Bayesian Estimation of Pareto Distribution Under Failure-Censored Step-Stress Life Test Model
Ali A. Ismail, Bayram El-Sayed

Bayesian Estimation of Pareto Distribution Under Failure-Censored Step-Stress Life Test Model

Reference

ABSTRACT
In this paper, both maximum likelihood and Bayesian estimators for a partially accelerated step-stress life test model are considered using type II censored data from Pareto distribution of the second kind. The posterior means and posterior variances are obtained under the squared error (SE) loss function using Lindley’s approximation procedure. The maximum likelihood estimators and analogous Bayes estimators are compared in terms of their mean-square errors based on simulated samples from Pareto distribution.

Keywords
reliability, partial acceleration, step-stress test, Pareto distribution, maximum likelihood estimation, Bayesian estimation, Lindley’s approximation, non-informative priors, posterior mean, posterior variance, type II censoring

Nomenclature
ALT = accelerated life test(s)
BE = Bayes estimate/estimator
MLE = maximum likelihood estimate/estimator
MSE = mean-square error
n = number of step-stress test units (total sample size)
n_c = number of censored units (n_c = n - n_a - n_u)
NIP = non-informative prior
n_a, n_u = numbers of test units failed at use and accelerated conditions, respectively
PALT = partially accelerated life test(s)
PD = Pareto distribution
pdf = probability density function
\( r = n_u + n_d \)
SE = squared error
SSPALT = step-stress PALT
\( T = \text{total lifetime of an item in a SSPALT} \)
\( y_1 \leq \ldots \leq y_{n(u)} \leq \tau \)
\( y_{(1)} \leq \ldots \leq y_{(n)} = \text{ordered failure times} \)
\( y_i = \text{observed value of the total lifetime} \)
\( Y_i = \text{of item } i, i = 1, \ldots, n \)
\( y_{(r)} = \text{time of the } r\text{th failure in a SSPALT} \)
\( x = \text{shape parameter } (x > 0) \)
\( \beta = \text{acceleration factor } (\beta > 1) \)
\( \theta = \text{scale parameter } (\theta > 0) \)
\( \tau = \text{stress change time in a step} \)
PALT (\( \tau < y_{(r)} \))
\( \lambda = \text{an implied maximum likelihood estimator} \)
\( \downarrow(\cdot) = \text{evaluated at } (\cdot) \)

Introduction

Accelerated reliability tests are commonly used as time- and cost-efficient industrial experiments. They consist of testing items under levels of stress higher than the normal working conditions. The main interest of such experiments is to obtain measures of the reliability of the devices under the normal working conditions via data acquired under stress levels. As indicated by Pathak et al. [1], the model of acceleration is chosen so that the relationship between the parameters of the failure distribution and the accelerated stress conditions is known. These relationships are usually derived from an analysis of the physical mechanisms of failure of the component. The tests performed under accelerated stress conditions are called fully accelerated life tests (FALT or, simply, ALT). Interested readers can refer to Meeker and Escobar [2] and Nelson [3], which are two comprehensive sources for ALT.

Sometimes, such relationships may not be known or cannot be assumed. So, in this case, ALT cannot be applied to predict products’ reliability. Instead, another type of test, called the partially accelerated life test (PALT), is used according to the proposed model by DeGroot and Goel [4].

As Nelson [3] indicates, the stress can be applied in various ways, commonly used method is step-stress. Under step-stress PALT, a test item is first run at use condition and, if it does not fail for a specified time, then it is run at accelerated condition until failure occurs or the test is terminated. Accelerated test stresses involve higher than usual temperature, voltage, pressure, load, humidity…, etc., or some combination of them. The objective of a PALT is to collect more failure data in a limited time without necessarily using high stresses to all test units. So, PALT combines both ordinary and accelerated life tests.

Most of literature performed on PALT considered the classical approach to estimate the parameters of interest, for example, see Goel [5], Bhattacharyya and Soejoeti [6], Bai and Chung [7], Bai et al. [8], Attia et al. [9], Abdel-Ghaly et al. [10], Madi [11], Abdel-Ghani [12], Aly and Ismail [13], Ismail and Sarhan [14], Ismail and Aly [15], Ismail and Abu-Youssef [16], and Ismail [17,18]. From the Bayesian point of view, few of studies have been considered on PALT. Goel [5] used the Bayesian approach for estimating the acceleration factor and the parameters in the case of step-stress PALT (SSPALT) with complete sampling for items having exponential and uniform distributions. DeGroot and Goel [4] investigated the optimal Bayesian design of a PALT in the case of the exponential distribution under complete sampling. Abdel-Ghani [12] considered the Bayesian approach to estimate the parameters of Weibull distribution in SSPALT with censoring. Ismail [19] considered the Bayesian approach to estimate the parameters of Gompertz distribution with time censoring.

