Bayesian Control Charts for the Two-parameter Exponential Distribution if the Location Parameter Can Take on Any Value Between Minus Infinity and Plus Infinity

R. van Zyl¹, A.J. van der Merwe²

¹Quintiles International, ruaanvz@gmail.com

²University of the Free State
Abstract

By using data that are the mileages for some military personnel carriers that failed in service given by Grubbs (1971) and Krishnamoorthy and Mathew (2009) a Bayesian procedure is applied to obtain control limits for the location and scale parameters, as well as for a one-sided upper tolerance limit in the case of the two-parameter exponential distribution. A comparison between the assumptions of $-\infty < \mu < \infty$ and $0 < \mu < \infty$ are also made. An advantage of the upper tolerance limit is that it monitors the location and scale parameter at the same time. By using Jeffreys' non-informative prior, the predictive distributions of future maximum likelihood estimators of the location and scale parameters are derived analytically. The predictive distributions are used to determine the distribution of the “run-length” and expected “run-length”. This paper illustrates the flexibility and unique features of the Bayesian simulation method.

Keywords: Jeffreys' prior, two-parameter exponential, Bayesian procedure, run-length, control chart

1 Introduction

In this section the same notation will be used as given in Krishnamoorthy and Mathew (2009) with the exception that the location parameter can now take on values between $-\infty$ and $\infty$, similarly as in some literature, see for example Johnson and Kotz (1970).

Therefore the two-parameter exponential distribution has the probability density function

\[ f(x; \mu, \theta) = \frac{1}{\theta} \exp\left\{ -\frac{(x - \mu)}{\theta} \right\} \quad x > \mu, \quad -\infty < \mu < \infty, \quad \theta > 0 \]

where $\mu$ is the location parameter and $\theta$ the scale parameter.

As before, let $X_1, X_2, \ldots, X_n$ be a sample of $n$ observations from the two-parameter exponential distribution. The maximum likelihood estimators for $\mu$ and $\theta$ are given by

\[ \hat{\mu} = X_{(1)} \]

and

\[ \hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} (X_i - X_{(1)}) = \bar{X} - X_{(1)} \]

where $X_{(1)}$ is the minimum or the first order statistic of the sample. It is well known (see Johnson and Kotz (1970); Lawless (1982); Krishnamoorthy and Mathew (2009)) that $\hat{\mu}$ and $\hat{\theta}$ are independently distributed with

\[ \frac{(\hat{\mu} - \mu)}{\theta} \sim \chi^2_{2n} \quad \text{and} \quad \frac{\hat{\theta}}{\theta} \sim \chi^2_{2n-2} \]. \hspace{1cm} (1.1)\]

Let $\hat{\mu}_0$ and $\hat{\theta}_0$ be observed values of $\hat{\mu}$ and $\hat{\theta}$ then it follows from Equation 1.1 that a generalized pivotal quantity (GPQ) for $\mu$ is given by

\[ G_{\mu} = \hat{\mu}_0 - \frac{\chi^2_{2n}}{\chi^2_{2n-2}} \hat{\theta}_0 \] \hspace{1cm} (1.2)

and a GPQ for $\theta$ is given by
\[ G_\theta = \frac{2n\hat{\theta}_0}{\chi^2_{2n-2}} \tag{1.3} \]

From a Bayesian perspective it will be shown that \( G_\mu \) and \( G_\theta \) are actually the posterior distributions of \( \mu \) and \( \theta \) if the prior \( p(\mu, \theta) \propto \theta^{-1} \) is used.

\section{Bayesian Procedure}

In this section it will be shown that the Bayesian procedure is the same as the generalized variable approach.

If a sample of \( n \) observations are drawn from the two-parameter exponential distribution, then the likelihood function is given by

\[
L(\mu, \theta|\text{data}) = \left(\frac{1}{\theta}\right)^n \exp\left\{\frac{-1}{\theta} \sum_{i=1}^{n} (x_i - \mu)\right\}.
\]

As prior the Jeffreys’ prior

\[ p(\mu, \theta) \propto \theta^{-1} \]

will be used.

The joint posterior distribution of \( \mu \) and \( \theta \) is

\[
p(\theta, \mu|\text{data}) \propto p(\mu, \theta) L(\mu, \theta|\text{data})
\]

\[ = K_1 \left(\frac{1}{\hat{\theta}}\right)^{n+1} \exp\left\{\frac{-n}{\theta} (\bar{x} - \mu)\right\} \quad -\infty < \mu < \text{x(1)}, \quad 0 < \theta < \infty \tag{2.1} \]

It can easily be shown that

\[
K_1 = \frac{n^n \left(\hat{\theta}\right)^{n-1}}{\Gamma(n-1)} \text{ where } \hat{\theta} = \bar{x} - \text{x(1)}.
\]

The posterior distribution of \( \mu \) is

\[
p(\mu|\text{data}) = \int_0^\infty p(\theta, \mu|\text{data}) d\theta
\]

\[ = (n-1) \left(\hat{\theta}\right)^{n-1} \left(\frac{1}{\bar{x}-\mu}\right)^n \quad -\infty < \mu < \text{x(1)} \tag{2.2} \]

and the posterior distribution of \( \theta \) is

\[
p(\theta|\text{data}) = \int_{-\infty}^{\text{x(1)}} p(\theta, \mu|\text{data}) d\mu
\]

\[ = K_2 \left(\frac{1}{\hat{\theta}}\right)^n \exp\left\{\frac{-2\theta}{\theta}\right\} \text{ where } 0 < \theta < \infty \tag{2.3} \]

an Inverse Gamma distribution where
The conditional posterior distribution of \( \theta \) given \( \mu \) is given by

\[
p(\theta | \mu, \text{data}) = \frac{p(\theta, \mu | \text{data})}{p(\mu | \text{data})} = K_3 \left( \frac{1}{\theta} \right)^{n+1} \exp \left\{ - \frac{n}{\theta} (\bar{x} - \mu) \right\} \quad \text{where } 0 < \theta < \infty
\]

an Inverse Gamma distribution where

\[
K_3 = \frac{n \left( \bar{x} - \hat{\mu} \right)}{\Gamma(n)}.
\]

Also the conditional posterior distribution of \( \mu \) given \( \theta \) is

\[
p(\mu | \theta, \text{data}) = \frac{p(\theta, \mu | \text{data})}{p(\theta | \text{data})} = \frac{2}{\theta} \exp \left\{ - \frac{n}{\theta} (x(1) - \mu) \right\} \quad \text{where } -\infty < \mu < x(1)
\]

The following theorem can now easily be proved:

**Theorem 2.1.** The distribution of the generalized pivotal quantities \( G_\mu \) and \( G_\theta \) defined in Equations 1.2 and 1.3 are exactly the same as the posterior distributions \( p(\mu | \text{data}) \) and \( p(\theta | \text{data}) \) given in Equations 2.2 and 2.3.

