STATISTICS | RESEARCH ARTICLE

Classical and Bayesian estimation of Weibull distribution in presence of outliers

Puneet Kumar Gupta1 and Alok Kumar Singh1*

Abstract: This study deals with the classical and Bayesian estimation of the parameters of Weibull distribution in presence of outlier. In classical setup, the maximum likelihood estimates of the model parameters along with their standard errors (SEs) and confidence intervals are computed. Bayes estimates along with their posterior SEs and highest posterior density credible intervals of the parameters are also obtained. Markov chain Monte Carlo technique such as Metropolis–Hastings algorithm has been used to simulate sample from the posterior densities of the parameters. Finally, a real data study illustrates the applicability of the proposed model.

Subjects: Science; Mathematics & Statistics; Applied Mathematics; Statistics & Probability; Statistics; Statistical Computing; Statistics & Computing; Statistical Theory & Methods

Keywords: Weibull distribution; outlier; Gibbs sample; Metropolis–Hastings algorithm; MCMC method

AMS subject classification: 62F15

1. Introduction
In life testing experiment, sometimes, an experimenter faces a situation where some of the observations in a sample are too small or too large in comparison to the remaining sets of observations. In other words, during an experimentation, we come across circumstances where one or more observations may not be homogeneous to rest of the observations and hence can be treated as outliers which may arise due to a variety of reasons (Barnett & Lewis, 1994).

ABOUT THE AUTHORS
Puneet Kumar Gupta received his M Phil degree in Statistics from the CCS University, Meerut, India. Currently, he is a research scholar at Department of Statistics, University of Allahabad, Allahabad, India. His research interests include Applied Statistics, Bayesian inference, load-sharing models, distribution theory, and reliability analysis.

Alok Kumar Singh received his master’s degree in Statistics from the University of Allahabad, Allahabad. Currently, he is a research scholar at Department of Statistics, University of Allahabad, Allahabad, India. His research interests include Spatial Statistics, Bayesian inference, and outlier data.

PUBLIC INTEREST STATEMENT
Detecting outliers can be important either because the outlying observations themselves are of interest, or because one wants to prevent outlier contamination of subsequent estimates. The problem of estimation of parameters has been used for many years but now there is a growing concern about its use when the underlying assumptions are violated, especially when an outlier or outliers is present. Even a single grossly outlying observation may completely spoil the parameter estimates involved in the model and hence the efficiency of the estimators if the outlier(s) are ignored. In life testing experiment, sometimes, an experimenter faces a situation where some of the observations in a sample are too small or too large in comparison to the remaining sets of observations. This study considers the classical and Bayesian estimation of the parameters of Weibull distribution by taking into account the presence of outlier.
According to Dixit, Moore, and Barnett (1996), we assume that a set of random variables represents the distance of an infected sampled plant from a plot of plants inoculated with a virus. Some of the observations are derived from the airborne dispersal of the spores and are distributed according to the gamma distribution. The remaining observations out of n random variables (say k) are present because aphids that are known to be carriers of barley yellow mosaic dwarf virus (BYMDV) have passed the virus into the plants when the aphids feed on the sap.

Detecting outliers can be important either because the outlying observations themselves are of interest, or because one wants to prevent outlier contamination of subsequent estimates. The problem of estimation of parameters has been used for many years but now there is a growing concern about its use when the underlying assumptions are violated, especially when an outlier or outliers is present. Even a single grossly outlying observation may completely spoil the parameter estimates involved in the model and hence the efficiency of the estimators if the outlier(s) are ignored. First of all, Kale and Sinha (1971) proposed the estimation of expected life in presence of an outlier in the context of life testing experiments by considering exponential distribution as a life time model. Pettit (1988) adopted a Bayesian approach to the modeling of outliers and used it in relation to members of the exponential family. Further, Dixit (1989) adopted a new approach for estimation of parameters of gamma model in presence of outlier. Using the approach of Dixit (1989), number of authors like Dixit and Nasiri (2001), Nasiri and Pazira (2010), and Deiri (2011, 2012) considered estimation of parameters in presence of outlier for different lifetime models.

