THE EGYPTINA STATISTICA JOURNAL
ISSR, CAIRO UNIV., VOL., 45, NO.2, 2001

ESTIMATION OF PARAMETERS OF LOMAX DISTRIBUTION BASED ON RECORD VALUES

Khalaf S. Sultan[1], Mohamed E. Moshref[1] and Nagi S. Abd-El-Hakim[2]

[1]*Department of Mathematics, Al-Azhar University, Nasr City, Cairo 11884, Egypt*
[2]*Department of Mathematics, El-Minia University, El-Minia, Egypt*

Key Words and Phrases: *Upper record values; Lomax distribution; BLUE's and MLE's.*

Abstract:
In this paper, we derive exact explicit expressions for the single and double moments of the upper record values from Lomax distribution. We then use these expressions to compute t‥ mean, variance and the best linear unbiased estimates (BLUE's) of the location and scale parameters of lomax distribution. Finally, we obtain the maximum likelihood estimates (MLE's) and compare them with the BLUE's.

1 Introduction

Record values arise naturally in many real life applications involving data relating to weather, sports, economics and life testing studies. Many authors have studied record values and associated statistics. Among others, are Chandler (1952), Ahsanullah (1980, 1988, 1990, 1993, 1995), and Arnold, Balakrishnan and Nagaraja (1992, 1998). Ahsanullah (1980, 1990), Balakrishnan and Chan (1993), and Balakrishnan, Ahsanullah and Chan (1995) have discussed some inferential methods for exponential, Gumbel, Weibull and logistic distributions, respectively. Abd-El-Hakim and Sultan (2001) have obtained the maximum likelihood estimates of the location and scale parameters of Weibull record values and have compared them with the BLUE's given by Balakrishnan and Chan (1993).

Lomax distribution has been used in connection with studies of income, size of cities and r‥iability modeling. The Lomax distribution is also known as the Pareto II distribution (see Arnold (1983)). Lomax (1954) used this distribution in the analysis of business failure data.

The two-parameter Lomax distribution has its density as

$$f(y) = \begin{cases} \frac{\lambda}{\sigma}[1 + (y-\theta)/\sigma]^{-(\lambda+1)}, & y \geq \theta, \lambda > 0, \\ 0, & \text{otherwise}, \end{cases} \tag{1.1}$$

while the standard form of Lomax distribution is given by

$$f(x) = \begin{cases} \lambda(1+x)^{-(\lambda+1)}, & x \geq 0, \lambda > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{1.2}$$

and

$$F(x) = \begin{cases} 1-(1+x)^{-\lambda}, & x \geq 0, \lambda > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{1.3}$$

Balakrishnan and Ahsanullah (1994) have established some recurrence relations satisfied by the single and double moments of upper record values from the Lomax distribution in (1.2).

In this paper, we derive exact explicit forms for the single and double moments of record values from Lomax distribution in Section 2. Then, we use these moments to obtain the BLUE's. In section 3, we obtain the MLE's of Lomax parameters. In Section 4, we discuss an application. Finally, conclusion and comparisons are made in Section 5.

2 Moments and Best Linear Unbiased Estimates (BLUE's)

Let $X_{U(1)}, X_{U(2)}, \ldots X_{U(n)}$ be the first n upper record values from the Lomax distribution (. 2), then the single and double moments are derived as follows:

2.1 Single moments

The single moments of the upper record value $X_{U(m)}$ is obtained to be

$$\begin{aligned} \mu_m^{(r)} &= \frac{1}{\Gamma(m)} \int_0^\infty x^r[-\log\{1-F(x)\}]^{m-1} f(x)dx, \\ &= \sum_{i=0}^{r} \binom{r}{i} (-1)^{r-i} (1-i/\lambda)^{-m}, \quad \lambda \geq r \text{ and } r = 1, 2, \ldots. \end{aligned} \tag{2.1}$$

As a check put $r = 1$ and $m = 1$ in the above expression of $\mu_m^{(r)}$, we get

$$\mu_1^{(1)} = \frac{1}{\lambda - 1},$$

which gives the mean of Lomax distribution given in (1.2).