**Proof.** The proof is given in Appendix A.

\[ \square \]

3 The Predictive Distributions of Future Sample Location and Scale Maximum Likelihood Estimators, \( \hat{\mu}_f \) and \( \hat{\theta}_f \)

Consider a future sample of \( m \) observations from the two-parameter exponential population: \( X_{1f}, X_{2f}, \ldots, X_{mf} \). The future sample mean is defined as \( \bar{X}_f = \frac{1}{m} \sum_{j=1}^{m} X_{jf} \). The smallest value in the sample is denoted by \( \hat{\mu}_f \) and \( \hat{\theta}_f = \bar{X}_f - \hat{\mu}_f \). To obtain control charts for \( \hat{\mu}_f \) and \( \hat{\theta}_f \), their predictive distributions must first be derived. If \( -\infty < \mu < \infty \), then the descriptive statistics, posterior distributions and predictive distributions will be denoted by a tilde (\( \tilde{\cdot} \)).

**Theorem 3.1.** The predictive distribution of a future sample location maximum likelihood estimator, \( \hat{\mu}_f \) is given by

\[
\tilde{f}(\hat{\mu}_f | \text{data}) = \begin{cases} 
K^* \left[ \frac{1}{n(x - \hat{\mu}_f)} \right]^n & -\infty < \hat{\mu}_f < x(1) \\
K^* \left[ \frac{1}{n\hat{\theta} + m(\hat{\mu}_f - x(1))} \right]^n & x(1) < \hat{\mu}_f < \infty 
\end{cases}
\]

where

\[
K^* = \frac{n^n (n-1) m^m}{n^m (n+m)} \left( \hat{\theta} \right)^{n-1}.
\]
Proof. The proof is given in Appendix B.

The reason why \( \tilde{f}(\hat{\mu}_f|data) \) (Equation 3.1) differs from \( f(\hat{\mu}_f|data) \) is that it is assumed that \( 0 < \mu < \infty \) which results in a posterior distribution of \( p(\mu|data) = (n-1) \left\{ \left( \frac{1}{\theta} \right)^{n-1} - \left( \frac{1}{\bar{x}} \right)^{n-1} \right\} (\bar{x} - \mu)^{-n} \), while for Equation 3.1 it is assumed that \( -\infty < \mu < \infty \) and the posterior distribution for \( \mu \) is therefore
\[ \tilde{p}(\mu|data) \propto (n-1)^{-1} (\hat{\theta})^{n-1} (\bar{x} - \mu)^{-n}. \]

**Theorem 3.2.** The mean and variance for \( \hat{\mu}_f \) is given by
\[ \tilde{E}(\hat{\mu}_f|data) = \bar{x} - \frac{mn - m - n}{m (n-2)} \hat{\theta} \] (3.2)
and
\[ \tilde{V}ar(\hat{\mu}_f|data) = \frac{n^3 (n-2) + (mn - m - n)^2}{m^2 (n-1)(n-3)(n-2)\hat{\theta}^2} \] (3.3)

Proof. By deleting the term \( \left( \frac{1}{\bar{x}} \right) \) from results from the previous report where \( 0 < \mu < \infty \) Equation 3.2 and Equation 3.3 follow.

**Theorem 3.3.** The predictive distribution of a future sample scale maximum likelihood estimator, \( \hat{\theta}_f \) is given by
\[ \tilde{f}(\hat{\theta}_f|data) = \frac{\Gamma(m + n - 2)m^{m-1}(\hat{\theta})^{n-1}(\hat{\theta}_f)^{m-2}}{\Gamma(m-1)\Gamma(n-1)(m\hat{\theta}_f + n\hat{\theta})^{m+n-2}} 0 < \hat{\theta}_f < \infty \] (3.4)

Proof. The proof is given in Appendix C.

**Corollary 3.4.** \( \hat{\theta}_f|data \sim \hat{\theta} \frac{n}{m} \frac{(m-1)}{(n-1)} F_{2m-2,2n-2}. \)

Proof. The proof is given in Appendix D.

**Theorem 3.5.** The mean and variance of \( \hat{\theta}_f \) is given by
\[ \tilde{E}(\hat{\theta}_f|data) = \frac{(m-1)}{m} \frac{n}{(n-2)} \hat{\theta} \] (3.5)
and
\[ \tilde{V}ar(\hat{\theta}_f|data) = \frac{n^2 (m-1)}{m^2 (n-2)^2 (n-3)} (n + m - 3) \hat{\theta}^2 \] (3.6)

Proof. By deleting the term \( \left( \frac{1}{\bar{x}} \right) \) from results from the previous report where \( 0 < \mu < \infty \), Equation 3.5 and Equation 3.6 follow.
4 Example

The following data is given in Grubbs (1971) as well as in Krishnamoorthy and Mathew (2009). The failure mileages given in Table 4.1 fit a two-parameter exponential distribution.

<table>
<thead>
<tr>
<th>162</th>
<th>200</th>
<th>271</th>
<th>302</th>
<th>393</th>
<th>508</th>
<th>539</th>
<th>629</th>
<th>706</th>
<th>777</th>
</tr>
</thead>
<tbody>
<tr>
<td>884</td>
<td>1008</td>
<td>1101</td>
<td>1182</td>
<td>1463</td>
<td>1603</td>
<td>1984</td>
<td>2355</td>
<td>2880</td>
<td></td>
</tr>
</tbody>
</table>

For this data, the maximum likelihood estimates are $\hat{\mu} = x_{(1)} = 162$, $\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} (x_i - x_{(1)}) = \bar{x} - x_{(1)} = 835.21$ and $n = 19$.

As mentioned in the introductory section, the aim of this article is to obtain control charts for location and scale maximum likelihood estimates as well as for a one-sided upper tolerance limit.

4.1 The Predictive Distribution of $\hat{\mu}_f (\mu > -\infty)$

By using Equation 3.1 the predictive distribution $\tilde{f}(\hat{\mu}_f|\text{data})$ for $m = 19$ future failure mileage data is illustrated in Figure 4.1.

Figure 4.1: Distribution of $\hat{\mu}_f$, $n = 19$, $m = 19$

$\text{mean}(\hat{\mu}_f) = \text{median}(\hat{\mu}_f) = \text{mode}(\hat{\mu}_f) = 162$, $\text{var}(\hat{\mu}_f) = 5129.2$

95% interval ($\hat{\mu}_f$) = (11.2; 312.8)
99.73% interval ($\hat{\mu}_f$) = (−154.85; 477.36)
96.08% interval ($\hat{\mu}_f$) = (−2; 326)
For $-\infty < \mu < \infty$, $n = 19$, $m = 19$, the predictive distribution $\tilde{p}(\hat{\mu}_f|data)$ is symmetrical. $\text{mean}(\hat{\mu}_f) = \text{median}(\hat{\mu}_f) = \text{mode}(\hat{\mu}_f) = 162$. $\text{Mode}(\hat{\mu}_f)$ for $-\infty < \mu < \infty$ is exactly the same as that for $0 < \mu < \infty$. A further comparison of Figure 4.1 and results from previous report shows that $\text{var}(\hat{\mu}_f) = 5129.2$ is somewhat larger than $\text{var}(\hat{\mu}_f) = 3888.7$, when $0 < \mu < \infty$. Also the predictive intervals are somewhat wider. A dissatisfactory aspect for $-\infty < \mu < \infty$ is that the $99.73\%$ interval $(\hat{\mu}_f) = (-154.85; 477.36)$, i.e., contains negative values.

