Estimation of the Parameters of Burr Distribution Based on Order Statistics

Munir Ahmad
ESTIMATION OF THE PARAMETERS OF BURR DISTRIBUTION
BASED ON ORDER STATISTICS

Munir Ahmad
University of Petroleum and Minerals
Dhahran, Saudi Arabia.

Abstract

Asymptotically best unbiased estimators of the two Burr parameters $\alpha$ and $\beta$ based on some selected order statistics from a sample of size $n$ are considered. It is shown that the minimum variance of the estimator of $\beta$ for the known $\alpha$ is $1.544168^2/n$ with efficiency of about 64% as compared to the maximum likelihood estimator of $\beta$. The minimum variance of the estimator of $\alpha$ for known $\beta$ is attained when 76th sample order statistics is used. When $\alpha$ and $\beta$ are unknown, variances are essentially minimized using three order statistics (.24,.54,.77) for $\alpha$ and (.27,.61,.81) for $\beta$. An example is given to illustrate the method.

1. Introduction

Consider a random sample of size $n$ from the Burr distribution
first given by Burr (1942) having the probability density function

\[ f(x) = \alpha \beta x^{\alpha-1}/(1 + x^\alpha)^{\beta + 1}, \quad x > 0, \quad \beta > 1 \] (1)

and the distribution function

\[ F(x) = 1 - (1 + x^\alpha)^{-\beta}, \quad x > 0 \] (2)

Properties of the distribution (1) have been investigated by Austin (1971), Burr (1968), Burr and Cislak (1968) and Hatke (1949). Recently Austin (1971) has applied the distribution to the control charts for the maximum and minimum value in sampling from a normal distribution.

In modeling tensile loads on a machine element, Weibull and log normal distributions have been tried (see Bury, 1975). In this paper we have used some order statistics to estimate the Burr parameters for use in the modeling of tensile loads on a machine element.

2. Estimation

Let \( x_1 \leq x_2 \leq \cdots \leq x_n \) be the sample ordered statistics and let \( u_1, u_2, \cdots, u_n \) be the corresponding population quantities. Let \( x_{n_1} < x_{n_2} \cdots < x_{n_k}, \quad k = 1, 2, \cdots \) be the selected \( k \) sample
quantities where \( k < n \). A set of \( k \) real numbers such that
\[
0 = \lambda_0 < \lambda_1 < \cdots < \lambda_{k+1} = 1
\]
is called a spacing. We define \( \lambda_i \) by \( \lambda_i = F(u_i) \), \( i = 1, 2, \cdots, k \). Then \( u_i = F^{-1}(\lambda_i) \) and therefore from (2)
\[
u_i = [(1 - \lambda_i)^{-1/\beta} - 1]^{1/\alpha}
\] (3)
and from (1)
\[
f[F^{-1}(\lambda_i)] = \alpha \beta (1 - \lambda_i)^{1+1/\beta} u_i^{\alpha-1}.
\] (4)
Taking the logarithm of (3) and (4), we get
\[
\ln u_i = \frac{1}{\alpha} \ln [(1 - \lambda_i)^{-1/\beta} - 1]
\] (5)
and
\[
\ln f[F^{-1}(\lambda_i)] = \ln(\alpha \beta) + (1 + 1/\beta) \ln(1 - \lambda_i) + (\alpha - 1) \ln u_i
\] (6)

Using some selected order statistics, we shall obtain best unbiased estimators of \( \alpha \) and \( \beta \) by minimizing the asymptotic variances of the estimators.

**Case 1.** (\( \beta \) is known): Given a spacing \( \{\lambda_i\} \), the estimator of \( \alpha \) when \( \beta = \beta_0 \), using (3) is
\[ \hat{a}_k = (k)^{-1} \sum_{i=1}^{k} \{ \ln[(1 - \lambda_i)^{-1/\beta_0} - 1]y_i \} = (k)^{-1} \sum_{i=1}^{k} a_i y_i \]  \hspace{1cm} (7)

where \( y_i = (\ln x_{n_i})^{-1} \), \( a_i = \ln[(1 - \lambda_i)^{-1/\beta_0} - 1] \) and \( x_{n_i} \) is a selected sample ordered statistic corresponding to \( u_{n_i} \). The asymptotic variance of \( \hat{a}_k \) is given by

\[ \text{var}(\hat{a}_k) = k^{-2} \sum_{i} a_i^2 \text{var}(y_i) + \sum_{i \neq j} a_i a_j \text{cov}(y_i, y_j) \]  \hspace{1cm} (8)

where \( \text{var}(y_i) = y_i^2 x_{n_i}^{-2} \text{var}(x_{n_i}) \)

and \( \text{cov}(y_i, y_j) = y_i^2 y_j^2 (x_{n_i} x_{n_j})^{-1} \text{cov}(x_{n_i}, x_{n_j}) \).

