J. King Saud Univ., Vol. 14, *Admin. Sci.* (1), pp. 1-9, Riyadh (1422/2002)

Estimation in a Doubly Truncated Burr Distribution

M.H. Al-Yousef
Professor, Department of Quantitative Methods
College of Administrative Sciences, King Saud University,
Riyadh, Saudi Arabia

(Received 4-2-1421; accepted for publication 19-8-1421)

Abstract. In this article we discuss the problem of estimating the parameters of the doubly truncated Burr distribution when truncation points are unknown. The estimation is based on type II censored data. Maximum likelihood estimators and approximate estimator of the variance-covariance matrix are derived.

1. Introduction

Truncated distributions arise when sample selection and /or observation is not possible in some subregion of the sample space. This can occur as a consequence of actual elimination of part of the original data.

Charernkavanich and Cohen [1] discussed this problem for complete samples with a variety of estimation problems involving truncated normal, gamma, Weibull, lognormal and various other truncated distributions. Bain and Weeks [2], and Deemer and Votaw [3] gave the main results in the case of truncated exponential distribution using censored data; Shalaby [4], and Al-Yousef [5,6] have discussed the problem in the case of Weibull, Gompertz and logistic distributions.

The present study is concerned with a Burr distribution that is doubly truncated at unknown truncation points with censored type II data. The truncation points then become additional parameters which must be estimated from sample data along with the primary distribution parameters.

2. The Burr Distribution

The Burr distribution was first introduced in the literature by Burr [7]. Burr and Cesłak [8], and Burr [9] have shown that if one chooses the parameters appropriately, the Burr distribution covers a large portion of the curve shape characteristics of type I, IV, VI in the Pearson family of distributions. Thus the use of the Burr distribution as a failure model is appropriate and useful in applied statistics, specially in survival analysis and actuarial studies.

The probability density function (p.d.f.) of this distribution can be written as:

$$f(x;r,s) = rsx^{s-1}(1+x^{s})^{-(r+1)} \tag{2.1}$$

for : $0 \le x < \infty; r > 0, s > 0$ (0 otherwise).

The corresponding cumulative distribution function (c.d.f.) is:

$$F(x;r,s) = 1-(1+x^{s})^{-r} \quad ; x \ge 0 \tag{2.2}$$

The mth moment of this distribution exists for m < rs, and we have :

$E(X^{m}) = ms^{-1}B_m \; ; m < rs$

where : $B_m = B(\frac{m}{s}, r - \frac{m}{s})$, and B(u, v) is Beta function defined by : $B(u,v) = \int_0^{\infty} t^{u-1}(1+t)^{-(u+v)}dt$

The mean μ, variance σ^2 and third standard moment α_3 of the distribution (2.1) are :

$\mu = s^{-1}B_1$, $\sigma^2 = 2s^{-1}B_2 - s^{-2}B_1^{2}$

$$\alpha_3 = \frac{3s^2B_3 - 6sB_1B_2 + B_1^{3}}{[2sB_2 - B_1^{2}]^{3/2}} \tag{2.3}$$

3. Doubly Truncated Burr Distribution

When the distribution with p.d.f.(2.1) is doubly truncated over the interval [a,b], the resulting truncated distribution becomes:

$$f_T(x;a,b,r,s) = f(x;r,s)[F(b;r,s) - F(a;r,s)]^{-1} \tag{3.1}$$

for $a \le x \le b$ (0 otherwise). And it follows from (2.1) and (2.2) that :

$$f_T(x;a,b,r,s) = rs(u-v)^{-1}x^{s-1}(1+x^{s})^{-(r+1)} \tag{3.2}$$

for $a \le x \le b$; zero elsewhere. Where;

$$u = (1+a^{s})^{-r} \quad \text{and} \quad v = (1+b^{s})^{-r} \tag{3.3}$$

Suppose that a sample of size n with failure time distribution given by (3.2) has been subjected to life testing and that the test is terminated at the time that the k th failure becomes available and k n. The likelihood function of this type of censoring is given by:

$$L(x_1, x_2, \ldots x_n; a, b, r, s) = r^k s^k (u - v)^{-n} \prod_{j=1}^{k} y_j^{s-1} (1 + y_j^s)^{-(r+1)} [(1 + y_k^s)^{-r} - v]^{n-k} \tag{3.4}$$

where: $y_j = x_{(j)}$ the jth order statistic in a random sample of size n.