In Table 4.2 descriptive statistics are given for the run-length and expected run-length.

Table 4.2: Descriptive Statistics for the Run-length and Expected Run-length in the case of $\hat{\mu}_f$: $-\infty < \mu < \infty$ and $\beta = 0.039$ for $n = 19$ and $m = 19$.

| Descriptive Statistics | $f(r|data)$ Expected Run-length |
|------------------------|----------------------------------|
|                        | Equal Tail                      |
| $\text{mean}$          | 367.84                          |
| $\text{median}$        | 44.50                           |
| $\text{var}$           | $6.663 \times 10^7$            |
| $95\%$ interval        | $(0; 740.50)$                   |

A comparison of Table 4.2 with results from previous report shows that if $-\infty < \mu < \infty$ the expected (mean) run-length $\approx 370$ if $\beta = 0.0392$ while this is the case if $\beta = 0.0258$ for $0 < \mu < \infty$. Also the $95\%$ intervals are somewhat shorter for $-\infty < \mu < \infty$.

### 4.2 A Comparison of the Predictive Distributions for $\hat{\theta}_f$

In Figure 4.2 comparisons are made between

$$f(\hat{\theta}_f|data) = \frac{\Gamma(m + n - 2) m^{n-1} \left(\hat{\theta}\right)^{n-1} \left(\hat{\theta}_f\right)^{m-2}}{\Gamma(m-1) \Gamma(n-1) \left(\hat{\theta} + \hat{n}\right)^{m+n-2}} 0 < \hat{\theta}_f < \infty$$

and

$$f(\hat{\theta}_f|data) = \frac{\Gamma(m + n - 2)}{\Gamma(m-1) \Gamma(n-1)} \frac{1}{\left(\hat{\theta} + \hat{n}\right)^{m+n-2}} \left(\frac{1}{\hat{\theta}}\right)^{n-1} \left(\frac{1}{\hat{x}}\right)^{n-1} \left(\hat{\theta}_f\right)^{m-2}$$

$$\times \left(\frac{1}{\hat{\theta} + \hat{n}}\right)^{m+n-2} \left(\frac{1}{\hat{\theta} + \hat{x}}\right)^{m+n-2} 0 < \hat{\theta}_f < \infty.$$
In Table 4.3 descriptive statistics are given for $\tilde{f}(\hat{\theta}|data)$ and $f(\hat{\theta})$.

Table 4.3: Descriptive Statistics of $\tilde{f}(\hat{\theta}|data)$ and $f(\hat{\theta}|data)$

| Descriptive Statistics | $\tilde{f}(\hat{\theta}|data)$ | $f(\hat{\theta}|data)$ |
|------------------------|-----------------------------|-----------------------------|
| Mean $(\hat{\theta})$  | 884.34                      | 876.98                      |
| Median $(\hat{\theta})$ | 835.2                       | 829.1                       |
| Mode $(\hat{\theta})$  | 747.3                       | 743.3                       |
| Var $(\hat{\theta})$   | 95037                       | 91991                       |
| 95% Equal - tail Interval $(\hat{\theta})$ | (430; 1622) | (428; 1601) |
| Length                  | 1192                        | 1173                        |
| 95% HPD Interval $(\hat{\theta})$ | (370.8; 1500.5) | (369.2; 1482.6) |
| Length                  | 1129.7                      | 1113.4                      |

Also, the exact means and variances are given by:

$$\tilde{E}\left(\hat{\theta}|data\right) = \frac{n(m - 1)}{m(n - 2)} \hat{\theta} = 884.34,$$

$$Var\left(\hat{\theta}\right) = \frac{n^2(m - 1)}{m^2(n - 2)^2(n - 3)} (n + m - 3) \hat{\theta}^2 = 95042,$$
\[ E \left( \hat{\theta}_f | \text{data} \right) = \frac{n(\theta - 1)}{m(n - 2)}KL = 976.98, \]

\[ Var \left( \hat{\theta}_f | \text{data} \right) = \frac{n^2(m - 1)}{m(n - 1)} \left\{ \frac{KL}{n - 2} - \frac{m - 1}{m(n - 1)}KL^2 \right\} = 91991 \]

where

\[ \bar{K} = (n - 1) \left\{ \left( \frac{1}{\theta} \right)^{n-1} - \left( \frac{1}{x} \right)^{n-1} \right\}^{-1} \]

\[ L = \frac{1}{n - 2} \left\{ \left( \frac{1}{\theta} \right)^{n-2} - \left( \frac{1}{x} \right)^{n-2} \right\} \]

and

\[ M = \frac{1}{n - 3} \left\{ \left( \frac{1}{\theta} \right)^{n-3} - \left( \frac{1}{x} \right)^{n-3} \right\} \]

From Figure 4.2 and Table 4.3 it can be seen that \( f \left( \hat{\theta}_f | \text{data} \right) \) and \( f \left( \hat{\theta}_f | \text{data} \right) \) are for all practical purposes the same. Also the exact means and variances are very much the same as the numerical values. It therefore seems that whether the assumption is that \(-\infty < \mu < \infty\) or that \(0 < \mu < \infty\) does not play a big role in the prediction of \(\hat{\theta}_f\).

In Table 4.4 comparisons are made between the run-lengths and expected run-lengths in the case of \(\hat{\theta}_f\) for \(-\infty < \mu < \infty\) and \(0 < \mu < \infty\).

### Table 4.4: Descriptive Statistics for the Run-length and Expected Run-lengths in the case of \(\hat{\theta}_f\), \(-\infty < \mu < \infty\), \(0 < \mu < \infty\) and \(\beta = 0.018\)

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>(-\infty &lt; \mu &lt; \infty)</th>
<th>Expected Run-Length</th>
<th>(0 &lt; \mu &lt; \infty)</th>
<th>Expected Run-Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(f(r</td>
<td>data)) Equal Tail</td>
<td>(E(r</td>
<td>data)) Equal Tail</td>
</tr>
<tr>
<td>Mean</td>
<td>369.29</td>
<td>370.42</td>
<td>375.25</td>
<td>375.25</td>
</tr>
<tr>
<td>Median</td>
<td>122.8</td>
<td>229.28</td>
<td>127.4</td>
<td>238.38</td>
</tr>
<tr>
<td>Variance</td>
<td>(3.884 \times 10^5)</td>
<td>(1.2572 \times 10^5)</td>
<td>(3.9515 \times 10^5)</td>
<td>(1.2698 \times 10^5)</td>
</tr>
<tr>
<td>95% Interval</td>
<td>(0; 1586)</td>
<td>(7.423; 1089.7)</td>
<td>(0; 1602.5)</td>
<td>(7.809; 1096.7)</td>
</tr>
</tbody>
</table>

From Table 4.4 it can be seen that the corresponding statistics for \(-\infty < \mu < \infty\) and \(0 < \mu < \infty\) are for all practical purposes the same. So with respect to \(\hat{\theta}_f\) it does not really matter whether it is assumed that \(\mu\) is positive or not.
5 Phase I Control Chart for the Scale Parameter in the Case of the Two-parameter Exponential Distribution

Statistical quality control is implemented in two phases. In Phase I the primary interest is to assess process stability. Phase I is the so-called retrospective phase and Phase II the prospective or monitoring phase. The construction of Phase I control charts should be considered as a multiple testing problem. The distribution of a set of dependent variables (ratios of chi-square random variables) will therefore be used to calculate the control limits so that the false alarm probability (FAP) is not larger than $FAP_0 = 0.05$. To obtain control limits in Phase I, more than one sample is needed. Therefore in the example that follows there will be $m = 5$ samples for a subgroup each of size $n = 10$.