The objective of this paper is to apply a Bayesian analysis of SSPALT considering two-parameter Pareto distribution with type II censoring assuming a squared error loss function. The Bayes estimators (BEs) of the acceleration factor and the distribution parameters are derived and compared with the maximum likelihood estimators (MLEs) counterparts by Monte Carlo simulations. To make the comparison more meaningful, the non-informative priors on both shape and scale parameters are considered.

The main contribution of this paper lies in studying the case of step-stress partially accelerated life tests (SSPALT) under type II censoring assuming the Pareto distribution within a Bayesian framework. It is a very important case that has not been investigated before.

This case is very important because of the following advantages. The advantages of the Pareto distribution over other distributions used in step-stress PALT can be presented as follows:

- The simple mathematical structure of the Pareto distribution enables it to be used effectively for modeling lifetime data with a larger tail weight or a heavy-tailed distribution.
- The interesting point is that the Pareto hazard rate function is a decreasing function. Another indication of heavy tail weight is that the distribution has a decreasing hazard rate function. Any distribution having an increasing hazard rate function is a light-tailed distribution such as Weibull. The Pareto is heavier in the tails.
- It observed that the Pareto distribution can be used quite effectively to analyze skewed data sets over the more popular Weibull distribution. In many situations, the Pareto
distribution provides a better fit than a Weibull distribution for modeling lifetime data with a larger tail weight.

- The Pareto distribution has many applications in modeling incomes and survival times. Because it is a heavy-tailed distribution, it is a good candidate for modeling survival times above a theoretical value. So the Pareto distribution is useful in modeling extreme value data, because of its long-tail feature.
- Within a Bayesian framework, the Pareto family has the potential for modeling reliability and life data.
- Moreover, type II censoring has received considerable interest among the statisticians. It is more effective and more practical than type I censoring. Type II censoring enables the experimenter to obtain a sufficient number of failure times to make a good statistical inference.

The rest of the paper is organized as follows. In the next section, the model and test methods are described. Approximate BEs of the parameters under consideration are derived in the following section. Next, BEs derived previously are obtained numerically using Lindley’s approximation and compared with the MLEs. Finally, we present conclusions.

The Model and Test Method

THE PARETO DISTRIBUTION AS A LIFETIME MODEL

In this paper, the two-parameter Pareto distribution of the second kind is considered as a lifetime model. The Pareto distribution was introduced by Pareto [20] as a model for the distribution of income. In recent years, its models in several different forms have been studied by many authors: Davis and Feldstein [21], Cohen and Whitten [22], and Grimshaw [23], among others. The Pareto distribution of the second kind is also known as Lomax or Pearson’s type VI distribution (see Johnson et al. [24]). It has been found to be a good model in biomedical problems, such as survival time following a heart transplant (Bain and Engelhardt [25]). Using the Pareto distribution, Dyer [26] studied annual wage data of production line workers in a large industrial firm. Lomax [27] used this distribution in the analysis of business failure data. The length of wire between flaws also follows a Pareto distribution (Bain and Engelhardt [25]). Because the Pareto distribution has a decreasing hazard or failure rate, it has often been used to model incomes and survival times (Howlader and Hossain [28]).

The probability density function of the Pareto distribution of the second kind is given by

\[
    f_T(t; \theta, \alpha) = \frac{\theta^\alpha}{(\theta + t)^{\alpha+1}}, \quad t > 0, \theta > 0, \alpha > 0
\]

The survival function takes the form

\[
    R(t) = \frac{\theta^\alpha}{(\theta + t)^\alpha}, \quad t > 0
\]

and the corresponding failure rate function is

\[
    h(t) = \frac{\alpha}{\theta + t}
\]

According to McCune and McCune [29], the Pareto distribution has classically been used in economic studies of income, size of cities and firms, service time in queuing systems, and so on. Also, it has been used in connection with reliability theory and survival analysis (see Davis and Feldstein [21]).

THE PARETO DISTRIBUTION AS A MIXTURE

The Pareto distribution is a mixture of exponential distributions with gamma mixing weights. We now elaborate more on this point. Through looking at various properties of the Pareto distribution, we also demonstrate that the Pareto distribution is a heavy-tailed distribution. The Pareto distribution is a great way to open up a discussion on heavy-tailed distribution. When a distribution significantly puts more probability on larger values, the distribution is said to be a heavy-tailed distribution (or said to have a larger tail weight). The Pareto distribution is a handy example.

