Applied Mathematical Sciences, Vol. 5, 2011, no. 58, 2883 - 2890

Estimation of Parameters of the Exponential-Bernoulli Distribution Based on Progressively Censored Data

M. M. Mohie El-Din and M. M. Amein [1]

Department of Mathematics, Faculty of Science
Al-Azhar University, Nasr City (11884), Cairo, Egypt

Abstract. Jones (1953) has derived approximate maximum likelihood method to estimate the location parameter based on order statistics technique and called the method "*approximating the mode from weighted sample values*". In this paper, we developed the Jones's method to estimate the location parameter in the progressive Type-II censoring case. Further, we used the setup proposed by Balakrishnan and Aggarwala (2000) to compute approximate best linear unbiased estimators (ABLUEs) of the location and scale parameters. Finally, we compared between the techniques considered for the estimation.

1. Introduction

Progressive Type-II censoring schemes are the most popular censoring schemes which are used in practice. It can be briefly described as follows: Suppose n independent items are put on life test with continuous, identically distributed failure times $X_1, X_2, \cdots, X_n$. Suppose a censoring scheme $(R_1, R_2, \cdots, R_m)$ is prefixed such that, at the first failure, R_1 surviving items are removed from the experiment at random; at the second observed failure, R_2 surviving items are removed from the experiment at random; this process continues until the $m-th$ observed failure, R_m items are removed from the test at random, $n = m + \sum_{i=1}^{m} R_i$. We will denote the m order observed failure times by $X_{1:m:n}^{(R_1,R_2,\cdots,R_m)}, X_{2:m:n}^{(R_1,R_2,\cdots,R_m)}, \cdots, X_{m:m:n}^{(R_1,R_2,\cdots,R_m)}$ and call them the progressive Type-II right censored order statistics from a sample of size n with progressive censoring scheme $(R_1, R_2, \cdots, R_m)$. Many authors have studied progressive censoring. Among them are Cohen (1963, 1976), Viveros and Balakrishnan (1994), Aggarwala, and Balakrishnan (1996), Balakrishnan and Sandhu (1995), Shuo-Jye (2002), Balakrishnan and et al. (2002) and Balakrishnan (2007). For extensive survey of progressive censored see Balakrishnan and Aggarwala (2000).

The joint probability density function of $X_{1:m:n}^{(R_1,R_2,\cdots,R_m)}, X_{2:m:n}^{(R_1,R_2,\cdots,R_m)}, \cdots, X_{m:m:n}^{(R_1,R_2,\cdots,R_m)}$ is given by (Balakrishnan and Sandhu (1995)):

[1] alamen_2003@yahoo.com

(1.1)

$$f_{1,2,\cdots,m:m:n}(x_1,x_2,\cdots,x_m) = A(n,m-1)\prod_{i=1}^{m} f(x_i)(1-F(x_i))^{R_i}, \quad a < x_1 < x_2 < \cdots < x_m < b,$$

where
$A(n,m-1) = n(n-R_1-1)(n-R_1-R_2-2)\cdots(n-R_1-R_2-\cdots-R_{m-1}-m+1)$.

The probability density function (PDF) for exponential-Bernoulli distribution (EBD) can be written as:

$$f(x) = q\alpha e^{-\alpha x} + p(\alpha+\beta)e^{-(\alpha+\beta)x}, \quad 0 \le p,q \le 1,\ p+q=1, \alpha,\beta \ge 0,\ 0 \le x < \infty \tag{1.2}$$

and the cumulative distribution function (CDF) is:

$$F(x) = 1 - e^{-\alpha x}(q + pe^{-\beta x}) \quad 0 \le p,q \le 1,\ p+q=1, \alpha,\beta \ge 0,\ 0 \le x < \infty \tag{1.3}$$

(for more details about this distribution see Teamah and El-Alosey (2004)).