Example 5.1

The data in Table 5.1 are simulated data obtained from the following two-parameter exponential distribution:

$$f(x_{ij}; \theta, \mu) = \frac{1}{\theta} \exp \left\{ -\frac{x_{ij} - \mu_i}{\theta} \right\}, \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n; \quad x_{ij} > \mu_i$$

\[ \theta = 8; \quad \mu_i = 2i; \quad m = 5; \quad n = 10. \]

Table 5.1: Simulated Data for the Two-parameter Exponential Distribution

<table>
<thead>
<tr>
<th>$\mu_i$</th>
<th>2</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7916</td>
<td>4.2388</td>
<td>32.6582</td>
<td>35.5781</td>
<td>17.7079</td>
<td></td>
</tr>
<tr>
<td>18.5094</td>
<td>4.3502</td>
<td>7.3084</td>
<td>18.2721</td>
<td>12.1376</td>
<td></td>
</tr>
<tr>
<td>2.749</td>
<td>9.7827</td>
<td>6.5463</td>
<td>32.6032</td>
<td>11.8333</td>
<td></td>
</tr>
<tr>
<td>5.6664</td>
<td>5.7823</td>
<td>9.1002</td>
<td>26.6535</td>
<td>23.4186</td>
<td></td>
</tr>
<tr>
<td>12.2267</td>
<td>10.9065</td>
<td>8.3750</td>
<td>10.9177</td>
<td>16.4669</td>
<td></td>
</tr>
<tr>
<td>6.8282</td>
<td>4.7042</td>
<td>13.4873</td>
<td>17.1883</td>
<td>13.4918</td>
<td></td>
</tr>
<tr>
<td>2.3474</td>
<td>6.8636</td>
<td>9.3791</td>
<td>8.4085</td>
<td>12.7474</td>
<td></td>
</tr>
<tr>
<td>2.2859</td>
<td>4.3308</td>
<td>20.1200</td>
<td>34.9469</td>
<td>12.2616</td>
<td></td>
</tr>
</tbody>
</table>

From the simulated data in Table 5.1 we have

$$\hat{\theta}_i = \bar{X}_i - X_{(1,i)} = \left[ \begin{array}{c} 5.4780 \ 4.5974 \ 5.9107 \ 12.0817 \ 34.023 \end{array} \right],$$

$$\sum_{i=1}^{m} \hat{\theta}_i = 31.4701$$

and

$$\hat{\theta} = \frac{1}{m} \sum_{i=1}^{m} \hat{\theta}_i = 6.2940.$$
It is well known that \( \hat{\theta}_i \sim \frac{\theta}{2n} \chi^2_{2n-2} = \frac{\theta}{2n} Y_i \) and therefore \( \sum_{i=1}^{m} \hat{\theta}_i \sim \frac{\theta}{2n} \sum_{i=1}^{m} Y_i \).

Let

\[
Z_1 = \frac{\hat{\theta}_i}{\sum_{i=1}^{m} \hat{\theta}_i} = \frac{\frac{\theta}{2n} Y_i}{\frac{\theta}{2n} \sum_{i=1}^{m} Y_i} = \frac{Y_i}{\sum_{i=1}^{m} Y_i} \quad i = 1, 2, \ldots, m
\]

where

\( Y_i \sim \chi^2_{2n-2} \).

For further details see also Human, Chakraborti, and Smit (2010).

To obtain a lower control limit for the data in Table 5.1, the distribution of \( Z_{\min} = \min (Z_1, Z_2, \ldots, Z_m) \) must be obtained.

Figure 5.1: Distribution of \( Z_{\min} = \min (Z_1, Z_2, \ldots, Z_m) \), 100 000 simulations

The distribution of \( Z_{\min} \) obtained from 100 000 simulations is illustrated in Figure 5.1. The value \( Z_{0.05} = 0.0844 \) is calculated such that the FAP is at a level of 0.05. The lower control limit is then determined as

\[
LCL = Z_{0.05} \sum_{i=1}^{m} \hat{\theta}_i = (0.0844)(31.4701) = 2.656.
\]

Since \( \hat{\theta}_i > 2.656 \) \( (i = 1, 2, \ldots, m) \) it can be concluded that the scale parameter is under statistical control.
6 Lower Control Limit for the Scale Parameter in Phase II

In the first part of this section, the lower control limit in a Phase II setting will be derived using the Bayesian predictive distribution.

The following theorem can easily be proved:

**Theorem 6.1.** For the two-parameter exponential distribution

\[
  f(x_{ij}; \theta, \mu_i) = \left(\frac{1}{\theta}\right) \exp\left\{ -\frac{1}{\theta} (x_{ij} - \mu_i) \right\} \quad i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n, \quad \text{and } x_{ij} > \mu_i
\]

the posterior distribution of the parameter \( \theta \) given the data is given by

\[
p(\theta | \text{data}) = \frac{(n\hat{\theta})^{m(n-1)}}{\Gamma(m(n-1)) \left(\frac{1}{\theta}\right)^{m(n-1)+1}} \exp\left( -\frac{n\hat{\theta}}{\theta} \right) \quad \theta > 0
\]

an Inverse Gamma Distribution.

**Proof.** The proof is given in Appendix E.

**Theorem 6.2.** Let \( \hat{\theta}_f \) be the maximum likelihood estimator of the scale parameter in a future sample of \( n \) observations, then the predictive distribution for \( \hat{\theta}_f \) is

\[
f\left(\hat{\theta}_f | \text{data}\right) = \frac{\Gamma[m(n-1)+n-1]}{\Gamma(n-1) \Gamma[m(n-1)]} \left(\frac{\hat{\theta}_f}{\hat{\theta}_f + \theta}\right)^{n-2} \left(\hat{\theta}_f + \theta\right)^{m(n-1)+n-1} \quad \hat{\theta}_f > 0
\]

which means that

\[
\hat{\theta}_f | \text{data} \sim \frac{\hat{\theta}}{m} F_{2(n-1):2m(n-1)}
\]

where

\[
\hat{\theta} = \sum_{i=1}^{m} \hat{\theta}_i.
\]

**Proof.** The proof is given in Appendix F.

At \( \beta = 0.0027 \) the lower control limit is obtained as \( \frac{\hat{\theta}}{m} F_{2(n-1):2m(n-1)} (0.0027) = \frac{31.4701}{5} (0.29945) = 1.8847 \) for \( m = 5 \) and \( n = 10 \).