3.1.Maximum likelihood estimation (M.L.E.)

On taking logarithms of (3.4) and differentiating we have:

$$\frac{\partial \ln L}{\partial a} = nrsu\,(u - v)^{-1} a^{-1} (1 + a^{-s})^{-1}$$

$$\frac{\partial \ln L}{\partial b} = -rsv\,\{n(u - v)^{-1} - (n - k)[(1 + y_k^s)^{-r} - v]^{-1}\} b^{-1} (1 + b^{-s})^{-1}$$

$$\frac{\partial \ln L}{\partial r} = kr^{-1} - n(u - v)^{-1} [v \ln(1 + b^s) - u \ln(1 + a^s)] + (n - k)[(1 + y_k^s)^{-r}$$

$$- v]^{-1} [v \ln(1 + b^s) - (1 + y_k^s)^{-r} \ln(1 + y_k^s)] - \sum_{j=1}^{k} \ln(1 + y_j^s) \tag{3.5}$$

$$\frac{\partial \ln L}{\partial s} = ks^{-1} - nr(u - v)^{-1} [v(1 + b^{-s})^{-1} \ln b - u(1 + a^{-s})^{-1} \ln a] + (n - k)r[(1 + y_k^s)^{-r} - v]^{-1}$$

$$[v(1 + b^{-s})^{-1} \ln b - (1 + y_k^s)^{-r} (1 + y_k^{-s})^{-1} \ln y_k] + \sum_{j=1}^{k} \ln y_j - (r + 1) \sum_{j=1}^{k} (1 + y_j^{-s})^{-1} \ln y_j$$

The function (3.4) attains its maximum value when a is as large as possible. Since

$$a \le x; \text{we then have } \hat{a} = Y_1 \tag{3.6}$$

where Y_1 is the smallest sample observation; i.e. Y_1 is the first order statistic in a random sample of size n. Estimators for b, r and s follow as solutions of equations:

$$n(\hat{u} - \hat{v})^{-1} - (n - k)\left[\left(1 + y_k^{\hat{s}}\right)^{-\hat{r}} - \hat{v}\right]^{-1} = 0$$

$$k\hat{r}^{-1} - n(\hat{u} - \hat{v})^{-1}\left[\hat{v} \ln\left(1 + \hat{b}^{\hat{s}}\right) - \hat{u} \ln\left(1 + y_1^{\hat{s}}\right)\right] + (n - k)\left[\left(1 + y_k^{\hat{s}}\right)^{-\hat{r}} - \hat{v}\right]^{-1}.$$

$$\left[\hat{v} \ln\left(1 + \hat{b}^{\hat{s}}\right) - \left(1 + y_k^{\hat{s}}\right)^{-\hat{r}} \ln\left(1 + y_k^{\hat{s}}\right)\right] - \sum_{j=1}^{k} \ln\left(1 + y_j^{\hat{s}}\right) = 0$$

$$k\hat{s}^{-1} - n\hat{r}(\hat{u}-\hat{v})^{-1}\left[\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\ln\hat{b} - \hat{u}\left(1+y_1{}^{-\hat{s}}\right)^{-1}\ln y_1 + (n-k)\hat{r}\left(1+y_k{}^{\hat{s}}\right)^{-\hat{r}} - \hat{v}\right]^{-1}.$$

$$\left[\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\ln\hat{b} - \left(1+y_k{}^{\hat{s}}\right)^{-\hat{r}}\left(1+y_k{}^{-\hat{s}}\right)^{-1}\ln y_k\right] + \sum_{j=1}^{k}\ln y_j - (\hat{r}+1)\sum_{j=1}^{k}\left(1+y_j{}^{-\hat{s}}\right)^{-1}\ln y_j = 0 \quad (3.7)$$

where : $\hat{u} = \left(1+y_1{}^{\hat{s}}\right)^{-\hat{r}}$ and $\hat{v} = \left(1+\hat{b}^{\hat{s}}\right)^{-\hat{r}}$

The solution of the set of equations given in (3.7) can be obtained using iterative techniques for solving a system of simultaneous non-linear equations in three unknowns.