The Pareto probability density function (pdf) indicated above can be obtained by mixing exponential distributions using gamma distributions as weights. Suppose that \( T \) follows an exponential distribution (conditional on a parameter value \( \beta \)). The following is the conditional pdf of \( T \):

\[
    f_{T|\beta}(t|\beta) = \beta e^{-\beta t}, \quad t > 0
\]

There is uncertainty in the parameter, which can be viewed as a random variable \( \beta \). Suppose that \( \beta \) follows a gamma distribution with scale parameter \( \theta \) and shape parameter \( \alpha \). The following is the pdf of \( \beta \):

\[
    f_\beta(\beta) = \frac{\theta^\alpha}{\Gamma(\alpha)} \beta^{\alpha-1} e^{-\theta \beta}, \quad \beta > 0
\]

The unconditional pdf of \( T \) is the weighted average of the conditional pdf with the gamma pdf as weight:

\[
    f_T(t) = \int_0^\infty f_{T|\beta}(t|\beta) f_\beta(\beta) d\beta
\]

\[
    = \int_0^\infty \left[ \beta e^{-\beta t} \frac{\theta^\alpha}{\Gamma(\alpha)} \beta^{\alpha-1} e^{-\theta \beta} \right] d\beta
\]

\[
    = \int_0^\infty \frac{\theta^\alpha}{\Gamma(\alpha)} \beta^{\alpha-1} e^{-\theta \beta(1+t)} d\beta
\]

\[
    = \frac{\theta^\alpha}{\Gamma(\alpha)} \Gamma(\alpha+1) \int_0^\infty \frac{(\theta + t)^{\alpha+1}}{(\alpha + 1)^{\alpha+1}} \beta^{\alpha+1-1} e^{-\theta \beta(1+t)} d\beta
\]

\[
    = \frac{\theta^\alpha}{(\theta + t)^{\alpha+1}}
\]

which is the Pareto pdf as indicated above.
It is observed that the exponential distribution is considered a light-tailed distribution. Yet mixing exponentials produces the heavy-tailed Pareto distribution. Mixture distributions tend to be heavy tailed and again the Pareto distribution is a handy example.

**TEST METHOD**

**Basic Assumptions**

1. Two stress levels, \( x_1 \) and \( x_2 \) (design and high), are used.
2. For any level of stress, the life distribution of the test unit is Pareto.
3. The total lifetime \( y \) of an item is as follows:
   \[
   Y = \begin{cases} 
   T, & \text{if } T \leq \tau \\
   \tau + \beta^{-1}(T-\tau), & \text{if } T > \tau 
   \end{cases}
   \]
   where:
   \( T \) = the lifetime of an item at use condition.
   This model is called the tampered random variable (TRV) model. It was proposed by DeGroot and Goel [4].
4. The life distribution of the test item is assumed to follow Pareto distribution with scale and shape parameters \( \theta \) and \( \alpha \), respectively.
5. The lifetimes \( Y_1, \ldots, Y_n \) of the \( n \) test items are independent and identically distributed random variables (i.i.d.r.v’s).

**Test Procedure**

1. Each of the \( n \) test items is first run at use condition.
2. If it does not fail at use condition by a pre-specified time \( \tau \), then it is put on accelerated condition and run until it fails or the test is terminated.

The lifetime of the test item is assumed to follow Pareto distribution with scale and shape parameters \( \theta \) and \( \alpha \), respectively. Therefore, the probability density function of total lifetime \( Y \) of an item under SSPALT is given by

\[
 f(y) = \begin{cases} 
 0, & y \leq 0 \\
 f_1(y) & 0 < y \leq \tau \\
 f_2(y) & y > \tau 
\end{cases}
\]

where:
\( \theta > 0 \),
\( \alpha > 0 \), and
\( f_2(y) \) is obtained by the transformation-variable technique using \( f_1(y) \) and the model given in Eq 4.

**Bayesian Estimation**

In this section, the SE loss function is considered. Under SE loss function, the Bayes estimator of a parameter is its posterior expectation. The Bayes estimators cannot be expressed in explicit forms. Approximate Bayes estimators will be obtained under the assumption of non-informative priors using Lindley’s approximation.

In most of the applied problems, the information about the parameters is available in an independently (see Basu et al. [30]). Thus, here it is assumed that the parameters are independent a priori and let the non-informative prior (NIP) for each parameter be represented by the limiting form of the appropriate natural conjugate prior.