Assuming that the process remains stable, the predictive distribution for \( \hat{\theta}_f \) can also be used to derive the distribution of the run-length, that is the number of samples until the control chart signals for the first time.

The resulting region of size \( \beta \) using the predictive distribution for the determination of the run-length is defined as

\[
\beta = \int_{R(\beta)} f\left(\hat{\theta}_f | \text{data}\right) d\hat{\theta}_f
\]

12
where

\[ R(β) = (0; 1.8847) \]

is the lower one-sided control interval.

Given \( θ \) and a stable process, the distribution of the run-length \( r \) is Geometric with parameter

\[ ψ(θ) = \int_{R(β)} f(\hat{θ}|θ) d\hat{θ} \]

where \( f(\hat{θ}|θ) \) is the distribution of a future sample scale parameter estimator given \( θ \).

The value of the parameter \( θ \) is however unknown and its uncertainty is described by the posterior distribution \( p(θ|data) \).

The following theorem can also be proved.

**Theorem 6.3.** For given \( θ \) the parameter of the Geometric distribution is

\[ ψ(θ) = ψ(\chi_{2m(n−1)}^2) \]

for given \( χ_{2m(n−1)}^2 \) which means that the parameter is only dependent on \( χ_{2m(n−1)}^2 \) and not on \( θ \).

**Proof.** The proof is given in Appendix G.

As mentioned, by simulating \( θ \) from \( p(θ|data) \) the probability density function of \( f(\hat{θ}|θ) \) as well as the parameter \( ψ(θ) \) can be obtained. This must be done for each future sample. Therefore, by simulating a large number of \( θ \) values from the posterior distribution a large number of \( ψ(θ) \) values can be obtained.

A large number of Geometric and run-length distributions with different parameter values \( (ψ(θ_1), ψ(θ_2), ..., ψ(θ_l)) \) will therefore be available. The unconditional run-length distribution is obtained by using the Rao-Blackwell method, i.e., the average of the conditional run-length distributions.

In Table 6.1 results for the run-length at \( β = 0.0027 \) for \( n = 10 \) and different values for \( m \) are presented for the lower control limit of the scale parameter estimator.
<table>
<thead>
<tr>
<th>n</th>
<th>m</th>
<th>mean (ARL)</th>
<th>median (ARL)</th>
<th>Mean (PDF)</th>
<th>Median (PDF)</th>
<th>One-sided</th>
<th>Two-sided</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>5</td>
<td>1040.878616</td>
<td>561.9041434</td>
<td>1016.107825</td>
<td>365.604745</td>
<td>120.668063</td>
<td>328.560474</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>859.2946037</td>
<td>528.043143</td>
<td>850.6496099</td>
<td>340.107825</td>
<td>132.1636771</td>
<td>263.779211</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>752.7840665</td>
<td>502.703642</td>
<td>747.964747</td>
<td>330.107825</td>
<td>137.135507</td>
<td>2163.362738</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>692.3117213</td>
<td>478.8131128</td>
<td>688.5627905</td>
<td>320.107825</td>
<td>147.78909272</td>
<td>1904.509651</td>
</tr>
<tr>
<td>10</td>
<td>9</td>
<td>644.1233737</td>
<td>467.3792855</td>
<td>640.6357323</td>
<td>310.107825</td>
<td>149.1918895</td>
<td>1707.552491</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>616.4008914</td>
<td>456.271994</td>
<td>613.9915372</td>
<td>305.107825</td>
<td>154.9653041</td>
<td>1581.526121</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>582.5441167</td>
<td>445.4805119</td>
<td>579.3080327</td>
<td>295.107825</td>
<td>159.4756421</td>
<td>1466.351811</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>563.8471571</td>
<td>445.4805119</td>
<td>560.8517002</td>
<td>295.107825</td>
<td>165.742153</td>
<td>1340.987775</td>
</tr>
<tr>
<td>10</td>
<td>13</td>
<td>542.293566</td>
<td>440.1999559</td>
<td>539.3080327</td>
<td>290.107825</td>
<td>170.6404787</td>
<td>1264.446813</td>
</tr>
<tr>
<td>10</td>
<td>14</td>
<td>523.2584573</td>
<td>429.8629367</td>
<td>520.4260911</td>
<td>285.107825</td>
<td>170.6404787</td>
<td>1193.031076</td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>514.4424077</td>
<td>429.8629367</td>
<td>511.6810866</td>
<td>285.107825</td>
<td>177.4499555</td>
<td>1142.597024</td>
</tr>
<tr>
<td>10</td>
<td>16</td>
<td>498.7108368</td>
<td>424.804032</td>
<td>496.0401877</td>
<td>285.107825</td>
<td>179.2039115</td>
<td>1064.049573</td>
</tr>
<tr>
<td>10</td>
<td>17</td>
<td>493.4525335</td>
<td>419.816607</td>
<td>490.7979633</td>
<td>285.107825</td>
<td>179.2039115</td>
<td>1064.049573</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>491.980728</td>
<td>424.804032</td>
<td>489.3787352</td>
<td>285.107825</td>
<td>188.2980341</td>
<td>1019.999561</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>479.4053832</td>
<td>414.8994965</td>
<td>476.8199625</td>
<td>280.107825</td>
<td>184.5942927</td>
<td>978.1005367</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>470.1638992</td>
<td>410.0515567</td>
<td>467.6270067</td>
<td>275.107825</td>
<td>188.2980341</td>
<td>951.3034035</td>
</tr>
<tr>
<td>10</td>
<td>50</td>
<td>408.6904318</td>
<td>386.8108266</td>
<td>406.482335</td>
<td>265.107825</td>
<td>235.6714776</td>
<td>654.3843849</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>389.048816</td>
<td>377.960703</td>
<td>386.9577637</td>
<td>260.107825</td>
<td>267.625879</td>
<td>548.0661023</td>
</tr>
<tr>
<td>10</td>
<td>500</td>
<td>373.4715597</td>
<td>373.6274685</td>
<td>373.573508</td>
<td>255.107825</td>
<td>318.8269483</td>
<td>440.1999555</td>
</tr>
<tr>
<td>10</td>
<td>1000</td>
<td>372.9790501</td>
<td>373.6274685</td>
<td>370.9764548</td>
<td>255.107825</td>
<td>333.4301277</td>
<td>414.8904065</td>
</tr>
<tr>
<td>10</td>
<td>5000</td>
<td>371.9260097</td>
<td>373.6274685</td>
<td>369.9395457</td>
<td>255.107825</td>
<td>352.8387861</td>
<td>391.329318</td>
</tr>
<tr>
<td>10</td>
<td>10000</td>
<td>371.7307984</td>
<td>373.6274685</td>
<td>369.732037</td>
<td>255.107825</td>
<td>356.8826129</td>
<td>386.8108266</td>
</tr>
</tbody>
</table>

mean(ARL) and median(ARL) refer to results obtained from the expected run-length while mean(PDF) and median(PDF) refer to results obtained from the probability density function of the run-length.
From Table 6.1 it can be seen that as the number of samples increase (larger \( m \)) the mean and median run-lengths converge to the expected run-length of 370.