3.2 Sampling errors

The p.d.f. of the j th order statistic in random sample of size n from the truncated distribution (3.2) is:

$$g(y_j;a,b,r,s) = rsj\binom{n}{j}(u-v)^{-n}[u-(1+y_j{}^{s})^{-r}]^{j-1}[(1+y_j{}^{s})^{-r}-v]^{n-j}y_j{}^{s-1}$$
$$(1+y_j{}^{s})^{-(r+1)} \quad (3.8)$$
$$a \le y_j \le b \quad \text{, zero elsewhere.}$$

if $j = 1$, the p.d.f. of the first order statistic may be obtained as:

$$g(y_1;a,b,r,s) = rsn(u-v)^{-n}[(1+y_1{}^{s})^{-r}-v]^{n-1}y_1{}^{s-1}(1+y_1{}^{s})^{-(r+1)} \quad (3.9)$$

It can be shown that

$$E(Y_1^j) = L_j \quad (3.10)$$

where:

$$L_j = n(u-v)^{-n}\int_v^u (t^{\frac{-1}{r}}-1)^{\frac{j}{s}}(t-v)^{n-1}dt \quad (3.11)$$

The variance of the M.L.E.(3.6) can be found as:

$$V(\hat{a}) = L_2 - L_1{}^2 \quad (3.12)$$

The asymptotic variance – covariance matrix of M.L.E. of $\hat{b}, \hat{r}$ and $\hat{s}$ can be expressed as:

$$V(\hat{b}, \hat{r}, \hat{s}) = [Z_{ij}]^{-1}; i, j = 1,2,3$$

where

$$Z_{11} = -E\left(\frac{\partial^2 \ln L}{\partial b^2}\right), \; Z_{12} = Z_{21} = -E\left(\frac{\partial^2 \ln L}{\partial b \partial r}\right)$$

$$Z_{13} = Z_{31} = -E\left(\frac{\partial^2 \ln L}{\partial b \partial s}\right), \; Z_{22} = -E\left(\frac{\partial^2 \ln L}{\partial r^2}\right)$$

$$Z_{23} = Z_{32} = -E\left(\frac{\partial^2 \ln L}{\partial r \partial s}\right), \; Z_{33} = -E\left(\frac{\partial^2 \ln L}{\partial s^2}\right)$$

The exact expressions for the expectations in the above matrix are difficult to obtain. However an approximate estimate for variance-covariance matrix can be obtained using the approximation of Cohen [10], namely;

$$Z_{ij} \cong \hat{Z}_{ij} \quad ; i, j = 1,2,3$$

$$\hat{Z}_{11} = [-\frac{\partial^2 \ln L}{\partial b^2}]_{b=\hat{b}, r=\hat{r}, s=\hat{s}, a=y_1}, \; \hat{Z}_{12} = [-\frac{\partial^2 \ln L}{\partial b \partial r}]_{b=\hat{b}, r=\hat{r}, s=\hat{s}, a=y_1}$$

$$\hat{Z}_{13} = [-\frac{\partial^2 \ln L}{\partial b \partial s}]_{b=\hat{b}, r=\hat{r}, s=\hat{s}, a=y_1}$$

$$\hat{Z}_{22} = [-\frac{\partial^2 \ln L}{\partial r^2}]_{b=\hat{b}, r=\hat{r}, s=\hat{s}, a=y_1}, \; \hat{Z}_{23} = [-\frac{\partial^2 \ln L}{\partial r \partial s}]_{b=\hat{b}, r=\hat{r}, s=\hat{s}, a=y_1}$$

$$\hat{Z}_{33} = [-\frac{\partial^2 \ln L}{\partial s^2}]_{b=\hat{b}, r=\hat{r}, s=\hat{s}, a=y_1}$$