The Weibull distribution has been frequently used in life testing and survival analysis, especially for describing the fatigue failures. Weibull (1951) used this distribution for vacuum tube failures and Lieblein and Zelen (1956) consider it for ball bearing failures. Singh, Rathi, and Kumar (2013) analyzed the \( k \)-components load-sharing parallel system model assuming the failure time distribution of the components as Weibull. Recently, Chaturvedi, Pati, and Tomer (2014) consider robust Bayesian analysis of the Weibull failure model under a sigma-contamination class of priors for the parameters. We refer to Mann (1968) and the reference cited therein, where Weibull distribution gives a variety of situations in which this distribution can be used for various types of data.

Dixit (1994) obtained predictive distribution of \( r \)-th order statistics for the future samples based on the original sample from the Weibull distribution in the presence of \( k \) outlier. Hassan, Elsherpieny, and Shalaby (2013) discussed the estimation procedure of stress–strength reliability for the same in presence of \( k \) outlier. Also, Khokan, Bari, and Khan (2013) developed a weighted maximum likelihood approach for the Weibull distribution in the presence of \( k \) outlier. However, all the above-mentioned studies (Dixit, 1994; Hassan et al., 2013; Khokan et al., 2013) consider the estimation procedure for Weibull distribution in presence of \( k \) outlier, but the study fully devoted to estimation procedure in presence of outliers based on the approach given by Dixit (1989) is yet to be explored. Also, there is lack of real-life application in all the above-mentioned studies as they only considered the simulated data-set. In view of the above, the present study deals with the classical and Bayesian estimation of the parameters of Weibull model in presence of outlier generated from exponential distribution. The maximum likelihood estimates (MLEs) of the model parameters along with their standard errors (SEs) are computed. In Bayesian setup, Markov chain Monte Carlo technique such as Metropolis–Hastings algorithm has been used to simulate sample from the posterior densities of the parameters. Finally, the results obtained through simulation study have been concluded.

2. Description of the model with \( k \) outliers

Let random variables be \((X_1, X_2, \ldots, X_n)\) such that \( k \) of them are from exponential distribution with parameter \( \beta \) with p.d.f.

\[
g(x; \beta) = \beta e^{-\beta x}; \quad x, \beta > 0
\]

and \( n - k \) are from the Weibull distribution with shape parameter \( \alpha \) and scale parameter \( \beta \) with p.d.f.
\( f(x; \alpha, \beta) = \alpha \beta x^{\alpha-1} e^{-\beta x}; \alpha, \beta > 0. \)

According to Dixit (1989), the joint distribution of \((X_1, X_2, \ldots, X_n)\) in the presence of \(k\) outliers can be written as

\[
\begin{align*}
    f(x_1, x_2, \ldots, x_n; \alpha, \beta) &= \prod_{j=1}^{n} f(x_j; \alpha, \beta) \sum_{\Delta} \prod_{r=1}^{k} \frac{g(x_{\hat{r}}; \beta)}{f(x_{\hat{r}}; \alpha, \beta)} \\
    \Delta &= \sum_{A_1=1}^{n-k+1} \sum_{A_2=A_1+1}^{n-k+2} \cdots \sum_{A_n=A_{n-k+1}+1}^{n} \text{ and } k = 1, 2, \ldots, n
\end{align*}
\]

(1)

\[
L = L(x|\alpha, \beta) \propto \alpha^{-n} \beta^n \prod_{j=1}^{n} x_j^{\alpha-1} e^{-\beta \sum_{j} x_j} \sum_{\Delta} \prod_{r=1}^{k} \left( \frac{e^{\beta(x_{\hat{r}}-x_j)}}{x_{\hat{r}}^{\alpha-1}} \right)
\]

(2)

\[
\propto \alpha^{-n} \beta^n \prod_{j=1}^{n} x_j^{\alpha-1} e^{-\beta \sum_{j} x_j} \sum_{\Delta} Z(x_{\hat{r}}; \alpha, \beta) \sum_{\Delta} \prod_{r=1}^{k} \left( \frac{e^{\beta(x_{\hat{r}}-x_j)}}{x_{\hat{r}}^{\alpha-1}} \right)
\]