The probability density function for non-normalized (EBD) can be written as :

$$f(x) = \frac{q\alpha}{\theta_2} e^{-\alpha(\frac{x-\theta_1}{\theta_2})} + \frac{p(\alpha+\beta)}{\theta_2} e^{-(\alpha+\beta)(\frac{x-\theta_1}{\theta_2})}, 0 \le p,q \le 1,\ \alpha,\beta \ge 0,\ \theta_1 \le x < \infty, \tag{1.4}$$

where θ_1 is the location parameter and θ_2 is the scale parameter. The cumulative distribution functions for non-normalized (EBD) can be written as:

$$F(x) = 1 - e^{-\alpha(\frac{x-\theta_1}{\theta_2})}(q + pe^{-\beta(\frac{x-\theta_1}{\theta_2})})\ ,\ 0 \le p,q \le 1,\ \alpha,\beta \ge 0,\ \theta_1 \le x < \infty. \tag{1.5}$$

Recurrence relations and characterizations for doubly truncated (EBD) satisfied by 1.5 are established by Mohie El-Din and Amein.

The aim of this paper is estimating the parameters of (EBD) satisfied by 1.5. For this purpose, we established and applied approximate maximum likelihood estimation (AMLE) method to estimate the location parameter. We used the setup proposed by Balakrishnan and Aggarwala (2000) to compute (ABLUEs) of the location and scale parameters. Also, a relative efficiency which defined by Balakrishnan and Lee (1998) are employed for comparison purpose between the results obtained from the two methods.

2. Approximate Maximum Likelihood Estimation in Progressive Censoring

Jones (1953) has derived approximate maximum likelihood method to estimate the location parameter based on order statistics technique and called the method "*approximating the mode from weighted sample values*". Mohie El-Din (1991) has used that method to estimate the location parameter of the Burr distribution. In a similar way, we relied on the Jones method's to estimate the location parameter in the progressive Typr-II censoring technique. The method can be described as follows :

Assume that the probability density function can be written as

$$y = f(x_r - \theta_1, \theta_2, \theta_3, \cdots \theta_s), \tag{2.1}$$

where θ_1 is the location parameter and $\theta_2, \theta_3, \cdots \theta_s$ are the scale parameters of the distribution. Assume, for every admissible value of x where $y > 0$, that the derivative $y' = \frac{dy}{dx}$ exists, that y has a unique mode at $x = \theta_1$. Let y_i and $y_i^{'}$ respectively, denote the value of y and its first derivative at the point $x = x_i$; and let

$$\phi(x_i) = -\frac{y_i'}{y_i}, \phi^{'}(x_i) = \frac{d\phi(x_i)}{d\ x_i}. \tag{2.2}$$

The maximum likelihood estimate of θ_1 when $\theta_2, \theta_3, \cdots, \theta_s$ are known is, in general, a solution of the equation

$$\frac{d}{d\theta_1}\sum_{r=1}^{m} \log f(x_r - \theta_1, \theta_2, \theta_3, \cdots, \theta_s) = 0, \tag{2.3}$$

where x_r is the r-th order statistic in a random sample of size n, $1 \leq r \leq m$.

We have

$$\frac{d}{d\theta_1}\log f(x_r - \theta_1, \theta_2, \theta_3, \cdots, \theta_s) = -\frac{d}{dx_r}\log f(x_r - \theta_1, \theta_2, \theta_3, \cdots, \theta_s) = \phi(x_r). \tag{2.4}$$

Combining 2.3 and 2.4 , we obtain

$$\sum_{r=1}^{m} \phi(x_r) = 0. \tag{2.5}$$

Using Taylor series to expanding $\phi(\hat{\theta}_1)$ in the neighborhood of x_r yields :

$$\phi(\hat{\theta}_1) \simeq \phi(x_r) + (\hat{\theta}_1 - x_r)\phi^{'}(x_r) + \cdots$$

Neglecting all except the first two terms of the expansion (which procedure may introduce error when $\hat{\theta}_1 - x_r$ is large), and summing over the m observations, we obtain,

$$m\phi(\hat{\theta}_1) \simeq \sum_{r=1}^{m} \phi(x_r) + \sum_{r=1}^{m}(\hat{\theta}_1 - x_r)\phi^{'}(x_r), \tag{2.6}$$

from 2.5, we have

$$m\phi(\hat{\theta}_1) \simeq \sum_{r=1}^{m}(\hat{\theta}_1 - x_r)\phi^{'}(x_r). \tag{2.7}$$