It follows that an NIP for the acceleration factor \( \beta \) is given by

\[
\pi_1(\beta) \propto \beta^{-1}, \quad \beta > 1
\]

Also, the NIPs for the scale parameter \( \theta \) and the shape parameter \( \alpha \) are, respectively,

\[
\pi_2(\theta) \propto \theta^{-1}, \quad \theta > 0 \quad \text{and} \quad \pi_3(\alpha) \propto \alpha^{-1}, \quad \alpha > 0
\]

Therefore, the joint NIP of the three parameters can be expressed by

\[
\pi(\beta, \theta, \alpha) \propto (\beta \theta \alpha)^{-1}, \quad \beta > 1, \theta > 0, \quad \alpha > 0
\]

The observed values of the total lifetime \( Y \) are given by

\[
y_{(1)} \leq \cdots \leq y_{(n_u)} \leq \tau \leq y_{(n_u+1)} \leq \cdots \leq y_{(n_u+n_a)} \leq y_{(r)}
\]

where:
\( n_u \) = the number of items failed at use condition, and
\( n_a \) = the number of items failed at accelerated condition.

Let \( \delta_{1i} \) and \( \delta_{2i} \) be indicator functions such that \( \delta_{1i} \sim 1(Y_i \leq \tau), \delta_{2i} \sim 1(t < Y_i \leq y_{(i)}), \) where \( i = 1, \ldots, n. \) Because the total lifetimes \( y_1, \ldots, y_n \) of \( n \) items are independent and identically distributed random variables, then the total likelihood function for them is given by

\[
L(y|\beta, \theta, \alpha) = \prod_{i=1}^{n} \left[ \frac{\alpha \theta^2}{(\theta + y_i)^{\alpha+1}} \right]^{\delta_{1i}} \left[ \frac{\beta \theta^2}{(\theta + y_i + \beta(\tau - y_i))^{\alpha+1}} \right]^{\delta_{2i}}
\]

where:

\[
\tilde{\delta}_{1i} = 1 - \delta_{1i} \quad \text{and} \quad \tilde{\delta}_{2i} = 1 - \delta_{2i}
\]

Another form of the total likelihood function can be given by

\[
L(y|\beta, \theta, \alpha) = \prod_{i=1}^{n_u} \left[ \frac{\alpha \theta^2}{(\theta + y_i)^{\alpha+1}} \right] \cdot \prod_{i=1}^{n_a} \left[ \frac{\beta \theta^2}{(\theta + \tau + \beta(\tau - y_i))^{\alpha+1}} \right] \cdot \prod_{i=1}^{r} \left[ \frac{\theta^2}{(\theta + \tau + \beta(\tau - y_i))^{\alpha+1}} \right]
\]
where:

\[ n_c = n - n_u - n_a. \]

Forming the product of Eqs 6 and 8, the joint posterior density function of \( \beta, \theta, \) and \( z \) given the data can be written as

\[ g(\beta, \theta, z; y) \propto L(y|\beta, \theta, z) \cdot \Pi(\beta, \theta, z) \]

\[ \propto \frac{\beta^{n_u-1} \theta^{n-a-1} (\theta + \tau)^{-2}}{(\theta + \tau + \beta(y - \tau))^{n-y-n_u}} \cdot \left[ \prod_{i=n_u+1}^{n} \frac{1}{(\theta + \tau + \beta(y_i - \tau))^{n-y-n_u}} \right] \]

As mentioned earlier, under a squared error loss function, the Bayes estimator of a parameter is its posterior expectation. To obtain the posterior means and posterior variances of \( \beta, \theta, \) and \( z, \) non-tractable integrals will be met. It is not possible to obtain them analytically. The marginal posteriors are somewhat unwieldy and require a numerical integration that may not converge. Instead, an approximation attributed to Lindley [31] via an asymptotic expansion of the ratio of two non-tractable integrals is used to obtain approximate Bayes estimators. Lindley’s approximation is evaluated at the ML estimates of the model parameters.

Now, let \( \Theta \) be a set of parameters \( \{\Theta_1, \Theta_2, \ldots, \Theta_m\}, \) where \( m \) is the number of parameters, then the posterior expectation of an arbitrary function \( u(\Theta) \) can be asymptotically estimated by

\[ E(u(\Theta)) = \int_{\Theta} u(\Theta) \pi(\Theta)^{\frac{\partial \ln L(y|\Theta)}{\partial \Theta}} d\Theta \int \pi(\Theta)^{\frac{\partial \ln L(y|\Theta)}{\partial \Theta}} d\Theta 
\approx \left[ u + \frac{1}{2} \sum_{i,j} (u^{(2)}_{ij} + 2u^{(1)}_{ij} \rho^{(1)}_{ij}) \sigma_{ij} \right] + \frac{1}{2} \sum_{i,j,k} L^{(3)}_{ijk} \sigma_{ij} \sigma_{jk} u^{(1)}_{ik} \]