2.2 Double moments

The double moments of the upper record values $X_{U(m)}$ and $X_{U(n)}$, $m < n$ is given by

$$\begin{aligned}\mu_{m,n}^{(r,s)} &= \frac{1}{\Gamma(m)\Gamma(n-m)}\int_0^{\infty}\int_x^{\infty} x^r y^s[-\log\{1-F(x)\}]^{m-1} \\ &\times [-\log\{1-F(y)\}+\log\{1-F(x)\}]^{n-m-1}\frac{f(x)}{1-F(x)}f(y)dydx, \\ &= \sum_{j=0}^{s}\sum_{i=0}^{r}\binom{s}{j}\binom{r}{i}(-1)^{r+s-i-j}(1-j/\lambda)^{-(n-m)}\left(1-\frac{i+j}{\lambda}\right)^{-m}, \lambda \geq r+s.\end{aligned} \tag{2.2}$$

As a check put $s = 0$ in (2.2) and use (2.1), we have

$$\mu_{m,n}^{(r,0)} = \mu_m^{(r)}.$$

2.3 Best linear unbiased estimates (BLUE's)

Let $Y_{U(1)} \leq Y_{U(2)} \leq \cdots \leq Y_{U(n)}$ be the upper record values from Lomax distribution in (1.1), and let $X_{U(i)} = \left(Y_{U(i)} - \theta\right)/\sigma$, $i = 1, \ldots, n$, be the corresponding record values from the standard Lomax distribution in (1.2). Let us denote $E(X_{U(i)})$ by μ_i, $Var(X_{U(i)})$ by $\sigma_{i,i}$, and $Cov(X_{U(i)}, X_{U(j)})$ by $\sigma_{i,j}$; further, let

$$\begin{aligned} Y &= \left(Y_{U(1)}, Y_{U(2)}, \ldots, Y_{U(n)}\right)^T, \\ \mu &= (\mu_1, \mu_2, \ldots, \mu_n)^T, \\ 1 &= \underbrace{(1, 1, \ldots, 1)^T}_{n}, \\ \text{and } \Sigma &= (\sigma_{ij}),\ 1 \leq i,\ j \leq n. \end{aligned}$$

Then, the BLUE's of θ and σ are given by [see Balakrishnan and Cohen (1991)]

$$\theta^* = \left\{\frac{\mu^T\Sigma^{-1}\mu 1^T\Sigma^{-1} - \mu^T\Sigma^{-1}1\mu^T\Sigma^{-1}}{(\mu^T\Sigma^{-1}\mu)(1^T\Sigma^{-1}1) - (\mu^T\Sigma^{-1}1)^2}\right\}Y = \sum_{i=1}^{n} a_i Y_{U(i)}, \tag{2.3}$$

and

$$\sigma^* = \left\{\frac{1^T\Sigma^{-1}1\mu^T\Sigma^{-1} - 1^T\Sigma^{-1}\mu 1^T\Sigma^{-1}}{(\mu^T\Sigma^{-1}\mu)(1^T\Sigma^{-1}1) - (\mu^T\Sigma^{-1}1)^2}\right\}Y = \sum_{i=1}^{n} b_i Y_{U(i)}. \tag{2.4}$$

Furthermore, the variances and covariance of these BLUE's are given by [see Balakrishnan and Cohen (1991)]

$$Var(\theta^*) = \sigma^2\left\{\frac{\mu^T\Sigma^{-1}\mu}{(\mu^T\Sigma^{-1}\mu)(1^T\Sigma^{-1}1) - (\mu^T\Sigma^{-1}1)^2}\right\} = \sigma^2 V_1, \tag{2.5}$$

$$Var(\sigma^*) = \sigma^2 \left\{ \frac{\mathbf{1}^T\Sigma^{-1}\mathbf{1}}{(\mu^T\Sigma^{-1}\mu)(\mathbf{1}^T\Sigma^{-1}\mathbf{1}) - (\mu^T\Sigma^{-1}\mathbf{1})^2} \right\} = \sigma^2 V_2, \tag{2.6}$$

