163, Padang-Indonesia

*Corresponding author: ferrayanuar@sci.unand.ac.id

Abstract
This present study purposes to conduct Bayesian inference to estimate scale parameters, denoted by θ, with known location parameter or β, of Weibull distribution. There are two types of prior distributions used in this study, conjugate prior and non-informative prior. As conjugate prior is inverse gamma, and as non-informative prior is Jeffreys’ prior. This research also aims to study several theoretical properties of posterior distribution based on prior used implement it to generated data and make comparison between both Bayes estimator as well. The method used to evaluate as the best estimator is based on the smallest Mean Square Error (MSE) value. This study proveds that Bayes estimator using conjugate prior produces better estimated parameter value estimate non-informative prior since it produces smaller MSE value, for condition θ > 1 based on analytic and simulation study. Meanwhile for θ < 1 both priors could not yield acceptable estimated parameter value.

Keywords
Bayesian inference, Weibull distribution, prior conjugate, non-informative prior, Mean Square Error

Received: 20 June 2019, Accepted: 18 July 2019
https://doi.org/10.26554/sti.2019.4.3.82-87

1. INTRODUCTION
Weibull models are used to describe various types of observed failures of components and phenomena. The Weibull distribution is widely used in life data analysis, reliability engineering and elsewhere due to its versatility and also its relative simplicity. Much of the attractiveness of the Weibull distribution is due to the wide variety of shapes which can assume by altering its parameters. The two-parameter Weibull distribution has one shape and one scale parameter. The random variable X follows Weibull distribution with the shape and scale parameters respectively denoted by β, θ > 0 has probability density function of the form (Aslam et al., 2014)

\[ f(x|\theta, \beta) = \frac{\beta}{\theta} x^{\beta-1} \exp \left[-\frac{x^\beta}{\theta}\right], x \geq 0, \theta, \beta > 0 \]

In this study, we concern the estimation of the scale parameter of the Weibull distribution, with known shape parameter. Many studies has been done in estimating such parameter using the maximum likelihood estimator (MLE) and/or Bayes method, such as study conducted by Ieren and Oguntunde (2018), Aslam et al. (2014) or Abdulabaas et al. (2013). All studies obtained that Bayes method yielded better estimated value than MLE.

In using Bayes method, most of researchers used the non-informative prior as the prior distribution, those are uniform prior or Jeffreys’ prior. Meanwhile many studies suggested to use the conjugate prior in order to obtain better estimated values (Thamrin et al., 2018), (Corrales and Cepeda-Cuervo, 2019). No work has been done in comparing between conjugate prior and non-informative prior yet. Thus, this present study aims to explore the comparison between both priors distribution in Bayes estimation method analytically and using simulation study (Muharisa et al., 2018). The theoretical properties of both Bayes estimator are also studied. In Section 2 we presents the literature review related to Bayes methods including invers Gamma distribution as conjugate prior and Jeffreys’ method as non-informative prior. Results from theoretical and simulation study and also discussions are presented in Section 3. Meanwhile Conclusions from this study are provided in Section 4.

2. BASIC CONCEPTS IN BAYESIAN METHOD
Suppose that X_1, X_2, \ldots, X_n are random sample from the distribution with pdf f(x|\theta) where \theta is the parameter of the distribution. Estimation of parameters \theta will be based on random samples X_1, X_2, \ldots, X_n. The Bayes method is one parameter estimation method that can be used for
this purpose. The Bayes method combines the likelihood function and the prior distribution of the parameters to be estimated, so that the posterior distribution is obtained which will be the basis of the parameter estimation (Thamrin et al., 2018).

The best estimator is the estimator which produces the smallest MSE value. The MSE of \( \mu \) is calculated based on the following formulae \( MSE(\mu) = Var(\mu) + [b(\mu)]^2 \) (Corrales and Cepeda-Cuervo, 2019).