which is the Bayes estimator of \( u(\Theta) \) under a squared error loss function, where \( \pi(\Theta) \) is the prior distribution of \( \Theta, \) \( u \equiv u(\Theta), \) \( L \equiv L(\Theta) \) is the likelihood function, \( \rho \equiv \rho(\Theta) = \log \pi(\Theta), \) \( \sigma_{ij} \) are the elements of the inverse of the asymptotic Fisher’s information matrix of \( \beta, \theta, \) and \( z, \) and

\[ u^{(1)}_{ij} = \frac{\partial u}{\partial \Theta_i}, \quad u^{(2)}_{ij} = \frac{\partial^2 u}{\partial \Theta_i \partial \Theta_j}, \]

\[ \rho^{(1)}_{ij} = \frac{\partial \log \pi(\Theta)}{\partial \Theta_i}, \quad \text{and} \quad L^{(3)}_{ijk} = \frac{\partial^3 \ln L(y|\Theta)}{\partial \Theta_i \partial \Theta_j \partial \Theta_k}. \]

Such an approximation is easy to use and does not require innovative programming and extensive computer time. According to Green [32], the linear Bayes estimator in Eq 10 is a "very good and operational approximation for the ratio of multi-dimension integrals." As indicated by Sinha [33], it has led to many useful applications. However, if the domain of the parameters is a function of the parameters, Bayes estimators using Lindley’s rule are not obtainable unless the MLEs exist. The derivation of posterior means and posterior variances is shown in the Appendix.

Monte Carlo Simulation Study

A Monte Carlo simulation study is conducted for demonstrating and comparing the methods of ML and Bayes estimators, under the SE loss function. The posterior means and posterior variances of the three parameters \( \beta, \theta, \) and \( z \) are derived assuming the NIP for each parameter using SE loss function with type II censored data. Because the BEs of the model parameters cannot be obtained analytically, approximate BEs are obtained numerically using the approximation of Lindley. The behavior sampling of the approximate BEs are investigated and compared with the MLEs in terms of their variances and mean-squared errors (MSEs) for different sample sizes and different parameter values. The process is replicated 1000 times for each sample size and the average of estimates is computed. The results are reported in Tables 1 and 2.

### Table 1

<table>
<thead>
<tr>
<th>( n )</th>
<th>Parameter</th>
<th>Method</th>
<th>Estimate</th>
<th>MSE</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>( \beta )</td>
<td>ML</td>
<td>2.3144</td>
<td>0.0988</td>
<td>0.0531</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>1.7812</td>
<td>0.0837</td>
<td>0.0421</td>
</tr>
<tr>
<td>50</td>
<td>( \theta )</td>
<td>ML</td>
<td>1.0142</td>
<td>0.0566</td>
<td>0.0237</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>0.9107</td>
<td>0.0425</td>
<td>0.0117</td>
</tr>
<tr>
<td>75</td>
<td>( \beta )</td>
<td>ML</td>
<td>1.6952</td>
<td>0.0727</td>
<td>0.0345</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>1.6083</td>
<td>0.0555</td>
<td>0.0212</td>
</tr>
<tr>
<td>100</td>
<td>( \beta )</td>
<td>ML</td>
<td>1.5871</td>
<td>0.0516</td>
<td>0.0172</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>1.5311</td>
<td>0.0431</td>
<td>0.0062</td>
</tr>
<tr>
<td>25</td>
<td>( \theta )</td>
<td>ML</td>
<td>0.7619</td>
<td>0.0287</td>
<td>0.0054</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>0.7252</td>
<td>0.0221</td>
<td>0.0033</td>
</tr>
<tr>
<td>100</td>
<td>( \beta )</td>
<td>ML</td>
<td>1.5241</td>
<td>0.0098</td>
<td>0.0036</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>1.5162</td>
<td>0.0054</td>
<td>0.0029</td>
</tr>
</tbody>
</table>
TABLE 2  Average values of the MLEs and approximate BEs with associated estimated variances and MSEs when $\beta = 2.5$, $\theta = 2$, $x = 1.5$, $\tau = 5$, and $\gamma_{10} = 11$. 