From 2.5, we have $\phi(\hat{\theta}_1) = 0$, then 2.7 will become

$$\sum_{r=1}^{m}(\hat{\theta}_1 - x_r)\phi^{'}(x_r) = 0. \tag{2.8}$$

Then

(2.9) $$\hat{\theta}_1 = \sum_{r=1}^{m} w_r x_{r:m:n} \quad \text{and} \quad w_r = \frac{\phi^{'}(x_{r:m:n})}{\sum_{r=1}^{m} \phi^{'}(x_{r:m:n})},$$

where $w_r(1 = r = m)$ are the coefficients (weights) of the (AMLEs) in the progressive Type-II censoring technique and $x_{r:m:n}$ is the mode of the probability density function $f(x_{r:m:n})$ (density function of progressive Type-II censoring). The explicit representation for the marginal density function of order statistics under progressive Type-II censoring are established by Balakrishnan and et al. (2002) as follows:

(2.10)
$$f_{s:m:n} = c^{'} \sum_{i=0}^{s-1} c_{i,s-1}(R_1+1, R_2+1, \cdots, R_{s-1}+1) f(x_s)\{1-F(x_s)\}^{R_i^{''}-1}, -\infty \leq x_s < \infty,$$

where $R_i^{''} = (R_s^* + 1) + \sum_{j=s-i}^{s-1}(R_j + 1), R_s^* = n - s - R_1 - R_2 - \cdots - R_{s-1}$,

$a_r = (a_1, a_2, \cdots, a_r)$, $c_{i,r}(a_r) = \frac{(-1)^i}{[\prod_{j=1}^{i} \sum_{k=r-i+1}^{r-i+j} a_k\}\{\prod_{j=1}^{r-i} \sum_{k=i}^{r-j} a_k]}$ and
$c^{'} = n(n - R_1 - 1)(n - R_1 - R_2 - 2) \cdots (n - R_1 - \cdots - R_{s-1} - s + 1)$.

To find x_r the mode of this distribution, we set the derivative of $f_{s:m:n}(x_{s:m:n})$ equal to zero, we get

(2.11)
$$f_{s:m:n}^{'} = c^{'} \sum_{i=0}^{s-1} \{D_{i,s} f^{'}(x_s)\{1-F(x_s)\}^{R_i^{''}-1} + (R_i^{''}-1) f^2(x_s)\{1-F(x_s)\}^{R_i^{''}-2}\} = 0,$$
$$D_{i,s} = c_{i,s-1}(R_1+1, R_2+1, \cdots, R_{s-1}+1).$$

This equation can generally be solved for $x_r, r = 1, 2, \cdots, m$ by an numerical procedure.
The likelihood estimation for θ_1 is obtained in the form ,

(2.12) $$\hat{\theta}_1 = \sum_{r=1}^{m} w_r x_{r:m:n},$$

One can using the next algorithm to estimate the location parameter $\hat{\theta}_1$using (AMLEs) as follows :
Step 1. Finding $x_{r:m:n}$ the mode of (2. 10), $r = 1, 2, \cdots, m$
Step 2. Finding $\phi^{'}(x_{r:m:n})$.
Step 3. Calculate $w_r = \frac{\phi^{'}(x_{r:m:n})}{\sum_{r=1}^{m} \phi^{'}(x_{r:m:n})}$.

3. Approximate Best Linear Unbiased Estimators

The goal of this section is estimating the location and scale parameters for (EBD) with (CDF) given in 1.3. Let $X_{1:m:n}^{(R_1,R_2,\cdots,R_m)}, X_{2:m:n}^{(R_1,R_2,\cdots,R_m)}, \cdots, X_{m:m:n}^{(R_1,R_2,\cdots,R_m)}$ be a progressively Type-II right censored sample from a location-scale family of distributions with (PDF) $g(x;\theta_1,\theta_2) = \frac{1}{\theta_2} f(\frac{x-\theta_1}{\theta_2})$, where $f(.)$ is the standard density. Let μ and $\sum$ denote the mean vector and the variance-covariance matrix of the corresponding standardized progressively Type-II right censored sample $Y_{i:m:n}^{(R_1,R_2,\cdots,R_m)} = (X_{i:m:n}^{(R_1,R_2,\cdots,R_m)} - \theta_1)/\theta_2, i = 1,\cdots,m$ from $f(.)$. Then, the (ABLUEs) of θ_1 and θ_2 and their variances and covariance are (see Balakrishnan and Cohen (1991), Balakrishnan and Aggarwala (2000) and David and Nagaraja (2003)) :-