Further define \( \bar{\psi}(\theta) = \frac{1}{l} \sum_{i=1}^{l} \psi(\theta_i) \). From Menzefricke (2002) it is known that as \( l \to \infty \), \( \bar{\psi}(\theta) \to \beta = 0.0027 \) and the harmonic mean of the unconditional run-length will be \( \left( \frac{1}{\beta} \right) = \frac{1}{0.0027} = 370 \). Therefore it does not matter how small \( m \) and \( n \) is, the harmonic mean of the run-length will always be \( \frac{1}{\beta} \) if \( l \to \infty \). In the case of the simulated example the mean run-length is 1040.88 and the median run-length 561.90. The reason for these large values is the uncertainty in the parameter estimate because of the small sample size and number of samples (\( n = 10 \) and \( m = 5 \)). \( \beta \) however can easily be adjusted to get a mean run-length of 370.

7 A Comparison of the Predictive Distributions and Control Charts for a One-sided Upper Tolerance Limit, \( U_f \)

A future sample tolerance limit is defined as

\[
U_f = \hat{\mu}_f - \tilde{k}_2 \hat{\theta}_f \quad \text{where} \quad \hat{\mu}_f > \mu \text{ and } \hat{\theta}_f > 0.
\]

Also

\[ f(\hat{\mu}_f|\mu, \theta) = \left( \frac{m}{\theta} \right)^{(m)} \exp \left\{ -\frac{m}{\theta} (\hat{\mu}_f - \mu) \right\} \hat{\mu}_f > \mu \]

which means that

\[
f(U_f|\mu, \theta, \hat{\theta}_f) = \left( \frac{m}{\hat{\theta}} \right)^{(m)} \exp \left\{ -\frac{m}{\hat{\theta}} \left[ U_f - (\mu - \tilde{k}_2 \hat{\theta}_f) \right] \right\} \quad U_f > \mu - \tilde{k}_2 \hat{\theta}_f \tag{7.1}
\]

A comparison will be made between \( f(U_f|data) \) and \( \tilde{f}(U_f|data) \). The difference in the simulation procedure for these two density functions is that in the case of \( f(U_f|data) \) it is assumed that \( 0 < \mu < \infty \) which results in a posterior distribution of \( p(\mu|data) = (n - 1) \left\{ \left( \frac{1}{\theta} \right)^{(n-1)} - \left( \frac{1}{\hat{\theta}} \right)^{(n-1)} \right\}^{-1} (\bar{x} - \mu)^{-n} \) while for \( \tilde{f}(U_f|data) \) it is assumed that \( -\infty < \mu < \infty \) and the posterior distribution for \( \mu \) is then \( \tilde{p}(\mu|data) = (n - 1) \left( \hat{\theta} \right)^{(n-1)} (\bar{x} - \mu)^{-n} \).

In the following figure comparisons are made between \( f(U_f|data) \) and \( \tilde{f}(U_f|data) \).
Figure 7.1: Predictive Densities of $f(U_f|data)$ and $\tilde{f}(U_f|data)$

The descriptive statistics obtained from Figure 7.1 are presented in Table 7.1.

| Descriptive Statistics | $f(U_f|data)$ | $\tilde{f}(U_f|data)$ |
|------------------------|--------------|----------------------|
| Mean ($U_f$)           | 3394.7       | 3406.6               |
| Median ($U_f$)         | 3211.5       | 3225.8               |
| Mode ($\hat{U}_f$)    | 2900         | 2907                 |
| $\text{Var} (\hat{U}_f)$ | $1.2317 \times 10^6$ | $1.2694 \times 10^6$ |
| 95% Equal-tail Interval | (1736.5; 6027) | (1738; 6106)         |
| 99.73% Equal-tail Interval | (1249.05; 7973) | (1249.9; 8620)       |

Also the exact means and variances for $f(U_f|data)$ are

$$E(U_f|data) = \bar{x} + \tilde{K}L(aH - 1) = 3394.8$$

and

$$\text{Var}(U_f|data) = \frac{n^2}{m^2(n-1)(n-2)} \left\{ J + \frac{H^2}{n-1} \right\} \tilde{K}M + (1-aH)^2 \left\{ \tilde{K}M - \tilde{K}^2L^2 \right\} = 1.2439 \times 10^6$$
where
\[ a = \frac{n}{m(n-1)}, \]
\[ H = 1 - \tilde{k}_2(m-1), \]
\[ J = 1 + \tilde{k}_2^2(m-1) \]
and \( \tilde{K} , L \) and \( M \) defined as before.

The exact means and variances for \( \tilde{f}(U_f|data) \) are
\[ \tilde{E}(U_f|data) = \bar{x} + \frac{n-1}{n-2} (aH - 1) \hat{\theta} = 3415.0 \]
and
\[ \tilde{V}ar(U_f|data) = \frac{1}{(n-2)(n-3)} \left\{ \frac{n^2}{m^2} \left( J + \frac{H^2}{n-1} \right) + (1-aH)^2 \frac{n-1}{n-2} \right\} (\hat{\theta})^2 = 1.2911 \times 10^6. \]

\( \tilde{E}(U_f|data) \) and \( \tilde{V}ar(U_f|data) \) are derived from \( E(U_f|data) \) and \( Var(U_f|data) \) by deleting the term \( \left( \frac{1}{2} \right) \) in \( \tilde{K}, L \) and \( M \).

It seems that the predictive intervals for \( \tilde{f}(U_f|data) \) are somewhat wider than in the case of \( f(U_f|data) \).

In Table 7.2 comparisons are made between the run-lengths and expected run-lengths in the case of \( U_f \) for \(-\infty < \mu < \infty \) and \( 0 < \mu < \infty \).

Table 7.2: Descriptive Statistics for the Run-lengths and Expected Run-lengths in the Case of \( U_f \);
\(-\infty < \mu < \infty ; \) \( 0 < \mu < \infty \) and \( \beta = 0.018 \)

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>(-\infty &lt; \mu &lt; \infty )</th>
<th>( 0 &lt; \mu &lt; \infty )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Equal Tail</td>
<td>Equal Tail</td>
</tr>
<tr>
<td>Mean</td>
<td>444.95</td>
<td>444.95</td>
</tr>
<tr>
<td>Median</td>
<td>136.5</td>
<td>258.12</td>
</tr>
<tr>
<td>Variance</td>
<td>( 7.6243 \times 10^5 )</td>
<td>( 2.8273 \times 10^5 )</td>
</tr>
<tr>
<td>95% Interval</td>
<td>( (0; 1892.4) )</td>
<td>( (0; 1803.6) )</td>
</tr>
</tbody>
</table>

It is clear from Table 10.1 that the corresponding statistics do not differ much. The mean, median and variance and 95% interval are however somewhat larger for \(-\infty < \mu < \infty \).