<table>
<thead>
<tr>
<th>$n$</th>
<th>Parameter</th>
<th>Method</th>
<th>Estimate</th>
<th>MSE</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>$\beta$</td>
<td>ML</td>
<td>3.1688</td>
<td>0.0799</td>
<td>0.0403</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>2.7415</td>
<td>0.0555</td>
<td>0.0254</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>ML</td>
<td>2.2314</td>
<td>0.0634</td>
<td>0.0182</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>2.1622</td>
<td>0.0611</td>
<td>0.0147</td>
</tr>
<tr>
<td></td>
<td>$x$</td>
<td>ML</td>
<td>1.9255</td>
<td>0.0406</td>
<td>0.0075</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>1.8436</td>
<td>0.0333</td>
<td>0.0047</td>
</tr>
<tr>
<td>50</td>
<td>$\beta$</td>
<td>ML</td>
<td>2.6861</td>
<td>0.0521</td>
<td>0.0327</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>2.6391</td>
<td>0.0401</td>
<td>0.0207</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>ML</td>
<td>2.1172</td>
<td>0.0407</td>
<td>0.0071</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>2.0533</td>
<td>0.0388</td>
<td>0.0047</td>
</tr>
<tr>
<td></td>
<td>$x$</td>
<td>ML</td>
<td>1.7213</td>
<td>0.0281</td>
<td>0.0027</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>1.6542</td>
<td>0.0143</td>
<td>0.0022</td>
</tr>
<tr>
<td>75</td>
<td>$\beta$</td>
<td>ML</td>
<td>2.5713</td>
<td>0.0392</td>
<td>0.0187</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>2.5370</td>
<td>0.0235</td>
<td>0.0103</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>ML</td>
<td>2.0382</td>
<td>0.0254</td>
<td>0.0024</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>2.0161</td>
<td>0.0142</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td>$x$</td>
<td>ML</td>
<td>1.5544</td>
<td>0.0129</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>1.5217</td>
<td>0.0084</td>
<td>0.0012</td>
</tr>
<tr>
<td>100</td>
<td>$\beta$</td>
<td>ML</td>
<td>2.5211</td>
<td>0.0102</td>
<td>0.0069</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>2.5114</td>
<td>0.0067</td>
<td>0.0046</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>ML</td>
<td>2.0124</td>
<td>0.0083</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>2.0081</td>
<td>0.0056</td>
<td>0.0009</td>
</tr>
<tr>
<td></td>
<td>$x$</td>
<td>ML</td>
<td>1.5120</td>
<td>0.0041</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bayes</td>
<td>1.5032</td>
<td>0.0023</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

When we compare the MLEs with the approximate BEs using Lindley’s technique in terms of their variances and MSEs, it is observed that the approximate BEs perform better than the MLEs. That is, the approximate BEs become smaller variances and smaller MSEs as the sample size increases. These results coincide with the note of Achcar [34]. He said that the use of approximate Bayesian methods could be a good alternative for the usual asymptotically classical methods in accelerated life testing.

Conclusion

In this paper, the ML and Bayes estimations of the parameters of Pareto distribution and the acceleration factor have been obtained. The Bayes estimators have been obtained under the assumptions of squared error loss functions and non-informative priors. It has been observed that the Bayes estimators cannot be obtained analytically. Instead, Lindley’s approximation has been used to obtain the Bayesian estimates numerically. It has been seen that the approximation works very well even for small sample sizes. Also, it has been noted that Lindley’s method usually provides posterior variances smaller than the variances of the maximum likelihood estimators. That is, it gives better estimates, which is an advantage of this method. It can be said that the intrinsic appeal of that method can be expressed as being a sort of adjustment to the maximum likelihood approach to reduce variability. However, it has been also observed for very large sample sizes that the Bayesian estimates and the MLEs become closer in terms of MSEs and variances. That is, for very large sample sizes, the performances are so far similar as expected. But if we consider informative priors, then the performances of BEs will be much better than those of MLEs and then there is no need for comparisons. As a future work, a Bayesian analysis via another approximation such as the Laplace approximation method or Markov Chain Monte Carlo (MCMC) algorithm will be considered.

ACKNOWLEDGMENTS

This project was supported by King Saud University, Deanship of Scientific Research, College of Science Research Center.

APPENDIX

Here, there are three parameters in the model. That is, $m = 3$. Let the subscripts 1, 2, and 3 refer to $\beta$, $\theta$, and $x$, respectively. Therefore, the posterior means (BEs) of the three parameters can be expressed by

\[
\beta^* = E(\beta|y) = \left[ \beta - \left( \frac{\sigma_{11}}{\beta} + \frac{\sigma_{12}}{\theta} + \frac{\sigma_{13}}{x} \right) \right] \downarrow \hat{\beta}
\]

\[
\theta^* = E(\theta|y) = \left[ \theta - \left( \frac{\sigma_{21}}{\beta} + \frac{\sigma_{22}}{\theta} + \frac{\sigma_{23}}{x} \right) \right] \downarrow \hat{\theta}
\]

\[
x^* = E(x|y) = \left[ x - \left( \frac{\sigma_{31}}{\beta} + \frac{\sigma_{32}}{\theta} + \frac{\sigma_{33}}{x} \right) \right] \downarrow \hat{x}
\]