2.2 Invers Gamma Distribution as Conjugate Prior

If \( X_1, X_2, \ldots \) are random samples from Weibull distribution, written as \( \text{Wei}(\theta, \beta) \) with unknown \( \beta \), so that its likelihood function is:

\[
    f(x|\theta) = \frac{n}{\theta^\beta} \prod_{i=1}^{n} x_i^{\beta-1} \exp\left[-\left(\frac{x_i}{\theta}\right)^\beta\right]
    \]

\[
    = \left(\frac{\beta}{\theta}\right)^n \left(\prod_{i=1}^{n} x_i^{\beta-1}\right) \exp\left[-\left(\frac{\sum_{i=1}^{n} x_i^\beta}{\theta}\right)\right] \tag{1}
    \]

2.2.1 Invers Gamma Distribution as Conjugate Prior

If a random variable \( Y \) has Invers Gamma distribution with scale parameter \( \omega > 0 \) and shape parameter \( \tau > 0 \) written as \( Y \sim IG(\tau, \omega) \), then the following formulae

\[
    f(y; \tau; \omega) = \frac{\omega^\tau}{\Gamma(\tau)} y^{-(\tau+1)} e^{-\frac{\omega}{y}}, \quad y > 0 \quad \text{and} \quad 0
    \]

The natural logarithm of the likelihood function from data that has Weibull distribution is formulated as

\[
    \ln(f(x|\theta)) = \ln\left(\frac{\beta}{\theta}\right)^n \prod_{i=1}^{n} x_i^{\beta-1} \exp\left[-\left(\frac{\sum_{i=1}^{n} x_i^\beta}{\theta}\right)\right] = n \ln \beta - \ln \theta + (\beta - 1) \sum_{i=1}^{n} \ln x_i \frac{\sum_{i=1}^{n} x_i^\beta}{\theta}
    \]

Therefore Fisher’s information for the parameter \( \theta \) that has Weibull distribution is given by

\[
    I(\theta) = -nE\left[\frac{\partial^2}{\partial \theta^2} \ln f(x|\theta)\right]
    \]

2.3 Jefferys’ Prior As Non-informatif Prior

Jeffreys’ prior is the most widely used noninformative prior in Bayesian analysis. This method is also attractive because it is proper under mild conditions and requires no elicitation of hyperparameters whatsoever. Jeffreys’ rule is derived from likelihood function then take the prior distribution to be the determinant of the square root of the Fisher information matrix (Corrales and Cepeda-Cuervo, 2019), denoted by \( f(\theta) \propto \sqrt{I(\theta)} \).

Based on Eq. (3) it’s identified that \( \theta|x \) using Invers Gamma distribution with parameters \( n+1 \) and \( \sum_{i=1}^{n} x_i^\beta + 1 \) denoted by \( \theta_{IG}|x \sim IG(n+1, \sum_{i=1}^{n} x_i^\beta + 1) \).
Then posterior distribution for $\theta$ is constructed as following

\[
(f(x|\theta) = \frac{f(x|\theta)f(\theta)}{\int f(x|\theta)f(\theta)d\theta} = \frac{f(x, \theta)}{f(x)}
\]

\[
= \frac{n^\beta n^\beta \prod_{i=1}^{n} x_i^{\beta-1} \theta^{-\theta(n+1)}}{\Gamma(n) \prod_{i=1}^{n} x_i^{\beta-1} \Gamma(n)} \exp\left[-\left(\frac{\sum_{i=1}^{n} x_i^\beta}{\theta}\right)\right]
\]

\[
= \frac{(\sum_{i=1}^{n} x_i^\beta)^2}{\Gamma(n)} \theta^{-(n+1)} \exp\left[-\left(\frac{\sum_{i=1}^{n} x_i^\beta}{\theta}\right)\right]
\]

Thus, (1 - $\alpha$)100% Bayesian credible interval for $\theta_{IG}$ is given by

\[
\sum_{i=1}^{n} x_i^\beta + 1 n - Z_\alpha \left(\frac{(\sum_{i=1}^{n} x_i^\beta)^2}{n^2(n-1)}\right) \leq \theta_{IG}
\]

3. RESULTS AND DISCUSSION

In this section we will present the results of several theoretical properties of Bayes inference for scale parameter from Weibull distribution using prior conjugate and noninformative prior. Both Bayes estimators will be compared using the best estimator criteria, that is the smallest value of Mean square Error (MSE).

