DIFFERENT METHOD OF ESTIMATIONS FOR A THREE-PARAMETER LINDLEY DISTRIBUTION

Lazim Kamberi, Rushadije Ramani, Dashmir Ibishi, Albana Gashi-Kadolli
State University of Tetovo, 1200,Tetovo Macedonia
rushadije.ramani@unite.edu.mk, lazim.kamberi@unite.edu.mk, dashmir.ibishi@unite.edu.mk, gashi_460@hotmail.com

ABSTRACT

The aim of this paper is to compare through Monte Carlo simulations the finite sample properties of the estimates of the parameters of a three parameter Lindley distribution obtained by four estimation methods: maximum likelihood, ordinary least squares, product estimating and Bayes estimation. The bias and mean-squared error are used as the criterion for comparison. The study reveals that the Bayes estimation and least squares methods are highly competitive with the maximum likelihood method and product estimating methods in small and large samples.

***Key Words**: Lindley distribution, Maximum likelihood estimation, Least squares estimation, Bayes estimation, Monte Carlo simulations.*

1 INTRODUCTION

The Lindley distribution was introducend by Lindley [14] in thecontext of Bayesian statistics, as a counter example of fudicial statistics.Its probability density function is given by

$$f(x,\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}, \quad x>0,\theta>0. \tag{1.1}$$

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

$$F(x) = 1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}, \quad x>0,\theta>0 \tag{1.2}$$

A three parameter Lindley distributions has pdf (see [12])

$$g(x) = \frac{\theta(\theta+1+\theta x)^{\beta-1}(1+x)e^{-\theta\beta x}}{B(\alpha,\beta)(\theta+1)^{\beta}}\left[1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}\right]^{\alpha-1} \tag{1.3}$$

The corresponding cumulative distribution function (c.d.f.) and hazart rate function , are given by:

$$G(x) = \frac{\left(1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}\right)^{\alpha}}{\alpha B(\alpha,\beta)}\,{}_2F_1\left(\alpha,1-\beta;\alpha+1;1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}\right) \tag{1.4}$$

And

$$h(t,\theta,\alpha,\beta) = \frac{\frac{\theta(\theta+1+\theta x)^{\beta-1}(1+x)e^{-\theta\beta x}}{B(\alpha,\beta)(\theta+1)^{\beta}}\left[1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}\right]^{\alpha-1}}{1-\frac{\left(1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}\right)^{\alpha}}{\alpha B(\alpha,\beta)}\,{}_2F_1\left(\alpha,1-\beta;\alpha+1;1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x}\right)}$$

2 MAXIMUM LIKELIHOOD ESTIMATES

In this section we consider the maximum likelihood estimation of the unknown parameter based on a complete sample. Let us assume that we have a sample of size *n,* namely $\{x_1, x_2, \ldots, x_n\}$ drawn from the density (1.3). The log-likelihood function is given by The corresponding Log-likelihood function is given by:

$$\ell = \ln L = n\{ \log(\theta) - \log\Gamma(\alpha) - \log\Gamma(\beta) + \log\Gamma(\alpha+\beta) - \beta\log(\theta+1) \}$$

$$+ \sum_{i=1} (1+x_i) + (\beta-1)\log(\theta+1+\theta x_i) - \theta\beta\sum_{i=1}$$

$$+ (\alpha-1)\sum_{i=1}^{n} log\left(1-\frac{\theta+1+\theta x_i}{\theta+1}e^{-\theta x_i}\right)$$

Normal equations can be obtained by taking the first derivates of the log-likelihood function with respect to α, β and θ are equate them to zeros as follows:

$$n\psi(\alpha+\beta) - n\psi(\alpha)$$

$$+\sum_{i=1}^{m}\left(-\frac{(\theta+1+\theta x_i)e^{-\theta x_i}}{\theta+1}\right) = 0,$$

$$n\psi(\alpha+\beta) - n\psi(\beta) - n\log(\theta+1)$$

$$+\sum_{i=1} \log(\theta+1+\theta x_i) - \theta\sum_{i=1} = 0,$$

$$\frac{2m}{\theta} - \frac{m\beta}{\theta+1} + (\beta-1)\sum_{i=1}\frac{1+x_i}{\theta+1+\theta x_i}$$