(3)

where \(Z(x_{\hat{r}}; \alpha, \beta) = \prod_{r=1}^{k} \left( \frac{e^{\beta(x_{\hat{r}}-x_j)}}{x_{\hat{r}}^{\alpha-1}} \right)\)

3. Estimation procedures

In this section, we consider the estimation procedure for the model defined in Equation (3) using maximum likelihood (ML) and Bayesian techniques.

3.1. Maximum Likelihood Estimation (MLE)

The log-likelihood function of Equation (3) is given by

\[
\ln L = (n - k) \ln \alpha + n \ln \beta + (\alpha - 1) \ln \sum_j x_j - \beta \sum_j x_j^{\alpha} + \ln Z(x_{\hat{r}}; \alpha, \beta)
\]

(4)

To obtain the MLE of \((\alpha, \beta)\), say \((\hat{\alpha}, \hat{\beta})\), one can solve the following equations

\[
\frac{\partial \ln L}{\partial \alpha} = \frac{n - k}{\alpha} - \ln \sum_j x_j - \beta \sum_j x_j^{\alpha} \ln x_j + \frac{\sum_{\Delta} Z(x_{\hat{r}}; \alpha, \beta)}{\sum_{\Delta} Z(x_{\hat{r}}; \alpha, \beta)}
\]

(5)

\[
\frac{\partial \ln L}{\partial \beta} = \frac{n}{\beta} - \beta \sum_j x_j^{\alpha} + \beta \sum_{\Delta} Z(x_{\hat{r}}; \alpha, \beta)
\]

(6)

Equations (5) and (6) can be solved for \(\hat{\alpha}\) and \(\hat{\beta}\) using any numerical iterative procedure. Since the MLEs of \(\alpha\) and \(\beta\) are not in the closed forms, therefore, it is not possible to derive their exact distributions. Thus, using large sample theory of MLE, the asymptotic sampling distribution of \((\hat{\alpha} - \alpha, \hat{\beta} - \beta)'\) is \(N_{2}(0, \Delta^{-1})\) where \(\Delta\) is the observed Fisher information matrix. The elements of \(\Delta\) are given by

\[
\begin{align*}
    \Delta_{11} &= -\frac{\partial^2 \ln L}{\partial \alpha^2} \bigg|_{\alpha = \hat{\alpha}, \beta = \hat{\beta}} \\
    \Delta_{12} &= \Delta_{21} = -\frac{\partial^2 \ln L}{\partial \alpha \partial \beta} \bigg|_{\alpha = \hat{\alpha}, \beta = \hat{\beta}} \\
    \Delta_{22} &= -\frac{\partial^2 \ln L}{\partial \beta^2} \bigg|_{\alpha = \hat{\alpha}, \beta = \hat{\beta}}
\end{align*}
\]
Here,

\[
\frac{\partial^2 \ln L}{\partial \alpha^2} = -\frac{(n-k)}{\alpha^2} - \beta \sum_j x_j^2 (\ln x_j)^2 + \left[ \sum_{\Delta} Z_\alpha(x_\alpha; \alpha, \beta) \sum_{\Delta} Z'_\alpha(x_\alpha; \alpha, \beta) - \left( \sum_{\Delta} Z'(x_\alpha; \alpha, \beta) \right)^2 \right] \\
\frac{\partial^2 \ln L}{\partial \alpha \partial \beta} = \frac{\partial \ln L}{\partial \beta \partial \alpha} = \sum_j x_j^2 \ln x_j + \left[ \sum_{\Delta} Z_\alpha(x_\alpha; \alpha, \beta) \sum_{\Delta} Z'_\beta(x_\alpha; \alpha, \beta) - \left( \sum_{\Delta} Z'(x_\alpha; \alpha, \beta) \right)^2 \right] \\
\frac{\partial^2 \ln L}{\partial \beta^2} = -\frac{n}{\beta^2} + \left[ \sum_{\Delta} Z_\beta(x_\beta; \alpha, \beta) \sum_{\Delta} Z'_{\beta}(x_\beta; \alpha, \beta) - \left( \sum_{\Delta} Z'(x_\beta; \alpha, \beta) \right)^2 \right]
\]