These approximations are derived to be:

$$\hat{Z}_{11} = \hat{r}^2 \hat{s}^2 \hat{v}^2 \hat{b}^{-2} (1+\hat{b}^{-\hat{s}})^{-2} \{ n(\hat{u}-\hat{v})^{-2} - (n-k)[(1+y_k{}^{\hat{s}})^{-\hat{r}} - \hat{v}]^{-2} \}$$

$$\hat{Z}_{12} = \hat{r}\hat{s}\hat{v}\hat{b}^{-1}\left(1+\hat{b}^{-\hat{s}}\right)^{-1} \left\{ \begin{array}{l} \left[\hat{u} \ln\left(1+y_1{}^{\hat{s}}\right) - \hat{v} \ln\left(1+\hat{b}^{\hat{s}}\right)\right] - (n-k)\left[\left(1+y_k{}^{\hat{s}}\right)^{-\hat{r}} - \hat{v}\right]^{-2} . \\ \left[\left(1+y_k{}^{\hat{s}}\right)^{-\hat{r}} \ln\left(1+y_k{}^{\hat{s}}\right) - \hat{v} \ln\left(1+\hat{b}^{\hat{s}}\right)\right] \end{array} \right\}$$

$$\hat{Z}_{13}=\hat{r}^2\hat{s}\hat{v}\hat{b}^{-1}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\left\{\begin{array}{l}(n-k)\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}\left(1+y_k^{-\hat{s}}\right)^{-1}\ln y_k-\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\ln\hat{b}\right]\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}-\hat{v}\right]^{-2}\\ \quad -n\left[\hat{u}\left(1+y_1^{-\hat{s}}\right)^{-1}\ln y_1-\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\ln\hat{b}\right](\hat{u}-\hat{v})^{-2}\end{array}\right\}$$

$$\hat{Z}_{22}=k\hat{r}^{-2}-n(\hat{u}-\hat{v})^{-2}\left[\hat{v}\ln\left(1+\hat{b}^{\hat{s}}\right)-\hat{u}\ln\left(1+y_1^{\hat{s}}\right)\right]^2+n(\hat{u}-\hat{v})^{-1}\left[\hat{u}\ln^2\left(1+y_1^{\hat{s}}\right)-\hat{v}\ln^2\left(1+\hat{b}^{\hat{s}}\right)\right]+$$
$$(n-k)\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}-\hat{v}\right]^{-2}\left[\hat{v}\ln\left(1+\hat{b}^{\hat{s}}\right)-\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}\ln\left(1+y_k^{\hat{s}}\right)\right]^2-(n-k)\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}-\hat{v}\right]^{-1}$$
$$\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}\ln^2\left(1+y_k^{\hat{s}}\right)-\hat{v}\ln^2\left(1+\hat{b}^{\hat{s}}\right)\right]$$