and

$$Cov(\theta^*, \sigma^*) = \sigma^2 \left\{ \frac{-\mu^T\Sigma^{-1}\mathbf{1}}{(\mu^T\Sigma^{-1}\mu)(\mathbf{1}^T\Sigma^{-1}\mathbf{1}) - (\mu^T\Sigma^{-1}\mathbf{1})^2} \right\} = \sigma^2 V_3, \tag{2.7}$$

where

$$\mu_m = (1 - 1/\lambda)^{-m} - 1, \tag{2.8}$$

$\Sigma^{-1} = (\sigma^{ij})$ is a symmetric tridiagonal matrix, and for $i \le j$, σ^{ij} is obtained to be

$$\sigma^{ij} = \begin{cases} -\left(\frac{\lambda-2}{\lambda}\right)^i (\lambda-2)(\lambda-1), & j = i+1, i = 1, ..., k-1, \\ \left(\frac{\lambda-2}{\lambda}\right)^i (2\lambda^2 - 4\lambda + 1), & i = j = 1 \text{ to } k-1, \\ (\lambda-1)^2 (\frac{\lambda-2}{\lambda})^k, & i = j = k, \\ 0, & j > i+1, \end{cases} \tag{2.9}$$

for details, we refer for example to Balakrishnan and Cohen (1991). By using the above expressions, the BLUEs of θ and σ can be calculated for different sample sizes and different values of λ. The coefficients of the BLUE's a_i and b_i in (2.3) and (2.4) were calculated in Table 1 for $n = 3(1)7$ and $\lambda = 5$.

Table 1: Coefficients of the BLUE's when $\lambda = 5.0$, $\theta = 0.0$ and $\sigma = 1.0$.

n	a_i						
3	1.425	-.125	-.300				
4	1.310	-.102	-.061	-.147			
5	1.260	-.092	-.055	-.033	-.079		
6	1.234	-.087	-.052	-.031	-.019	-.045	
7	1.220	-.084	-.050	-.030	-.018	-.011	-.026
n	b_i						
3	-1.700	.500	1.200				
4	-1.241	.408	.245	.588			
5	-1.038	.368	.221	.132	.318		
6	-.935	.347	.208	.125	.075	.180	
7	-.878	.336	.201	.121	.073	.044	.104

y making use of the entries in Table 1, the BLUE's θ^* and σ^* given in (2.3) and (2.4), were ated based on 10000 simulated upper record sets of sizes $n = 3(1)7$ (with $\theta = 0.0$, 1.0 and $\lambda = 5.0$). Table 2 represents the bias and MSE of these BLUE's.

Table 2: The Bias and MSE of the BLUE's when $\lambda = 5.0$, $\theta = 0.0$ and $\sigma = 1.0$

	θ^*		σ^*	
n	Bias	MSE	Bias	MSE
3	-.0036	.1361	.0114	1.1779
4	.0010	.1234	-.0066	.9737
5	-.0101	.1178	-.0002	.8837
6	-.0124	.1150	-.0330	.8388
7	-.0336	.1144	-.0665	.8170

3 Maximum Likelihood Estimates (MLE's)

The joint density function of the first n upper record values $X_{U(1)}, X_{U(2)}, \dots, X_{U(n)}$ is given by

$$f_{1,2,\dots,n}(x_{U(1)},\dots,x_{U(n)}) = \prod_{i=1}^{n} f(x_{U(i)}) \Big/ \prod_{i=1}^{n-1}[1 - F(x_{U(i)})], \tag{3.1}$$

where $f(.)$ is given by (1.1) and $F(.)$ is the corresponding cdf.