It is well known [see Mosteller (1946)] that given a spacing, the joint distribution of the \( k \) sample quantities is asymptotically normal.

Suppose \( k = 1 \). The \( i \)th sample quantity is asymptotically normal with mean \( u_i \) and variance \( \lambda_i (1 - \lambda_i)/[n \text{var}(x_{n_i})] \). Minimizing \( \text{var}(\hat{a}_i) \) with respect to \( \lambda_i \), we get

\[ (1 - 2\lambda_i)(1 - (1 - \lambda_i)^{1/\beta_0}) \ln[(1 - \lambda_i)^{-1/\beta_0} - 1] + \lambda_i / \beta_0 = 0 \]  \hspace{1cm} (9)

Table 1 gives the solution of \( \lambda_i \) of equation (9) that minimizes
var(\hat{a}_1) for some given values of \( \beta \).

<table>
<thead>
<tr>
<th>( \beta )</th>
<th>( \lambda_i )</th>
<th>( a_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>.023950</td>
<td>1.54340</td>
</tr>
<tr>
<td>1.5</td>
<td>.669111</td>
<td>1.05748</td>
</tr>
<tr>
<td>2.0</td>
<td>.904915</td>
<td>0.80773</td>
</tr>
<tr>
<td>2.5</td>
<td>.931742</td>
<td>0.65567</td>
</tr>
<tr>
<td>3.0</td>
<td>.951335</td>
<td>0.55332</td>
</tr>
<tr>
<td>3.5</td>
<td>.965447</td>
<td>0.47972</td>
</tr>
<tr>
<td>4.0</td>
<td>.975530</td>
<td>0.42421</td>
</tr>
<tr>
<td>5.0</td>
<td>.987780</td>
<td>0.34583</td>
</tr>
<tr>
<td>6.0</td>
<td>.993912</td>
<td>0.29282</td>
</tr>
<tr>
<td>7.0</td>
<td>.996969</td>
<td>0.25439</td>
</tr>
<tr>
<td>8.0</td>
<td>.998491</td>
<td>0.22514</td>
</tr>
<tr>
<td>9.0</td>
<td>.999248</td>
<td>0.20192</td>
</tr>
<tr>
<td>10.0</td>
<td>.999626</td>
<td>0.18355</td>
</tr>
</tbody>
</table>

In case of large \( \beta \), the largest observation may be used to estimate \( a \). For \( k = 2, 3, \ldots \), \( \lambda_5 \) can be determined by minimizing (8) numerically.
Case 2. (α is known): Suppose α = α₀. For a given spacing \( \{\lambda_i\} \) satisfying (2), we find that

\[
\hat{\beta}_k = k^{-1} \sum_{i=1}^{k} b_i z_i
\]

where \( b_i = \ln(1 - \lambda_i) \) and \( z_i = -(\ln(1 + x_{n_i}^{\alpha_0}))^{-1} \).

Its asymptotic variance is given by

\[
\text{var}(\hat{\beta}_k) = k^{-2} \left[ \sum_{i} b_i^2 \text{var}(z_i) + \sum_{i \neq j} b_i b_j \text{cov}(z_i, z_j) \right]
\]

where \( \text{var}(z_i) = z_i^2 \left[ \beta(1 - \lambda_i) \right]^{-2} \text{var}(x_{n_i}) \), \( f_i = f(x_{n_i}) \)

and \( \text{cov}(z_i, z_j) = z_i^2 z_j^2 f_i f_j \beta^{-2} \left[ (1 - \lambda_i)(1 - \lambda_j) \right]^{-1} \text{cov}(x_{n_i}, x_{n_j}) \).