8 Conclusion

This paper develops a Bayesian control chart for monitoring the scale parameter, location parameter and upper tolerance limit of a two-parameter exponential distribution. In the Bayesian approach prior knowledge about the unknown parameters is formally incorporated into the process of inference by assigning a prior distribution to the parameters. The information contained in the prior is combined with the likelihood function to obtain the posterior distribution. By using the posterior distribution the predictive distributions of \( \hat{\mu}_f, \hat{\theta}_f \) and \( U_f \) can be obtained.

The theory and results described in this paper have been applied to the failure mileages for military carriers analyzed by Grubbs (1971) and Krishnamoorthy and Mathew (2009). The example illustrates the flexibility and unique features of the Bayesian simulation method for obtaining posterior distributions and “run-lengths” for \( \hat{\mu}_f, \hat{\theta}_f \) and \( U_f \).
References


A Proof of Theorem 2.1

(a) The posterior distribution $p(\theta|\text{data})$ is the same as the distribution of the pivotal quantity $G_\theta = \frac{2n\hat{\theta}}{\chi^2_{2n-2}}$.

Proof:

Let $Z \sim \chi^2_{2n-2}$

$\therefore f(z) = \frac{1}{2^{n-1}(n-1)!} z^{n-2} \exp \left\{ -\frac{1}{2} z \right\}$

We are interested in the distribution of $\theta = \frac{2n\hat{\theta}}{Z}$.

Therefore $Z = \frac{2n\hat{\theta}}{\theta}$ and $\left| \frac{dZ}{d\theta} \right| = \frac{2n\hat{\theta}}{Z}$.

From this it follows that

$$f(G_\theta) = f(\theta) = \frac{1}{2^{n-1}(n-1)!} \left( \frac{2n\hat{\theta}}{\theta} \right)^{n-2} 2n\hat{\theta} \exp \left\{ -\frac{n\hat{\theta}}{\theta} \right\}$$

$$= \frac{n^{n-1}(\hat{\theta})^{n-1}}{\Gamma(n-1)} \left( \frac{\hat{\theta}}{\theta} \right)^n \exp \left\{ -\frac{n\hat{\theta}}{\theta} \right\} = p(\theta|\text{data})$$

See Equation 2.3.

(b) The posterior distribution $p(\mu|\text{data})$ is the same as the distribution of the pivotal quantity $G_\mu = \hat{\mu} - \frac{\chi^2}{2n-2}$.

Proof:

Let $F = \frac{\chi^2/2}{\chi^2_{2n-2}/(2n-2)} \sim F_{2,2n-2}$

$\therefore g(f) = \left( 1 + \frac{1}{n-1} f \right)^{-n}$ where $0 < f < \infty$

We are interested in the distribution of $\mu = \hat{\mu}_0 - \frac{\chi^2}{2n-2} F$ which means that $F = \frac{(n-1)}{\theta} (\hat{\mu} - \mu)$ and $\left| \frac{dF}{d\mu} \right| = \frac{(n-1)}{\theta}$.

Therefore

$$g(\mu) = \left( 1 + \frac{1}{\theta} (\hat{\mu}_0 - \mu) \right)^{-n} \frac{n-1}{\theta}$$

$$= (n-1) \hat{\theta}^{n-1} \left( \frac{1}{\theta} \right)^n \text{ where } -\infty < \mu < \hat{\mu}$$

$$= p(\mu|\text{data})$$

See Equation 2.2.

(c) The posterior distribution of $p(\mu|\theta,\text{data})$ is the same as the distribution of the pivotal quantity $G_{\mu|\theta} = \hat{\mu} - \frac{\hat{\theta}^2}{2n}$ (see Equation 1.1).
Proof:

Let $\tilde{Z} \sim \chi^2_2$ then

$$g(\tilde{z}) = \frac{1}{2} \exp \left\{ -\frac{1}{2} \tilde{z} \right\}.$$ 

Let $\mu = \bar{\mu} - \frac{\tilde{z}}{2n} \theta$, then $\tilde{z} = \frac{2n}{\theta} (\bar{\mu} - \mu)$ and $\left| \frac{d\mu}{d\theta} \right| = \frac{2n}{\theta}$.

Therefore

$$g(\mu|\theta) = \frac{n}{\theta} \exp \left\{ -\frac{n}{\theta} (\bar{\mu} - \mu) \right\} -\infty < \mu < \bar{\mu}$$

$$= p(\mu|\theta, \text{data})$$

See Equation 2.5.

\section{B Proof of Theorem 3.1}

As before

$$f(\hat{\mu}_f|\mu, \theta) = \left( \frac{m}{\theta} \right) \exp \left\{ -\frac{m}{\theta} (\hat{\mu}_f - \mu) \right\} \hat{\mu}_f > \mu$$

and therefore

$$\tilde{f}(\hat{\mu}_f|\mu, \text{data}) = \int_0^\infty f(\hat{\mu}_f|\mu, \theta) \tilde{p}(\theta|\mu, \text{data}) d\theta.$$ 

Since

$$\tilde{p}(\theta|\mu, \text{data}) = \frac{n(\bar{x} - \mu)^n}{\Gamma(n)} \left( \frac{1}{\theta} \right)^{n+1} \exp \left\{ -\frac{n}{\theta} (\bar{x} - \mu) \right\}$$

it follows that

$$\tilde{f}(\hat{\mu}_f|\mu, \text{data}) = \frac{n^{n+1} (\bar{x} - \mu)^n m}{[m(\hat{\mu}_f - \mu) + n(\bar{x} - \mu)]^{n+1}} \hat{\mu}_f > \mu.$$ 

For $-\infty < \mu < x_{(1)}$,

$$\tilde{p}(\mu|\text{data}) = (n-1) \left( \bar{x} - \mu \right)^{n-1} \left( \frac{1}{\bar{x} - \mu} \right)^n$$

and

$$\tilde{f}(\hat{\mu}_f, \mu|\text{data}) = \tilde{f}(\hat{\mu}_f|\mu, \text{data}) \tilde{p}(\mu|\text{data})$$

$$= \frac{n^{n+1} m(n-1) (\hat{\theta})^{n-1}}{[m(\hat{\mu}_f - \mu) + n(\bar{x} - \mu)]^{n+1}}.$$ 

Therefore
\[ f(\hat{\mu}_f|\text{data}) = \int_{-\infty}^{\hat{\mu}_f} f(\hat{\mu}_f, \mu|\text{data}) d\mu \quad -\infty < \hat{\mu}_f < x_{(1)} \]

\[ = \int_{-\infty}^{x_{(1)}} f(\hat{\mu}_f, \mu|\text{data}) d\mu \quad x_{(1)} < \hat{\mu} < \infty \]

\[ = \tilde{K}^* \left[ \frac{1}{n(\hat{\mu}_f - \mu)} \right]^n \quad -\infty < \hat{\mu}_f < x_{(1)} \]

\[ = \tilde{K}^* \left[ \frac{1}{n\hat{\theta} + m(\hat{\mu}_f - x_{(1)})} \right]^n \quad x_{(1)} < \hat{\mu}_f < \infty \]

where

\[ \tilde{K}^* = \frac{n^n(n - 1)m}{(n + m)} \left( \frac{\hat{\theta}}{\theta} \right)^{n-1}. \]