The the log-likelihood function of (3.1), is given by

$$L^*(\theta,\alpha,\lambda) = n\log\lambda - n\log\sigma - \lambda\log\left(1 + \frac{x_{U(n)} - \theta}{\sigma}\right) - \sum_{i=1}^{n}\log\left(1 + \frac{x_{U(i)} - \theta}{\sigma}\right),$$

and hence the MLE's of θ, λ and σ are obtained, respectively, by

$$\tilde{\theta} = x_{U(1)}, \tag{3.2}$$

$$\tilde{\lambda} = n\Big/\log\left(1 + \frac{x_{U(n)} - \tilde{\theta}}{\tilde{\sigma}}\right), \tag{3.3}$$

and

$$\psi(\tilde{\sigma}) = \tilde{\lambda}\left(\frac{x_{U(n)} - \tilde{\theta}}{\tilde{\sigma} + x_{U(n)} - \tilde{\theta}}\right) + \sum_{i=1}^{n}\frac{x_{U(i)} - \tilde{\theta} - n}{\tilde{\sigma} + x_{U(i)} - \tilde{\theta}} = 0. \tag{3.4}$$

From (3.2), (3.3) and (3.4), we observe that:

1. If θ and σ are known, then $E(\tilde{\theta}) = \theta + \sigma/(\lambda - 1), \lambda > 1$, that is $\tilde{\theta}$ given in (6.2) is a biased estimate of θ while

$$\hat{\theta} = x_{U(1)} - \frac{\sigma}{\lambda - 1}, \lambda > 1, \tag{3.5}$$

is an unbiased estimate of θ with variance is given by

$$Var(\hat{\theta}) = \frac{\lambda\sigma^2}{(\lambda-2)(\lambda-1)^2}, \lambda > 2. \tag{3.6}$$

So, we can replace (3.2) by

$$\hat{\theta} = x_{U(1)} - \frac{\tilde{\sigma}}{\tilde{\lambda}-1}, \tilde{\lambda} > 1, \tag{3.7}$$

2. If θ and σ are known, then $n/\tilde{\lambda} \sim \Gamma(n, 1/\lambda)$, and hence $1/\tilde{\lambda}$ is an unbiased estimate of $1/\lambda$, but $E(\tilde{\lambda}) = \frac{n}{n-1}\lambda$, that is $\tilde{\lambda}$ given in (3.3) is a biased estimate of λ, while

$$\hat{\lambda} = (n-1)/\log\left(1 + \frac{x_{U(n)} - \theta}{\sigma}\right) \tag{3.8}$$

is an unbiased estimate of λ with variance is given by

$$Var(\hat{\lambda}) = \frac{\lambda^2}{n-2}, n > 2, \tag{3.9}$$

and the efficiency of $\hat{\lambda}$ based on Cramér-Rao inequality is obtained to be $(n-2)/n$.

3. If θ and σ are known, then both of $\tilde{\lambda}$ and $\hat{\lambda}$ given in (3.3) and (3.8), respectively, are sufficient estimates of λ.

4. The nonlinear equation (3.4) has a unique solution, that is because $\psi(0) > 0$, $\psi(\infty) < 0$ and $\psi'(\tilde{\sigma}) < 0$.

When λ is known, the MLE's of θ and σ were calculated based on 10000 simulated records as presented below:

Table 3: The Bias and MSE of the MLE's when $\lambda = 5.0$, $\theta = 0.0$ and $\sigma = 1.0$.

	$\tilde{\theta}$		$\tilde{\sigma}$	
n	Bias	MSE	Bias	MSE
3	.0380	.1267	-.1625	.9813
4	.0137	.1251	-.0653	.9592
5	-.0001	.1240	-.0101	.9097
6	-.0077	.1224	.0203	.9034
7	-.0129	.1210	.0413	.9010

Application

An interesting application of the Lomax distribution stem from the fact that it is a mixture (or compound) of exponential distribution $f_{X|\Theta}(x \mid \theta) = \theta e^{-\theta x}$, where Θ has a gamma istribution, $f_{\Theta}(\theta) = [\sigma/\Gamma(\beta)](\sigma\theta)^{\beta-1}e^{-\sigma\theta}$. This is used in the analysis of heart transplant ata in Turnbull et al. (1974). The data represents the survival times and times to tansplant (in daya) for 82 patients from the Stanford heart transplantion program. From the data, ne gets the upper record values, then by using Table 1, the BLUE's for the location and cale parameters are obtained as given below:

Survival times	Upper record values	BLUE's
Nontransplant patients	49, 84, 101, 148	$\theta^* = 28$, $\sigma^* = 85$
Transplant patients	35, 50, 82	$\theta^* = 19$, $\sigma^* = 64$

5 Conclusion and Comparisons

In this work, the moments of the upper record values from Lomax distribution are derived. These moment are then used to obtain the BLUE's. In addtion, the MLE's of Lomax parameters are obtained. From Tables 1, 2 and 3, we conclude the following points:

1. The coefficients of the BLUE's, presented in Table 1 are cheked by the forms $\sum_{i=1}^{n} a_i = 1$ and $\sum_{i=1}^{n} b_i = 0$.

2. As we can see from Table 2, the MSE's of the BLUE's of θ and σ decrease as n increases.

3. From Table 3, we can see that the MSE's of the MLE's of θ and σ decrease as n increases.

   In conclusion, we can say that the BLUE of θ is better than its MLE except for $n = 3$, while the BLUE of σ is better than the MLE except for n=3 and 4.

Acknowledgments The authors would like to thank would like to thank the refrees for their useful comments that improved the presentation of the paper.

REFERENCES

Abd-El-Hakim, N. S. and Sultan, K. S. (2001). Maximum likelihood estimates of Weibull parameters based on record values, *Journal of the Egyptian Mathematical Society*, 9, 79 - 89.

Ahsanullah, M. (1980). Linear prediction of record values for the two parameter exponential distribution, *Ann. Inst. Statist. Math.* 32, 363 - 368.

Ahsanullah, M. (1988). *Intoduction to Record Values*, Ginn Press, Needham Heights, Massachusetts.

Ahsanullah, M. (1990). Estimation of the parameters of the Gumbel distribution based on the m record values, *Comput. Statist. Quart.* **3**, 231 - 239.

Ahsanullah, M. (1993). On the record values from univariate distributions, *National Institute of Standards and Technology Journal of Research, Special Publications* **866**, 1-6.

Ahsanullah, M. (1995). Record values, In *The Exponential Distribution: Theory, Methods and Applications* (Eds., N. Balakrishnan and A.P. Basu), Gordon and Breach Publishers, Newark, New Jersey.

Arnold, B. C. (1983). *The Pareto Distributions*, International Co-operative Publishing House, Fairland, MD.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). *A First Course in Order Statistics*, John Wiley & Sons, New York.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998). *Records*, John Wiley & Sons, New York (to appear).

Balakrishnan, N. and Ahsanullah, M. (1994). Relations for single and product moments of record values from Lomax distribution, *Sankhyā, Series B* **56**, 140 - 146.

Balakrishnan, N., Ahsanullah, M. and Chan, P. S. (1995). On the logistic record values and associated inference, *J. Appl. Statist. Sci.* **2**, 233-248.

Balakrishnan, N. and Chan, P. S. (1993). Record values from Rayleigh and Weibull distributions and associated inference, *National Institute of Standards and Technology Journal of Research, Special Publications* **866**, 41-51.

Balakrishnan, N. and Cohen, A. C. (1991). *Order Statistics and Inference: Estimation Methods*, Academic Press, San Diego.

Chandler, K. N. (1952). The distribution and frequency of record values, *J. Roy. Statist. Soc., Ser. B* **14**, 220 - 228.

Lomax, K. S. (1954). Business failures. Another example of the analysis of failure data. *J. Amer. Statist. Assoc.*, **49**, 847 - 852.

Turnbull, B.W., Brown, B.W., Jr. and Hu, M. (1974). Survivorship analysis of heart transplant data, *J. of Amr. Statist. Assoc.*, **69**, 74-80.