(7)
(8)
(9)

The respective asymptotic 100(1 - γ)% confidence intervals (C.I.) for α and β are \( \hat{\alpha} \pm z_{\gamma/2} \sqrt{V(\hat{\alpha})} \) and \( \hat{\beta} \pm z_{\gamma/2} \sqrt{V(\hat{\beta})} \) where \( V(\hat{\alpha}) \) and \( V(\hat{\beta}) \) are the variances of \( \hat{\alpha} \) and \( \hat{\beta} \), which can be obtained using Fisher information matrix. Here, \( z_{\gamma/2} \) is the upper 100 × (γ/2)th percentile of a standard normal distribution.

### 3.2. Bayesian estimation using Gibbs sampler

In classical frequentist approach, the parameters are considered to be fixed. However, in many real-life situations represented by life time models, the parameters cannot be treated as constant (Ibrahim, Chen, & Sinha, 2001; Martz & Waller, 1982; Singpurwalla, 2006) throughout the life testing period. In view of this, we propose Bayesian estimation procedure by assuming the following independent gamma priors for α and β:

\[
g_1(\alpha) = \frac{v_1}{\Gamma(\delta_1)} \alpha^{v_1-1} \exp(-v_1 \alpha); \quad (v_1, \delta_1, \alpha > 0)
\]

(10)

and

\[
g_2(\beta) = \frac{v_2}{\Gamma(\delta_2)} \beta^{v_2-1} \exp(-v_2 \beta); \quad (v_2, \delta_2, \beta > 0)
\]

(11)

Here, the hyper parameters \( v_1, \delta_1, v_2, \) and \( \delta_2 \) are assumed to be known real numbers. Using likelihood function in (3) and prior distributions in (10) and (11), the joint posterior distribution of \( \alpha, \beta \) and the data are given by

\[
\phi(\alpha, \beta | x) = L(x | \alpha, \beta)g_1(\alpha)g_2(\beta)
\]

(12)

It can be seen that the above expression cannot be obtained in nice closed form and one needs numerical approximation. Here, we use Gibbs sampler, an MCMC method, proposed by Geman and Geman (1984). It allows us to generate observations from the conditional distribution of each of the parameters using the current values of the given parameters (see Appendix A). MCMC is a class of methods in which one can simulate draws that are slightly dependent and approximately from the
posterior distribution. By means of this procedure, our aim is to get the ergodic chains of the parameters which are irreducible, aperiodic, and positive recurrent. For implementing Gibbs sampling procedure, the full conditional posterior distributions of $\alpha$ and $\beta$ are

$$\pi_1(\alpha|\beta, x) \propto \alpha^{n+i+k-1} \prod_{j=1}^{n} \lambda_j \sum_{A} \sum_{r=1}^{k} \left[ e^{\beta(x_{r}^{A} - x_{i})} \right]$$

(13)

$$\pi_2(\beta|\alpha, x) \propto \beta^{n+i+k-1} e^{-\beta \sum_{A} \sum_{r=1}^{k} \left[ \frac{A r}{e^{\frac{\alpha}{\alpha}} \lambda_i} \right]}$$

(14)

On putting in (13) and (14), one gets the respective conditional posterior distributions of the parameters under the assumptions of Jeffrey’s prior.