Thus, the posterior variances can be obtained by

\[
\text{Var}(\beta|y) = E(\beta^2|y) - (\beta^*)^2
\]

\[
= \sigma_{11} - \left( \frac{\sigma_{11}}{\beta} + \frac{\sigma_{12}}{\theta} + \frac{\sigma_{13}}{x} \right) ^2 \downarrow \hat{\beta}
\]

\[
\text{Var}(\theta|y) = E(\theta^2|y) - (\theta^*)^2
\]

\[
= \sigma_{22} - \left( \frac{\sigma_{21}}{\beta} + \frac{\sigma_{22}}{\theta} + \frac{\sigma_{23}}{x} \right) ^2 \downarrow \hat{\theta}
\]
\[ \text{Var}(z|y) = E(z^2|y) - (z')^2 \]
\[ = \sigma_{33} - \left[ \left( \frac{\sigma_{31}}{\beta} + \frac{\sigma_{32}}{\theta} + \frac{\sigma_{33}}{\tau} \right) \right]^2 - (1/2)(\sigma_{31}\Psi_1 + \sigma_{32}\Psi_2 + \sigma_{33}\Psi_3)^2 \]

where:
\[ \Psi_1 = \sum_{ij} \sigma_{ij} L^{(3)}_{ij} = \sigma_{11} L^{(3)}_{111} + 2\sigma_{12} L^{(3)}_{112} + 2\sigma_{13} L^{(3)}_{113} + 2\sigma_{21} L^{(3)}_{211} + 2\sigma_{22} L^{(3)}_{222} + 2\sigma_{23} L^{(3)}_{232} + 2\sigma_{31} L^{(3)}_{311} + 2\sigma_{32} L^{(3)}_{322} + 2\sigma_{33} L^{(3)}_{333} \]
\[ \Psi_2 = \sum_{ij} \sigma_{ij} L^{(3)}_{ij} = \sigma_{11} L^{(3)}_{111} + 2\sigma_{12} L^{(3)}_{112} + 2\sigma_{13} L^{(3)}_{113} + 2\sigma_{21} L^{(3)}_{211} + 2\sigma_{22} L^{(3)}_{222} + 2\sigma_{23} L^{(3)}_{232} + 2\sigma_{31} L^{(3)}_{311} + 2\sigma_{32} L^{(3)}_{322} + 2\sigma_{33} L^{(3)}_{333} \]

and
\[ \Psi_3 = \sum_{ij} \sigma_{ij} L^{(3)}_{ij} = \sigma_{11} L^{(3)}_{111} + 2\sigma_{12} L^{(3)}_{112} + 2\sigma_{13} L^{(3)}_{113} + 2\sigma_{21} L^{(3)}_{211} + 2\sigma_{22} L^{(3)}_{222} + 2\sigma_{23} L^{(3)}_{232} + 2\sigma_{31} L^{(3)}_{311} + 2\sigma_{32} L^{(3)}_{322} + 2\sigma_{33} L^{(3)}_{333} \]

To compute the posterior means and the posterior variances of the three parameters \( \beta, \theta, \) and \( \tau, \) both second and third derivatives of the natural logarithm of the likelihood function must be obtained.

The likelihood function is shown in Eq. 7. Its natural logarithm can be written as
\[ \ln L = \sum_{i=1}^{n} \left[ \delta_{1i} \ln \left( \frac{2\theta^2}{(\theta + y_i)^{x+1}} \right) \right] + \sum_{i=1}^{n} \left[ \delta_{2i} \ln \left( \frac{2\beta\theta^2}{(\theta + y_i + \beta(y_i - \tau))^{x+1}} \right) \right] + \sum_{i=1}^{n} \left[ \delta_{1i} \delta_{2i} \ln \left( \frac{\theta^2}{(\theta + y_i + \beta(y_i - \tau))^2} \right) \right] \]

Therefore, we have
\[ \ln L = (n_u \cdot n_u) \ln x + n \ln \theta + n_u \ln \beta \]
\[ - \left( \sum_{i=1}^{n} \delta_{1i} \ln (\theta + y_i) + \sum_{i=1}^{n} \delta_{2i} \ln (\theta + \tau + \beta(y_i - \tau)) \right) + (x + 1) \]

where:
\[ n_u = \sum_{i=1}^{n} \delta_{1i} \quad \text{and} \quad n_u = \sum_{i=1}^{n} \delta_{2i} \]