3.1 Analytical Study

In this section, it will be explained theoretically the estimation process of scale parameters $\theta$ of the Weibull distribution with known shape parameter $\beta$ using the Bayes method. In estimating parameter, it uses two types of prior distributions, namely prior conjugate, used inverse Gamma distribution and noninformative prior, that is Jeffrey’s method.

3.1.1 Bayes Inference with Inversive Gamma Distribution as Conjugate Prior

It has been proved that $\theta_{IG} \sim IG(n + 1, \sum_{i=1}^{n} x_i^\beta + 1)$. Thus the posterior mean for $\theta_{IG}$ is given by

\[
\hat{\theta}_{IG} = E(\theta_{IG}) = \frac{\sum_{i=1}^{n} x_i^\beta + 1}{(n + 1) - 1} = \frac{\sum_{i=1}^{n} x_i^\beta + 1}{n}, n > 0
\]

This posterior mean is also the estimator for $\theta_{IG}$ of Weibull distribution with known shape parameter. The posterior variance for $\theta_{IG}$ is expressed as

\[
(S'_{IG}) = Var(\theta_{IG}) = \frac{(\sum_{i=1}^{n} x_i^\beta + 1)^2}{n^2(n-1)}, n > 1
\]

The (1 - $\alpha$)100% Bayesian credible interval for $\theta$ is approximately

\[
\hat{\theta}_{Bayes} - Z_\alpha \left(\frac{(\sum_{i=1}^{n} x_i^\beta)^2}{n^2(n-1)}\right) \leq \theta \leq \hat{\theta}_{Bayes} + Z_\alpha \left(\frac{(\sum_{i=1}^{n} x_i^\beta)^2}{n^2(n-1)}\right)
\]

3.1.2 Bayes Inference with Jeffrey’s Prior As Noninformative Prior

Based on Eq. 5 is known that $\theta_{j} \sim IG(n, \sum_{i=1}^{n} x_i^\beta)$ Estimator of scale parameter $\theta_{j}$ is provided by the posterior mean that is given by

\[
\hat{\theta}_{j} = E(\theta_{j}) = \frac{\sum_{i=1}^{n} x_i^\beta}{n - 1}, n > 1
\]

And the posterior variance is expressed as

\[
(S'_{j}) = Var(\theta_{j}) = \frac{(\sum_{i=1}^{n} x_i^\beta)^2}{(n - 1)^2(n - 2)}, n > 2
\]

The (1 - $\alpha$)100% Bayesian credible interval for $\theta_{j}$ is given by

\[
\frac{\sum_{i=1}^{n} x_i^\beta}{n - 1} - Z_\alpha \left(\frac{(\sum_{i=1}^{n} x_i^\beta)^2}{(n - 1)^2(n - 2)}\right) \leq \theta \leq \frac{\sum_{i=1}^{n} x_i^\beta}{n - 1} + Z_\alpha \left(\frac{(\sum_{i=1}^{n} x_i^\beta)^2}{(n - 1)^2(n - 2)}\right)
\]
Bayes Inference for Scale Parameter, True value for \( \theta = 2 \)

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Sample Size</th>
<th>Conjugate Prior (Invers Gamma)</th>
<th>Non-informative Prior (Jeffreys’ Prior)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated Mean</td>
<td>n = 25</td>
<td>1.68624</td>
<td>1.71483</td>
</tr>
<tr>
<td></td>
<td>n = 100</td>
<td>1.80989</td>
<td>1.81807</td>
</tr>
<tr>
<td></td>
<td>n = 150</td>
<td>1.82652</td>
<td>1.83207</td>
</tr>
<tr>
<td>Estimated Variance</td>
<td>n = 25</td>
<td>0.11847</td>
<td>0.12785</td>
</tr>
<tr>
<td></td>
<td>n = 100</td>
<td>0.03308</td>
<td>0.03372</td>
</tr>
<tr>
<td></td>
<td>n = 150</td>
<td>0.02239</td>
<td>0.02267</td>
</tr>
<tr>
<td>MSE</td>
<td>n = 25</td>
<td>0.11533</td>
<td>0.13273</td>
</tr>
<tr>
<td></td>
<td>n = 100</td>
<td>0.03285</td>
<td>0.03406</td>
</tr>
<tr>
<td></td>
<td>n = 150</td>
<td>0.02228</td>
<td>0.02282</td>
</tr>
<tr>
<td>Width of</td>
<td>n = 25</td>
<td>1.34927</td>
<td>1.40166</td>
</tr>
<tr>
<td>Bayesian Credible Interval</td>
<td>n = 100</td>
<td>0.71305</td>
<td>0.71991</td>
</tr>
<tr>
<td></td>
<td>n = 150</td>
<td>0.58656</td>
<td>0.59033</td>
</tr>
</tbody>
</table>