Let \( k = i \), \( \hat{\beta}_i = b_i z_i \) and the asymptotic variance is given by

\[
\text{var}(\hat{\beta}_i) = \left[ \frac{\ln(1 - \lambda_i)}{\ln^2(1 + x_i^{\alpha_0})} \frac{\alpha_0 x_i^{\alpha_0-1}}{1 + x_i^{\alpha_0}} \right]^2 \text{var}(x_i)
\]

(11)
Using \( \text{var}(x_i) = \lambda_i (1 - \lambda_i)/nf_i^2 \),

\[
f_i = \alpha_0 \beta x_i^{\alpha_0 - 1} (1 - \lambda_i)/(1 + x_i^{\alpha_0}) \text{ and } \ln(1 + x_i^{\alpha_0}) = -(1/\beta) \ln(1 - \lambda_i).
\]

We have

\[
\text{var}(\hat{\lambda}_i) = \beta^2 \lambda_i [\ln(1 - \lambda_i) \ln^2(1 - \lambda_i)]
\]

(12)

We choose \( \lambda_i \) by minimizing \( \text{var}(\hat{\lambda}_i) \) with respect to \( \lambda_i \) and derive an equation in \( \lambda_i \) only

\[
2\lambda_i + \ln(1 - \lambda_i) = 0
\]

giving a value of \( \lambda_i = 0.79681 \).

The minimum variance of \( \hat{\beta} \) for any \( \alpha_0 \) is \( 1.54416 \beta^2/n \). The maximum likelihood estimate of \( \hat{\beta} \) for \( \alpha = \alpha_0 \) from a sample of size \( n \) is

\[
\hat{\beta} = \frac{n}{\sum_{i=1}^{n} \ln(1 + x_i^{\alpha})}
\]

and its asymptotic variance is \( \beta^2/n \). The efficiency of \( \hat{\beta} \) as compared to \( \hat{\beta} \) is given by

\[
E = (1 - \lambda_i) \ln (1 - \lambda_i)/\lambda_i
\]
and is about 65% for \( \lambda_i = 0.79681 \).

For \( k = 2, 3, \ldots \), \( \lambda_i \) can be determined such that (11) is minimized. Some special cases have been investigated separately.

**Case 3.** (\( \alpha \) and \( \beta \) are unknown): Suppose \( \alpha \) and \( \beta \) are unknown.

We write \( y_{n_i} = \ln x_{n_i} = \frac{1}{\alpha}[(1 - \lambda_i)^{1/\beta} - 1] \) and

\[
\ln\tilde{f}_{n_i} = \ln\tilde{\alpha} + \ln\tilde{\beta} + (\alpha - 1)y_{n_i} + (1 + 1/\tilde{\beta})\ln(1 - \lambda_i), \ i = 1, 2, 3.
\]

Using \( \ln\tilde{f}_{n_i}, \ i = 1, 2, 3, \) we have

\[
\ell_1 = (\tilde{\alpha} - 1)y_1 + (1 + 1/\tilde{\beta})\alpha_1
\]

\[
\ell_2 = (\tilde{\alpha} - 1)y_2 + (1 + 1/\tilde{\beta})\alpha_2
\]

where \( \ell_1 = \ln(f_{n_1}/f_{n_2}), \ \ell_2 = \ln(f_{n_3}/f_{n_2}) \)

\[
\alpha_1 = \ln[(1 - \lambda_1)/(1 - \lambda_2)], \ \alpha_2 = \ln[(1 - \lambda_2)/(1 - \lambda_3)]
\]

\[
y_1 = y_{n_1} - y_{n_2}, \ \ y_2 = y_{n_2} - y_{n_3} \ \text{and} \ \lambda_i = i(x_{n_i})
\]

Solving for \( \tilde{\alpha} \) and \( \tilde{\beta} \), we get the estimator of \( (\alpha, \beta) \) as
\[ \tilde{\alpha} = 1 + (\ell_1 a_2 - \ell_2 a_1)/(y_1 a_2 - y_2 a_1) \] (13)

and

\[ \tilde{\beta} = (u_1 y_2 - u_2 y_1)/[(\ell_1 - \alpha_1)y_2 - (\ell_2 - \alpha_2)y_1] \] (14)