The second derivatives of \( \ln L \) with respect to \( \beta, \theta, \) and \( \tau \) are given by
\[ \frac{\partial^2 \ln L}{\partial \beta^2} = -\frac{n_u}{\beta} + \frac{(n_u \cdot n_u) \beta}{\psi_{\beta}} + (x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(x - \tau)^2}{\psi_i^2} \]

where:
\[ \psi_{\beta} = \theta + \tau + \beta(y_i - \tau) \]
\[ \psi_i = \theta + \tau + \beta(y_i - \tau) \]

and
\[ n_u = n_u + n_u \]

\[ \frac{\partial^2 \ln L}{\partial \theta^2} = \frac{n_u}{\theta} + \frac{(n_u \cdot n_u) \theta}{\psi_{\theta}} + (x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau)^2}{\psi_i^2} \]

\[ \frac{\partial^2 \ln L}{\partial \tau^2} = \frac{n_u}{\tau} + \frac{(n_u \cdot n_u) \tau}{\psi_{\tau}} + (x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau)^2}{\psi_i^2} \]

Now, the third derivatives of \( \ln L \) with respect to \( \beta, \theta, \) and \( \tau \) are as follows:
\[ \frac{\partial^3 \ln L}{\partial \beta^3} = 2n_u \frac{2(n_u - n_u - n_u) \beta(y_i - \tau)^3}{(\beta(y_i - \tau) + \theta + \tau)^3} - 2(x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau)^3}{(\beta(y_i - \tau) + \theta + \tau)^3} \]
\[ \frac{\partial^3 \ln L}{\partial \theta^3} = 2n_u \frac{2(n_u - n_u - n_u) \theta(y_i - \tau)^3}{(\beta(y_i - \tau) + \theta + \tau)^3} - 2(x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau)^3}{(\beta(y_i - \tau) + \theta + \tau)^3} \]
\[ \frac{\partial^3 \ln L}{\partial \tau^3} = \frac{n_u}{\tau} + \frac{(n_u \cdot n_u) \tau}{\psi_{\tau}} + (x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau)^3}{\psi_i^2} \]

\[ \frac{\partial^3 \ln L}{\partial \beta \partial \theta} = -2n_u \frac{(n_u - n_u - n_u) \beta(y_i - \tau)^2}{(\beta(y_i - \tau) + \theta + \tau)^3} + 2(x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau)^2}{(\beta(y_i - \tau) + \theta + \tau)^3} \]
\[ \frac{\partial^3 \ln L}{\partial \beta \partial \tau} = -2n_u \frac{(n_u - n_u - n_u) \beta(y_i - \tau)^2}{(\beta(y_i - \tau) + \theta + \tau)^3} + 2(x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau)^2}{(\beta(y_i - \tau) + \theta + \tau)^3} \]
\[ \frac{\partial^3 \ln L}{\partial \theta \partial \tau} = \frac{n_u}{\tau} + \frac{(n_u \cdot n_u) \tau}{\psi_{\tau}} + (x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau)^2}{\psi_i^2} \]
\[ \frac{\partial^3 \ln L}{\partial \beta \partial \theta \partial \tau} = \frac{n_u}{\beta} + \frac{(n_u \cdot n_u) \beta}{\psi_{\beta}} - \frac{(n_u \cdot n_u) \theta}{\psi_{\theta}} + (x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau) \beta(y_i - \tau)}{\psi_i^2} \]
\[ \frac{\partial^3 \ln L}{\partial \beta \partial \theta \partial \tau} = \frac{n_u}{\beta} + \frac{(n_u \cdot n_u) \beta}{\psi_{\beta}} - \frac{(n_u \cdot n_u) \theta}{\psi_{\theta}} + (x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau) \beta(y_i - \tau)}{\psi_i^2} \]
\[ \frac{\partial^3 \ln L}{\partial \beta \partial \theta \partial \tau} = \frac{n_u}{\beta} + \frac{(n_u \cdot n_u) \beta}{\psi_{\beta}} - \frac{(n_u \cdot n_u) \theta}{\psi_{\theta}} + (x + 1) \sum_{i=1}^{n} \delta_{2i} \frac{(y_i - \tau) \beta(y_i - \tau)}{\psi_i^2} \]
\[
L^{(3)}_{232} = L^{(3)}_{232} = \frac{\partial^3 \ln L}{\partial \theta^3} = \frac{(n - n_a - n_b)}{(\beta(y_i - \tau) + \tau + \theta)^3} \cdot \frac{n}{\theta^3} + \sum_{i=1}^{n} \frac{\delta_{1i}}{(\theta + y_i)^3} + \sum_{i=1}^{n} \frac{\delta_{2i}}{(\beta(y_i - \tau) + \tau + \theta)^3}
\]

\[
L^{(3)}_{311} = L^{(3)}_{311} = \frac{\partial^3 \ln L}{\partial \beta^3} = \frac{(n - n_a - n_b)(y_i - \tau)^3}{(\beta(y_i - \tau) + \tau + \theta)^3} + \sum_{i=1}^{n} \frac{\beta(y_i - \tau)^3}{(\beta(y_i - \tau) + \tau + \theta)^3}
\]

References