$\theta_1^* = -\mu^T \Gamma Y$ and $\theta_2^* = 1^T \Gamma Y$

and

$Var(\theta_1^*) = \frac{\sigma^2}{\Delta}(\mu^T \sum^{-1} \mu), Var(\theta_2^*) = \frac{\sigma^2}{\Delta}(1^T \sum^{-1} 1)$ and $Cov(\theta_1^*, \theta_2^*) = -\frac{\sigma^2}{\Delta}(\mu^T \sum^{-1} 1)$,

where $A = (1,\mu)$ is a matrix of m rows and 2 columns, $Y = (y_1 \quad y_2 \cdots y_m\ )^T$, $\hat{\theta} = (\theta_1,\theta_2)^T$ and $\mu = (\mu_{i:m:n} \quad \mu_{2:m:n} \cdots \mu_{m:m:n}\ )^T, 1 = (1,\cdots,1), \Gamma = \sum^{-1}(1\mu^T - \mu 1^T)\sum^{-1}/\Delta$ and $\Delta = (1^T \sum^{-1} 1)(\mu^T \sum^{-1} \mu) - ((\mu^T \sum^{-1} 1))^2$.

Let the linear combination of these ordered sample values be

$$\theta_i^* = \sum_{j=1}^{m} \alpha_{ij} y_j, i = 1, 2, \tag{3.1}$$

where θ_i^* is an estimate of the parameter θ_i .

4. Numerical Results

This section deals with obtaining some numerical results. We calculated the coefficients, variances, bias and (MSE) of the (AMLEs) and Coefficients, variances and covariances of the (ABLUEs) for *(EBD)* in Tables (2-3), respectively. All results are given in Tables are obtained at $\alpha = 0.2, \beta = 0.5, p = 0.7, q = 0.3$ and different censoring schemes (see Table 1).

Table 1

n	m	*Censoringscheme* (CS)
10	2	$R_1 = \{5, 3\}$
15	3	$R_2 = \{4, 5, 3\}$
20	4	$R_3 = \{3, 4, 4, 5\}$
30	5	$R_4 = \{6, 5, 3, 7, 4\}$

Table 2. *Coefficients, variances, bias and MSE of the (AMLEs)*

n	m	CS	$AMLEs$			
			w_i	$Var\theta_1$	$Bias$	$MSE = Bias^2 + Var\theta_1$
10	2	R_1	3.81407×10^{-8}	$0.320638\theta_1^2$	$0.676453\theta_1$	$0.778227\theta_1^2$
			1			
15	3	R_2	0.0000716306	$0.0888621\theta_1^2$	$0.482879\theta_1$	$0.322035\theta_1^2$
			0.516372			
			0.483556			
20	4	R_3	0.00282237	$0.0831127\theta_1^2$	$0.52848\theta_1$	$0.362409\theta_1^2$
			0.343177			
			0.327			
			0.327			
30	5	R_4	0.0988325	$0.0252363\theta_1^2$	$0.340883\theta_1$	$0.141438\theta_1^2$
			0.23007			
			0.223699			
			0.223699			
			0.223699			

Table 3. *Coefficients, variances and covariances of the (ABLUEs)*

n	m	CS	$ABLUEs$				
			α_{1i}	α_{2i}	$Var\theta_1$	$Var\theta_2$	$Covar(\theta_1, \theta_2)$
10	2	R_1	1.57875	−3.12361	$0.0715187\theta_2^2$	$1.08905\theta_2^2$	$-0.197352\theta_2^2$
			−0.578754	3.12361			
15	3	R_2	1.37313	−3.04744	$0.0232901\theta_2^2$	$0.550987\theta_2^2$	$-0.06574492\theta_2^2$
			−0.222489	1.82256			
			−0.150643	1.22488			
20	4	R_3	1.24361	−2.665	$0.01154491\theta_2^2$	$0.375957\theta_2^2$	$-0.033453\theta_2^2$
			−0.0953289	1.04667			
			−0.0930941	1.01995			
			−0.0551894	0.598377			
30	5	R_4	1.21296	−3.50083	$0.004708898\theta_2^2$	$0.268398\theta_2^2$	$-0.0159423\ \theta_2^2$
			−0.072096	1.18762			
			−0.0362764	0.598278			
			−0.0523124	0.860473			
			−0.052272	0.854461			

Remark 4.1. $\sum_{j=1}^{m} \alpha_{1j} \simeq 1, \sum_{j=1}^{m} \alpha_{2j} \simeq 0$ and $\sum_{j=1}^{m} w_j \simeq 1$.