4. Simulation study

In this section, we conduct a simulation study to compare the performance of the classical and Bayesian methods of estimation. The SEs of the estimates and widths of the confidence/highest posterior density (HPD) credible intervals are used for comparison purpose. Assuming the following values of the model parameters $\alpha = 0.5$, $\beta = 2.5$ and $k = 2$, we generated the six sets of data containing, respectively, $n = 25$, 50, 100, 200, 300, and 500 observations (i.e. small, moderate, and large samples) and based on these data-sets, the MLEs and Bayes estimates for the parameters have been obtained. For Bayesian estimation of the parameters, we generated 10,000 realizations of the Markov chains using Gibbs and Metropolis–Hastings algorithms. The convergence of the sequences of parameters for their stationary distributions has been checked through different starting values. It was observed that after some burn-in periods, all the Markov chains reached their stationary condition. The MCMC runs of the parameters $\alpha$ and $\beta$ are plotted in Figures 1 and 3, which show fine mixing of the chains. We have also drawn the posterior densities of $\alpha$ and $\beta$ and found that both the distributions are uni-model (Figures 2 and 4). The results of the simulation study have been summarized in Tables 1–3. For all the numerical computations, the programs are developed in R-software (2016) and are available with the authors.

From the simulation results in Tables 1–3 and various plots in Figures 1–4, it is observed that both the methods of estimation considered in the study are precisely estimating the model parameters. The amount of errors in estimating all the parameters decreases as the sample size increases which shows increased precision in the estimation of the model parameters. Bayes estimation with
Figure 3. Plot of generated $\beta$ vs. iteration of the MCMC algorithm.

Figure 4. Posterior density of $\beta$.

Table 1. Maximum likelihood estimates with their SEs and widths of the 95% CI intervals of the parameters for fixed $k = 2, \alpha = 0.5, \beta = 2.5$, and varying sample sizes ($n$)

| Sample size | $\hat{\alpha}$ | SE($\hat{\alpha}$) | Width_CI $\hat{\alpha}$ | $\hat{\beta}$ | SE($\hat{\beta}$) | Width_CI $\hat{\beta}$ |
|-------------|----------------|-------------------|-------------------------|----------------|----------------|-------------------------|
| 25          | 0.27670        | 0.06687           | 0.26215                 | 2.57387        | 0.58935        | 2.31028                 |
| 50          | 0.27550        | 0.04008           | 0.15714                 | 2.88781        | 0.41051        | 1.60921                 |
| 100         | 0.31402        | 0.03186           | 0.12489                 | 2.57865        | 0.25826        | 1.01240                 |
| 200         | 0.31501        | 0.02244           | 0.08796                 | 2.60693        | 0.18450        | 0.72324                 |
| 300         | 0.32075        | 0.01861           | 0.07295                 | 2.54730        | 0.14716        | 0.57689                 |
| 500         | 0.32183        | 0.01443           | 0.05659                 | 2.57639        | 0.11527        | 0.45186                 |

Table 2. Jeffrey Bayes estimates with their PSEs and widths of the 95% HPD intervals of the parameters for fixed $k = 2, \alpha = 0.5, \beta = 2.5$, and varying sample sizes ($n$)

| Sample size | $\alpha^*$ | PSE($\alpha^*$) | HPD_CI $\alpha^*$ | $\beta^*$ | PSE($\beta^*$) | HPD_CI $\beta^*$ |
|-------------|------------|-----------------|-------------------|-----------|----------------|-------------------|
| 25          | 0.62243    | 0.06271         | 0.25181           | 3.14879   | 0.57139        | 2.27456           |
| 50          | 0.60956    | 0.03927         | 0.13131           | 2.84431   | 0.39278        | 1.56738           |
| 100         | 0.55108    | 0.03079         | 0.11798           | 2.73362   | 0.24754        | 0.92167           |
| 200         | 0.54207    | 0.02207         | 0.08713           | 2.40890   | 0.16984        | 0.69547           |
| 300         | 0.53447    | 0.01751         | 0.07278           | 2.64109   | 0.14104        | 0.56754           |
| 500         | 0.49898    | 0.014372        | 0.05659           | 2.54808   | 0.11514        | 0.45279           |