\[
\frac{n\theta^2}{(n-1)^2} \text{ and } \text{bias}(\hat{\theta}_J) = E(\hat{\theta}_J) - \theta = \frac{n\theta}{n-1} - \theta = \frac{\theta}{n-1}. \text{ Thus MSE is expressed as }
\]

\[
MSE(\hat{\theta}_J) = \text{var}(\hat{\theta}_J) + [\text{bias}(\hat{\theta}_J)]^2 = \frac{n\theta^2}{(n-1)^2} + \left[ \frac{\theta}{n-1} \right]^2 = \frac{(n+1)\theta^2}{(n-1)^2} (13)
\]

### 3.1.3 Comparison of Both MSES

Based on two formulae for MSE, \( \hat{\theta}_{IG} \) and \( MSE(\hat{\theta}_J) \) obtained above, it will be determined which one is better estimator.

It has been obtained that \( MSE(\hat{\theta}_J) = \frac{(n+1)\theta^2}{(n-1)^2} \). In the case \( \theta > 1 \), implies \( \theta^2 > 1 \), so

\[
MSE(\hat{\theta}_J) = \frac{n\theta^2 + \theta^2}{(n-1)^2} > \frac{n\theta^2 + \theta^2}{(n-1)^2} = MSE(\hat{\theta}_{IG})
\]

For condition \( \theta > 1 \), it is proven that \( MSE(\hat{\theta}_J) > MSE(\hat{\theta}_{IG}) \) it means the value of the MSE with prior conjugate (Inverse Gamma) is smaller than MSE with non-informative priors (Jeffreys’ prior). This means that the Bayes estimator with conjugate priors is better estimator than noninformative priors.

### 3.2 Simulation Study

This research then conducted a simulation study to implement the result of analytical study above. The simulation study also aims to examine how similar the estimates of the parameters of the models are, compared with the true values of the parameters.

In this simulation study, \( n = 25, 100 \) and 150 data were generated from Weibull distribution, \( WEi(2, 1) \). Based on

\[
n = 25, \text{ it is calculated that } \sum_{i=1}^{25} x_i^2 = 41.1560. \text{ Here, it is assumed that } \beta = 1, \text{ the value of } \theta \text{ will be estimated using the Bayes method. In this study, we consider two conditions for } \theta, \text{ these are } \theta > 1 \text{ and } \theta < 1.
\]

#### 3.2.1 Bayes Inference for Parameter \( \theta > 1 \) with Conjugate Prior

It has been obtained that the posterior distribution for \( \theta_{IG} \) is Invers Gamma distribution, written as \( IG(n + 1, \sum_{i=1}^{n} x_i^2 + 1) \). So that the posterior distribution for \( \theta_{IG} \) based on generated data is \( IG(26, 42.15606) \). Posterior mean for \( \theta_{IG} \) is \( E(\theta_{IG}) = \sum_{i=1}^{n} x_i^2 + 1 = 1.686242 \), posterior variance for \( \theta_{IG} \) is \( Var(\theta_{IG}) = \frac{\sum_{i=1}^{n} x_i^2}{(n-1)^2} = 0.1184755 \). Furthermore, the MSE value based on generated data is \( MSE(\theta_{IG}) = \frac{n\theta^2 + 1}{n^2} = 0.1153365 \). By taking \( \alpha = 5\% \), the 95\% Bayesian credible interval for \( \theta_{IG} \) is approximately \( 1.011605 ; 2.360879 \).