In order to compare the performance of the (ABLUEs) and (AMLEs), Balakrishnan and Lee (1998) have defined the relative efficiency between the two methods of estimation as follows :-

$EFF(\theta_1) = \frac{MSE(\theta_1)}{Var(\theta_1)} \times 100$ and $EFF(\theta_2) = \frac{MSE(\theta_2)}{Var(\theta_2)} \times 100$

The values of the relative efficiency can be interpreted as follows:

1- If $Eff(\theta_1) > 100$, we conclude that the estimation of θ_1 based on the (ABLUEs) is more efficient than that based on the (AMLEs).

2- If $Eff(\theta_1) < 100$, we conclude that the estimation of θ_1 based on the (AMLEs) is more efficient than that based on the (ABLUEs).

Table 4. Relative efficiency

n, CS	$n = 10, CS = R_1$	$n = 15, CS = R_2$	$n = 20, CS = R_3$	$n = 30, CS = R_4$
$EFF(\theta_1)$	1088.145	1382.712	3137.985	3003.58

Conclusion : Its clear form Table 4 that the (ABLUEs) is more efficient than the (AMLEs) for estimating the location parameter.

References

[1] **Aggarwala, R. and Balakrishnan, N. (1996).** Recurrence relations for single and single and product moments of progressive Type-II right censored order statistics from exponential and truncated exponential distributions, *Ann. Inst. Statist. Math.*, Vol. 48, No. 4, 757-771.

[2] **Balakrishnan, N., (2007)**. Progressive censoring methodology : an appraisal. Test 16, 211–296 .

[3] **Balakrishnan, N. and Cohen, A. C. (1991).** *Order statistics and inference* : Estimation methods. San Diego : Academic Press.

[4] **Balakrishnan, N. and Sandhu, R. A. (1995)**. A simple simulational algorithm for generating progressive Type-II censored samples, *Amer. Statist.*, 49, 229-230.

[5] **Balakrishnan, N. and Lee, S. K. (1998).** Order statistics from the Type III generalized logistic distribution and applications, (Eds. N. Balakrishnan and C. R. Rao) *Handbook of Statistics*, Vol.17, 127-155, North-Holland, Amsterdam.

[6] **Balakrishnan, N. and Aggarwala, R., (2000)**. *Progressive Censoring: Theory, Methods and Applications*. Birkhkauser, Boston.

[7] **Balakrishnan N, Childs A. and Chandrasekar B (2002).** An efficient computational method for moments of order statistics under progressive censoring. *Stat Probab Lett* , 60:359–365.

[8] **Cohen, A. C. (1963).** Progressively censored samples in life testing, *Technometrics*, 5, 327-329.

[9] **Cohen, A. C. (1976).** Progressively censored sampling in the three parameter log-normal distribution, *Technometrics*, 18, 99-103.

[10] **Jones, H. L. (1953).** Approximating the mode from weighted sample values, *Journal of The American Statistical Association*, 48, 113-127.

[11] **Mohie El-Din, M. M. (1991).** Maximum likelihood weights of the location of the Burr distribution, *Computers & Mathematics with Applications*, Vol. 21, Issue 4, 33-38.

[12] **Mohie El-Din, M. M. and Amein, M. M. (2010).** Order Statistics and Characterizations from Independent and Identical Doubly-Truncated Exponential-Bernoulli Random Variables, *Journal of Advanced Research in Statistics and Probability*, Vol. 2, Issue. 1, pp. 14-26.

[13] **Shuo-Jye, Wu. (2002).** Estimation of the parameters of the Weibull distribution with progressively censored data, *Journal of Japan Statistical Society*, 32, No.2, 155 - 163.

[14] **Teamah, A. A. and El-Alosey, A. R. (2004).** Random sum of mixtures of exponential distribution, *Internat. J. of Appl. Math.* , Vol. 16, No. 2 ,237-247.

[15] **Viveros, R. and Balakrishnan, N. (1994).** Interval estimation of parameters of life from progressively censored data, *Technometrics*, 36, 84 - 91.

Received: January, 2011