Table 3. Gamma Bayes estimates with their PSEs and widths of the 95% HPD intervals of the parameters for fixed $k = 2, \alpha = 0.5, \beta = 2.5$, and varying sample sizes ($n$)

| Sample size | $\alpha^*$ | PSE($\alpha^*$) | HPD_CI $\alpha^*$ | $\beta^*$ | PSE($\beta^*$) | HPD_CI $\beta^*$ |
|-------------|------------|-----------------|-------------------|-----------|----------------|-------------------|
| 25          | 0.61920    | 0.06728         | 0.24170           | 2.84776   | 0.51672        | 2.12480           |
| 50          | 0.57180    | 0.03175         | 0.13617           | 2.72116   | 0.32084        | 1.81712           |
| 100         | 0.54010    | 0.02048         | 0.10180           | 2.68005   | 0.21615        | 0.99721           |
| 200         | 0.52665    | 0.01576         | 0.06775           | 2.61119   | 0.11125        | 0.55452           |
| 300         | 0.51811    | 0.01275         | 0.06110           | 2.51557   | 0.09612        | 0.48675           |
| 500         | 0.50103    | 0.01005         | 0.04335           | 2.51905   | 0.07725        | 0.39180           |
gamma prior provides more precise estimates as compared to Bayes estimates with Jeffrey’s prior as well as ML estimates. Also Bayes estimates with Jeffrey priors perform better than the ML estimates even though they are quite similar when sample size becomes large. The size of CI/HPD intervals of the parameters tends to condense as the sample size increases.

5. Real data analysis
The following is the life distribution (in units of 100) of 20 electronic tubes:

|     |    |    | Log L(x; \hat{\alpha}, \hat{\beta}) |
|-----|----|----|----------------------------------|
| 0.1415 | 0.5937 | 2.3467 | 13.1356 | 3.5681 |
| 0.3484 | 1.1045 | 2.4651 | 3.2259 | 3.7287 |
| 0.3994 | 1.7323 | 2.6155 | 3.4177 | 9.2817 |
| 0.4174 | 1.8348 | 2.7425 | 3.5551 | 9.3208 |

The data are taken from Dixit and Nooghabi (2011). It is to be noted that the number of outliers \( k \) is unknown, so we estimate the model parameter for different values of \( k \). Thus, \( k \) can be chosen by observing the log-likelihood function with respect to \( k \). The estimates of \( \alpha \) and \( \beta \) and corresponding log-likelihood function for different values of \( k \) are shown in Table 4. From Table 4, it is observed that the log-likelihood function with respect to \( k \) is maximized for \( k = 2 \).

### Table 4. ML estimates and log-likelihood function for varying values of \( k \)

| \( k \) | \( \hat{\alpha} \) | \( \hat{\beta} \) | Log L(x; \hat{\alpha}, \hat{\beta}) |
|-----|-----|-----|-------------------------------|
| 1   | 1.15074 | 0.29264 | 34.92786 |
| 2   | 1.15904 | 0.29358 | 35.06586 |
| 3   | 1.15471 | 0.29317 | 34.11783 |

### References

Barnett, V. A., & Lewis, T. (1994). Outliers in statistical data. Chichester: John Wiley and Sons.

Chaturvedi, A., Pati, M., & Tomer, S. K. (2014). Robust Bayesian analysis of Weibull failure model. Metron, 72, 77–95.

Deiri, E. (2011). Estimations of the exponentiated gamma distribution parameters with presence of two outliers. International Journal of Academic Research, 3, 846–852.

Deiri, E. (2012). Classical estimator of parameters of the gamma case with presence of two outliers. Communications in Statistics-Simulation and Computation, 41, 590–597. doi:10.1080/03610918.2011.598993

Dixit, U. J. (1994). Bayesian approach to prediction in the presence of outliers for Weibull distribution. Metrika, 41, 127–136.

Dixit, U. J., Moore, K. L., & Barnett, V. (1996). On the estimation of the power of the scale parameter of the exponential distribution in the presence of outliers generated from uniform distribution. Metron, 54, 201–211.

Dixit, U. J., & Nasiri, P. F. (2001). Estimation of parameters of the exponential distribution in the presence of outliers generated from uniform distribution. Metron, 49, 187–198.