The asymptotic variances of \( \tilde{\alpha} \) and \( \tilde{\beta} \) are

\[ \text{var}(\tilde{\alpha}) = (\alpha - 1)^2 \sum_{i,j=1}^{4,4} a_i a_j \text{C}(n_i, n_j) \] (15)

and

\[ \text{var}(\tilde{\beta}) = \beta^2 \sum_{i,j=1}^{4,4} b_i b_j \text{C}(n_i, n_j) \] (16)

where \( \text{C}(n_i, n_j) = \text{var}(n_i), j = 1 \)

\[ \text{Cov}(n_i, n_j), j \neq i \]

\[ n_1 = \ell_1, \; n_2 = \ell_2, \; n_3 = y_1, \; n_4 = y_2, \; a_1 = a_2/(\ell_1 a_2 - \ell_2 a_1) \]

\[ a_2 = -a_1/(\ell_1 a_2 - \ell_2 a_1) \]

\[ a_3 = -a_2/(y_1 a_2 - y_2 a_1) \]

\[ a_4 = a_1/(y_1 a_2 - y_2 a_1) \]
\[ b_1 = -\frac{y_2}{((\xi_1 - \alpha_1) y_2 - (\xi_2 - \alpha_2) y_1)} \]

\[ b_2 = \frac{y_1}{((\xi_1 - \alpha_2) y_2 - (\xi_2 - \alpha_2) y_1)} \]

\[ b_3 = \frac{(\xi_2 - \alpha_2)/((\xi_1 - \alpha_1) y_2 - (\xi_2 - \alpha_2) y_1)} - a_3 \]

\[ b_4 = -a_4 - \frac{(\xi_1 - \alpha_1)/((\xi_1 - \alpha_1) y_2 - (\xi_2 - \alpha_2) y_1)} \]

Minimum variance can be obtained only numerically using variance values of unknown quantities involved in the expressions (15) and (16). For various values of \((\lambda_i)\) the variance functions at (15) and (16) were computed on the IBM 370 computer. For \(k = 3\), we find that the variance functions are essentially minimized by choosing \((0.24, 0.54, 0.77)\) for \(\text{var}(\tilde{a})\) and by choosing \((0.27, 0.61, 0.81)\) for \(\text{var}(\tilde{b})\).

3. Example

To illustrate the method, suppose that the machine element is subjected to 100 random tensile loads during its mission \(x\) and that the load has been represented by the Burr model with \(\alpha = 6\) and \(\beta = 1\). If \(\beta\) is known to have a value of 4, then \(\alpha\) is estimated from (7) when \(k = 1\) and \(\lambda_1 = 0.97553\) (see table 1). The values of \(x_{n_i}\) and \(y_i\) are
\[ x_{.98} = 1.075, \quad y_i = 0.93023 \text{ and } a_i = 0.42421. \quad \hat{\alpha} = 5.87 \text{ and } \var(\hat{\alpha}) = 13.07. \] If \( \alpha \) is known to have a value of 6, then \( \beta \) is estimated from (10) when \( k = 1, \quad \lambda_i = 0.79681. \) The values of \( x_{n_1} = 0.853, \quad z_i = -3.0689 \text{ and } b_i = -1.5936. \) Then \( \hat{\beta} = 4.89 \text{ and } \var(\hat{\beta}) = 0.3696. \) In this case, the maximum likelihood estimation of \( \beta \) is \( \hat{\beta} = 4.23 \text{ and } \var(\hat{\beta}) = 0.1786. \) If \( \beta = 4, \) then efficiency of \( \hat{\beta} \) is about 65\%. If we use \( \hat{\alpha} \) and \( \hat{\beta} \) in the variance formulae respectively the efficiency is about 48.3\%. If \( \alpha \) and \( \beta \) are unknown then using (13) and (14) as estimating equations we have for \( \alpha, \) the values of \( x_{n_1}, x_{n_2} \) and \( x_{n_3} \) are \( x_{n_1} = .585, \quad x_{n_2} = .744 \text{ and } x_{n_3} = .840. \) Then \( \hat{\alpha} = 5.53 \text{ and } \var(\hat{\beta}) = 20.31 \text{ and for } \beta, \) the values of \( x_{n_1}, x_{n_2} \) and \( x_{n_3} \) are \( x_{n_1} = .612, \quad x_{n_2} = .772 \text{ and } x_{n_3} = .878. \) Then \( \hat{\beta} = 3.42 \text{ and } \var(\hat{\beta}) = .2511. \)

4. References