#### 3.2.2 Bayes Inference for Parameter \( \theta > 1 \) with Non Informative Prior

It is obtained that posterior distribution for \( \theta_{IG} \) is Invers Gamma distribution, \( \theta_{J} \sim IG(n, \sum_{i=1}^{n} x_i^2) \). Based on generated data, it is obtained \( \theta_{J} \sim IG(25, 41.15606) \). The posterior mean is \( E(\theta_{J}) = \frac{\sum_{i=1}^{n} x_i^2}{n-1} = 1.714836 \) and posterior variance is \( Var(\theta_{J}) = \frac{(\sum_{i=1}^{n} x_i^2)^2}{(n-1)^2} = 0.1287549 \). Meanwhile MSE value is \( MSE(\theta_{J}) = \frac{(n+1)\theta^2}{n^2} = 0.1327382 \). The 95\% Bayesian credible interval for \( \theta_{J} \) is approximately \( 1.014003 ; 2.415669 \).

The same process then implemented to \( n = 100 \) and 150 generated from Weibull distribution as well. Complete results of calculation for all three groups of data are presented in Table 1.

From the Table 1 above, it can be seen that the Bayes estimator with prior Invers Gamma produces the estimated value is closer to the true value for \( \theta = 2 \) compared to
Jeffrey’s prior for all three groups data. In general it can be seen that both estimators are consistent estimator because more data, the estimated value increasingly approaching the true value. This present study also proved that the value of variance, the MSE value of Inverse Gamma prior results in smaller value than Jeffrey’s prior for all three selected sample sizes. Invers Gamma Prior also results narrower 95% Bayesian confidence interval for all three sample sizes compared to Jeffreys’ prior.

The same estimation process is also carried out for several true values for $\theta > 1$ and linear results are obtained. Thus it can be concluded that for $\theta > 1$, the prior conjugate is better estimator than prior non-informative for case of parameter of Weibull distribution.

### 3.2.3 Bayes Inference for Parameter $\theta < 1$

In the same ways as in the previous section, simulations are carried out for several values for $\theta < 1$. To save the space, simulation study was done for $n = 100$ with the true values for $\theta$ are 0.3, 0.5, 0.7 and 0.9. Complete results of the study are provided in Table 2.

Based on Table 2 above, it can be seen that the Bayes estimator with the two selected priors produces estimated mean that does not approach the corresponding true values. The estimated value is further away from the true value of $\theta$ which is set for the value of $\theta$ which is close to 0. The estimated value here are not acceptable. Therefore, this simulation study proved that the scale parameter of the Weibull distribution for $\theta < 1$ cannot be used.

The estimated value is further away from the true value of $\theta$ which is set for the value of $\theta$ which is close to $\theta$. The estimated value here are not acceptable. Therefore, this simulation study proved that the scale parameter of the Weibull distribution for $\theta < 1$ cannot be used.

### 4. CONCLUSIONS

This study aims to compare the results of estimating scale parameters of the Weibull distribution by using prior conjugates and non-informative priors on the posterior distribution. Comparison is done analytically and with simulation studies. Theoretically it is proven that for the value of $\theta > 1$, the MSE value of the prior conjugate will yield a smaller estimation value than non-informative priors, while the value for $\theta < 1$ does not produce a consistent predictive value.

While the results of simulation studies proved that the value of $\theta < 1$ will produce a value that is far from the true value so that the estimated value is not acceptable. Whereas for the value of $\theta > 1$, it is concluded that the estimated value approaches the true value by increasing the sample size. Invers Gamma prior produced better estimate mean compared Jeffreys’ prior. This conjugate prior also resulted smaller estimated variance, smaller MSE and narrower Bayesian credible interval than Jeffreys’ prior. Thus it can be concluded that inverse gamma prior (prior conjugate) gives the better estimator than Jeffreys’ prior.

### REFERENCES

Abdulabaas, F., Y. M. Al Mayali, and I. A. Neama (2013). A Comparison Between the Bayesian and the Classical...
Estimators of Weibull Distribution. *Journal of Kufa for Mathematics and Computer*, 1(8); 21–28