$$\hat{Z}_{23}=n(\hat{u}-\hat{v})^{-1}\left\{\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\left[1-\hat{r}\ln\left(1+\hat{b}^{\hat{s}}\right)\right]\ln\hat{b}-\hat{u}\left(1+y_1^{-\hat{s}}\right)^{-1}\left[1-\hat{r}\ln\left(1+y_1^{\hat{s}}\right)\right]\ln y_1\right\}-n\hat{r}(\hat{u}-\hat{v})^{-2}.$$
$$\left[\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\ln\hat{b}-\hat{u}\left(1+y_1^{-\hat{s}}\right)^{-1}\ln y_1\right]\left[\hat{v}\ln\left(1+\hat{b}^{\hat{s}}\right)-\hat{u}\ln\left(1+y_1^{\hat{s}}\right)\right]-(n-k)\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}-\hat{v}\right]^{-1}.$$
$$\left\{\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\left[1-\hat{r}\ln\left(1+\hat{b}^{\hat{s}}\right)\right]\ln\hat{b}-\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}\left(1+y_k^{-\hat{s}}\right)^{-1}\left[1-\hat{r}\ln\left(1+y_k^{\hat{s}}\right)\right]\ln y_k\right\}+(n-k)\hat{r}.$$
$$\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}-\hat{v}\right]^{-2}\left[\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\ln\hat{b}-\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}\left(1+y_k^{-\hat{s}}\right)^{-1}\ln y_k\right]\left[\hat{v}\ln\left(1+\hat{b}^{\hat{s}}\right)-\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}\ln\left(1+y_k^{\hat{s}}\right)\right]$$
$$-\sum_{j=1}^{k}\left(1+y_j^{-\hat{s}}\right)^{-1}\ln y_j$$

$$\hat{Z}_{33}=k\hat{s}^{-2}-n\hat{r}^2(\hat{u}-\hat{v})^{-2}\left[\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\ln\hat{b}-\hat{u}\left(1+y_1^{-\hat{s}}\right)^{-1}\ln y_1\right]^2-n\hat{r}(\hat{u}-\hat{v})^{-1}.$$
$$\left[\hat{u}y_1^{\hat{s}}\left(1-y_1^{\hat{s}}\right)\left(1+y_1^{\hat{s}}\right)^{-2}\ln^2 y_1-\hat{v}\hat{b}^{\hat{s}}\left(1-\hat{b}^{\hat{s}}\right)\left(1+\hat{b}^{\hat{s}}\right)^{-2}\ln^2\hat{b}\right]+(n-k)\hat{r}^2\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}-\hat{v}\right]^{-2}.$$
$$\left[\hat{v}\left(1+\hat{b}^{-\hat{s}}\right)^{-1}\ln\hat{b}-\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}\left(1+y_k^{-\hat{s}}\right)^{-1}\ln y_k\right]^2+(n-k)\hat{r}\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}-\hat{v}\right]^{-1}.$$
$$\left[\left(1+y_k^{\hat{s}}\right)^{-\hat{r}}y_k^{\hat{s}}\left(1-y_k^{\hat{s}}\right)\left(1+y_k^{\hat{s}}\right)^{-2}\ln^2 y_k-\hat{v}\hat{b}^{\hat{s}}\left(1-\hat{b}^{\hat{s}}\right)\left(1+\hat{b}^{\hat{s}}\right)^{-2}\ln^2\hat{b}\right]+$$
$$(\hat{r}+1)\sum_{j=1}^{k}y_j^{\hat{s}}\left(1+y_j^{\hat{s}}\right)^{-2}\ln^2 y_j$$

4. Illustrative Example

The practical application of estimators resulting in this work are illustrated with simulated data from the doubly truncated Burr distribution in which: a=5,b=20,r=1and s=2.

The individual observations for a random sample of size 40 are listed in Table 1.

Table 1. A random sample of 40 observations from
$f_T(x,a,b,r,s) = f_T(x,5,20,1,2)$

7.64	8.98	5.20	6.25	5.65	7.08	5.49	8.09
5.57	15.50	7.44	5.13	5.06	6.50	7.85	7.25
10.66	5.27	8.67	5.74	11.92	5.41	6.92	19.15
6.37	5.83	9.29	9.68	8.37	6.64	5.93	14.11
6.78	6.14	11.23	17.26	5.34	12.80	10.15	6.03

Estimators calculations for the case of complete sample and the case of censored sample when k=30 are summarized in Table 2. We note that from our data, the covariance between these estimators are very small (less than 0.001), which indicate that the estimators introduced in this work can be approximated by the normal distribution. The asymptotic confidence intervals (c.i.) for the actual parameters are calculated and given in Table 2.