C  Proof of Theorem 3.3

From Equation 1.1 it follows that

\[ \hat{\theta}_f|\theta \sim \chi^2_{2m-2}\theta \]

which means that

\[ f(\hat{\theta}_f|\theta) = \left( \frac{m}{\theta} \right)^{m-1} \left( \frac{\theta}{\hat{\theta}_f} \right)^{m-2} \exp \left( -\frac{m}{\theta} \hat{\theta}_f \right) \frac{1}{\Gamma(m-1)} \quad 0 < \hat{\theta}_f < \infty \quad \text{(C.1)} \]

The posterior distribution of \( \theta \) (Equation 2.3) is

\[ \hat{\theta}(\theta|\text{data}) = \left( \frac{n\theta}{n - 1} \right)^{n-1} \left( \frac{1}{\theta} \right)^n \exp \left( -\frac{n}{\theta} \right) \quad 0 < \theta < \infty. \]

Therefore

\[ \tilde{f}(\hat{\theta}_f|\text{data}) = \int_0^\infty f(\hat{\theta}_f|\theta) \hat{\theta}(\theta|\text{data}) d\theta \]

\[ = \frac{m^{m-1}}{\Gamma(m-1)\Gamma(n-1)} \left( \frac{\hat{\theta}_f}{\theta} \right)^{m-2} \int_0^\infty \left( \frac{1}{\theta} \right)^{m+n-1} \exp \left\{ -\frac{1}{\theta} \left( m\hat{\theta}_f + n\theta \right) \right\} d\theta \]

\[ = \frac{\Gamma(m+n-2)m^{m-1}(n\hat{\theta})^{n-1}(\hat{\theta}_f)^{m-2}}{\Gamma(m-1)\Gamma(n-1)(m\hat{\theta}_f + n\theta)^{m+n-2}} \quad 0 < \hat{\theta}_f < \infty. \]

D  Proof of Corollary 3.4

\[ \hat{\theta}_f|\theta \sim \chi^2_{2m-2}\theta \]

and

\[ \theta|\text{data} \sim \hat{\theta} \frac{2n}{\chi^2_{2n-2}}. \]
Therefore
\[
\hat{\theta}_f|\text{data} \sim \frac{\chi^2_{2m-2}}{2m} \cdot \frac{2n}{\chi^2_{2n-2}} \hat{\theta}_i
\]
\[
\frac{1}{2m} \cdot \frac{2n - 2}{2m} \hat{\theta}_f \sim F_{2m-2;2n-2}
\]
and
\[
\hat{\theta}_f \sim \frac{\theta}{m} \frac{(m-1)}{(n-1)} F_{2(m-1);2(n-1)}
\]

E Proof of Theorem 6.1

Let
\[
\hat{\theta} = \sum_{i=1}^{m} \hat{\theta}_i
\]
As mentioned in Section 6 (see also Krishnamoorthy and Mathew (2009)) that it is well known that
\[
\hat{\theta}_i \sim \frac{\theta}{2n} \chi^2_{2(n-1)}
\]
which means that
\[
\hat{\theta} \sim \frac{\theta}{2n} \chi^2_{2m(n-1)}
\]
Therefore
\[
f\left(\hat{\theta} | \theta\right) = \left(\frac{n}{\theta}\right)^{m(n-1)} \left(\hat{\theta}\right)^{m(n-1)-1} \exp \left(-\frac{n\hat{\theta}}{\theta}\right) \frac{1}{\Gamma[m(n-1)]} = L\left(\theta|\hat{\theta}\right)
\]
i.e, the likelihood function.
As before we will use as prior \( p(\theta) \propto \theta^{-1} \).
The posterior distribution
\[
p\left(\theta|\hat{\theta}\right) = p\left(\theta|\text{data}\right) \propto L\left(\theta|\hat{\theta}\right) p(\theta)
\]
\[
= \left(\frac{n\theta}{m(n-1)}\right)^{m(n-1)+1} \left(\frac{1}{\Gamma[m(n-1)]}\right)^{m(n-1)+1} \exp \left(-\frac{n\theta}{\theta}\right)
\]
An Inverse Gamma distribution.
F Proof of Theorem 6.2

\[ \hat{\theta}_f | \theta \sim \frac{\theta}{2m} \chi^2_{2(n-1)}. \]

Therefore

\[
f (\hat{\theta}_f | \text{data}) = \int_0^\infty f (\hat{\theta}_f | \theta) p (\theta | \text{data}) \, d\theta
\]

\[
= \int_0^\infty \left( \frac{n}{\theta} \right)^{n-1} \left( \frac{\hat{\theta}_f}{\theta} \right)^{n-2} \exp \left( -\frac{n \hat{\theta}_f}{\theta} \right) \times \]

\[
\frac{(n \hat{\theta})^{m(n-1)}}{\Gamma (m(n-1))} \left( \frac{1}{\theta} \right)^{m(n-1)+1} \exp \left( -\frac{n \hat{\theta}}{\theta} \right) \, d\theta
\]

\[
= \frac{(n)^{n-1} \left( \frac{\hat{\theta}_f}{\theta} \right)^{n-2} \theta^{m(n-1)}}{\Gamma (m(n-1)) \left[ \Gamma (m(n-1)) \right]} \int_0^\infty \left( \frac{1}{\theta} \right)^{m(n-1)+1} \exp \left\{ -\frac{n}{\theta} \left[ \hat{\theta}_f + \hat{\theta} \right] \right\} \, d\theta
\]

\[
= \frac{\Gamma [m(n-1)+n-1]}{\Gamma (m(n-1)) \left[ \Gamma (m(n-1)) \right]} \left( \frac{\hat{\theta}_f}{\theta + \hat{\theta}} \right)^{m(n-1)+n} \hat{\theta}_f > 0
\]

From this it follows that

\[ \hat{\theta}_f | \text{data} \sim \frac{\hat{\theta}}{m} F_{2(n-1);2m(n-1)} \]

where

\[ \hat{\theta} = \sum_{i=1}^m \hat{\theta}_i. \]

G Proof of Theorem 6.3

For given \( \theta \)

\[
\psi (\theta) = p \left( \hat{\theta}_f \leq \frac{\hat{\theta}}{m} F_{2(n-1);2m(n-1)} (\beta) \right)
\]

\[
= p \left( \frac{\theta}{2 \theta} \chi^2_{2(n-1)} \leq \frac{\hat{\theta}}{m} F_{2(n-1);2m(n-1)} (\beta) \right) \quad \text{given } \theta
\]

\[
= p \left( \frac{2n \hat{\theta}}{\chi^2_{2m(n-1)}} \chi^2_{2(n-1)} \leq \frac{\hat{\theta}}{m} F_{2(n-1);2m(n-1)} (\beta) \right) \quad \text{given } \chi^2_{2m(n-1)}
\]

\[
= p \left( \chi^2_{2(n-1)} \leq \frac{\chi^2_{2m(n-1)}}{m} F_{2(n-1);2m(n-1)} (\beta) \right) \quad \text{given } \chi^2_{2m(n-1)}
\]

\[
= \psi \left( \chi^2_{2m(n-1)} \right)
\]