Dixit, U. J., & Nooghabi, J. M. (2011). Efficient estimation of the parameters of the Pareto distribution in the presence of outliers. Communications of the Korean Statistical Society, 18, 817–835.

Geman, S., & Geman, D. (1984). Stochastic relaxation, gibbs distribution, and the Bayesian restoration of images. IEEE Transaction on Pattern Analysis and Machine Intelligence, 6, 721–741.

Hassan, A. S., Elsherpieny, E. A., & Shalaby, R. M. (2013). On the estimation of P(Y < X < Z) for Weibull distribution in the presence of k outliers. International Journal of Engineering Research and Applications, 3, 1727–1733. Retrieved from http://www.ijera.com/papers/Vol3 issue6/JZ3617271733.pdf

Ibrahim, J. G., Chen, M.-H., & Sinha, D. (2001). Bayesian survival analysis. New York, NY: Springer-Verlag.

Kale, B. K., & Sinha, S. K. (1971). Estimation of expected life in the presence of an outlier observation. Technometrics, 13, 755–759.

Khokan, M. R., Bari, W., & Khan, J. A. (2013). Weighted maximum likelihood approach for robust estimation: Weibull model (in the presence of outliers). Dhaka University Journal of Science, 61, 153–156.
Appendix A

Gibbs algorithm

1. Initialize $\beta_0$ as starting values of $\beta$.
2. For given $\beta_0$, generate $\alpha_1$ from $\pi_1(\alpha|\beta, \mathbf{x})$, a complex distribution given in (13).
3. For given $\alpha_1$, generate $\beta_1$ from $\pi_2(\beta|\alpha, \mathbf{x})$, a complex distribution given in (14).
4. Set $\beta_0 = \beta_1$, and repeat steps 2 to 3, $M$-times and record the sequence, $(\alpha, \beta)$ after N burn-in iterations have occurred to abolish the effects of the starting values.
5. Bayes estimates $(\alpha^*, \beta^*)$ of $(\alpha, \beta)$ are taken as the means of the generated draws of $(\alpha, \beta)$.

Note that the sampling in step 3 is done using following Metropolis–Hastings algorithm.

Metropolis–Hasting algorithm

1. Start with any arbitrary value of $\mathbf{x}$ say $x_0$ from the support of the target distribution.
2. Generate a proposal point $(x_{prop})$ from the proposal density $q(x_{prop}|x_{i-1})$, i.e. the probability of returning a value of $x_{prop}$ given a previous value of $x_{i-1}$. Calculate the ratio at the proposal point $(x_{prop})$ and current $(x_{i-1})$ as:

$$\rho = \frac{f(x_{prop})q(x_{prop}|x_{i-1})}{f(x_{i-1})q(x_{i-1}|x_{prop})}$$

3. Generate $U$ from uniform on $(0, 1)$ and take $Z = \log U$.
4. If $Z < \rho$, accept the move, i.e. $x_{prop}$, and set $x_0 = x_{prop}$ and return to step 1. Otherwise, reject it and return to step 2.

References

Lieblein, J., & Zelen, M. (1956). Statistical investigation of the fatigue life of deep groove ball bearings. *Journal of Research of the National Bureau of Standards, 57*, 273–315.

Mann, N. (1968). Results on statistical estimation and hypotheses testing with application to the Weibull and extreme value distribution. Dayton, OH: Aerospace Research Laboratories, Wright-Patterson Air Force Base.

Martz, H. F., & Waller, R. A. (1982). *Bayesian reliability analysis*. New York, NY: John Wiley.

Nasiri, P., & Pazira, H. (2010). Bayesian approach on the generalized exponential distribution in the presence of outliers. *Journal of Statistical Theory and Practice, 4*, 453–475.

Pettit, L. I. (1988). Bayes methods for outliers in exponential samples. *Journal of the Royal Statistical Society: Series B (Methodological), 50*, 371–380.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Singpurwalla, N. D. (2006). *Reliability and risk: A Bayesian perspective*. Chichester: John Wiley.

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics, 18*, 293–297.