Table 2. Estimators from doubly truncated Burr distribution

Case:	Complete sample	Censored sample k=30
E.of a	5.06	5.06
V.of a	3.68E-2	6.53E-1
95% c.i.	[4.94,5.18]	[3.48,6.64]
E.of b	19.15	18.19
V.of b	8.53E-1	9.58E-1
95% c.i.	[17.34,20.96]	[16.27,20.11]
E.of r	0.97564	1.02477
V.of r	2.36E-3	3.54E-2
95%c.i.	[0.88,1.07]	[0.66,1.39]
E.of s	1.95183	1.83942
V.of s	3.76E-2	5.36E-2
95% c.i.	[1.57,2.33]	[1.39,2.29]

It is of interest to note the following special cases:

- If b $\rightarrow \infty$, we have the case of left truncated Burr distribution.
- If a=0, we have the case of right truncated Burr distribution.
- If k=n, we have the problem of estimation for complete sample.
- The likelihood function when the parent distribution is two parameters Burr can be obtained as a special case from (3.4) if u=1 and v=0.

References

[1] Charernkavanich, D. and Cohen, A.C. "Estimation in the Sngly Truncated Weibull Distribution with an unknown Truncation Point". *Commun.Stat.Theor.Meth.* 13 (1984), 843-857.

[2] Bain, L.J. and Weeks, D.L. "A Note on the Truncated Exponential Distribution". *Ann.Math.Stat.* 35 (1964),1366-1367.

[3] Deemer, W.L. and Votaw, D.F. "Estimation of Parameters of Truncated or Censored Exponential Distribution". *Ann.Math.Stat.*, 26 (1955),498-504.

[4] Shalaby,O.A. "Estimation in Truncated Weibull Distribution". *J. King Saud Univ., Admin. Sci.*,2, No. 1 (1990),25-40.

[5] Al Yousef, M.H. "Estimation of Parameters and Truncation Point for the Gompertz Distribution". *J. King Saud Univ., Admin. Sci.*, 5, No. 2 (1993),35-48.

[6] Al Yousef, M.H. "Estimation in the Doubly Truncated Logistic Distribution with Unknown Truncation Points". *J. King Saud Univ. Admin. Sci.*, 11, No. 2, (1999), 51-58.

[7] Burr. I.W. "Cumulative Frequency Functions". *Ann.Math.Stat.*, 13 (1942), 215-232.

[8] Burr, I.W. and Cislak, P.J. "On a General System of Distributions I, Its Curve-shape Characteristics II. The Sample Median". *J.Amer.Statist.Assoc.*, 63 (1968), 627-635.

[9] Burr, I.W." On a General System of Distribution III. The Sample Rang". *J.Amer.Statist.Assoc.*, 63 (1968), 636-643.

[10] Cohen, A.C. "Maximum Likelihood Estimation in the Weibull Distribution Based on Complete and Censored Sample". *Technometrics*, 5 (1965), 579-588.

تقدير معالم توزيع بور المبتور من الطرفين

محمود حسين اليوسف
أستاذ بقسم الأساليب الكمية ، كلية العلوم الإدارية ، جامعة الملك سعود

(قدم للنشـــر في ١٤٢١/٢/٤ هـــ ، و قبل للنشـــر في ١٤٢١/٨/١٩ هـــ)

ملخص البحث . يتناول هذا البحث مشكلة تقدير معالم توزيع بور المبتور من الطرفين عندما تكون نقطتا البتر غير معلومتين ، حيث يمكن اعتبارهما عندئذ معلمتين إضافيتين يجب تقديرهما إلى جانب المعالم الأساسية للتوزيع انطلاقا من بيانات العينة .

و لقد تناولت هذا الدراسة تقدير هذه المعالم باستخدام طريقة الإمكان الأكبر في حالتين : عندما تكون العينة كاملة ، و عندما تكون العينة مراقبة . و لقد تم اختبار النتائج على بيانات مأخوذة من توزيع بور المبتور من الطرفين باستخدام الحاسب الآلي .