$$-\beta\sum_{i=1}^{m} x_i + (\alpha-1)\sum_{i=1}^{n}\frac{e^{-\theta x_i} x_i\{(\theta+1)(\theta+1+\theta x_i) - 1\}}{(\theta+1)^2\left(1-\frac{\theta+1+\theta x_i}{\theta+1}e^{-\theta x_i}\right)} = 0,$$

where $\psi(\cdot)$ is digamma function. The MLEs $(\hat{\alpha}, \hat{\beta}, \hat{\theta})$ of (α, β, θ), respectively is obtained by solving this non-linear system of equations.

3 MAXIMUM PRODUCT SPACING ESTIMATES

The maximum product spacing (MPS) method has been proposed by [2]. This method is based on an idea that the differences (spacings) of the consecutive points should be identically distributed. The geometric mean of the differences is given as

$$GM = \sqrt[n+1]{\prod_{i=1}^{n+1} D_i}$$

where, the difference D_i is defined as

$$D_i = \int_{x_{(i-1)}}^{x_{(i)}} f(x, \alpha, \beta, a, b, c)\, dx;\ i = 1, 2, \ldots, n+1. \tag{3.1}$$

where $F(x_{(0)},\alpha,\beta,\theta)=0$ and $F(x_{(n+1)},\alpha,\beta,\theta)=1$. The MPS estimators $\hat{\alpha}_{MPS},\hat{\beta}_{MPS}$ and $\hat{\theta}_{MPS}$ of α, β and θ are obtained by maximizing the geometric mean (GM) of the differences. Substituting pdf of a three parameter Lindley distribution in (3.1) and taking logarithm of the above expression, we will have

$$LogGM=\frac{1}{n+1}\sum_{i=1}^{n+1}\log\left[F(x_{(i)},\alpha,\beta,\theta)-F(x_{(i-1)},\alpha,\beta,\theta)\right]$$

The MPS estimators $\hat{\alpha}_{MPS},\hat{\beta}_{MPS}$ and $\hat{\theta}_{MPS}$ of α, β and θ can be obtained as the simultaneous solution of the following non-linear equations:

$$\frac{\partial LogGM}{\partial\alpha}=\frac{1}{n+1}\sum_{i=1}^{n+1}\left[\frac{F'_{\alpha}(x_{(i)},\alpha,\beta,\theta)-F'_{\alpha}(x_{(i-1)},\alpha,\beta,\theta)}{F(x_{(i)},\alpha,\beta,\theta)-F(x_{(i-1)},\alpha,\beta,\theta)}\right]=0$$

$$\frac{\partial LogGM}{\partial\beta}=\frac{1}{n+1}\sum_{i=1}^{n+1}\left[\frac{F'_{\beta}(x_{(i)},\alpha,\beta,\theta)-F'_{\beta}(x_{(i-1)},\alpha,\beta,\theta)}{F(x_{(i)},\alpha,\beta,\theta)-F(x_{(i-1)},\alpha,\beta,\theta)}\right]=0$$

$$\frac{\partial LogGM}{\partial\theta}=\frac{1}{n+1}\sum_{i=1}^{n+1}\left[\frac{F'_{\theta}(x_{(i)},\alpha,\beta,\theta)-F'_{\theta}(x_{(i-1)},\alpha,\beta,\theta)}{F(x_{(i)},\alpha,\beta,\theta)-F(x_{(i-1)},\alpha,\beta,\theta)}\right]=0$$

3.1 Least square estimates. Let $x_{(1)},x_{(2)},\ldots,x_{(n)}$ be the ordered sample of size n drawn the Kum-TEMWD population pdf. Then, the expectation of the empirical cumulative distribution function is defined as

$$E\left[F(X_{(i)})\right]=\frac{i}{n+1};i=1,2,\ldots,n \qquad (3.2)$$

The least square estimates (LSEs) $\hat{\alpha}_{LS},\hat{\beta}_{LS},\hat{a}_{LS},\hat{b}_{LS}$ and $\hat{c}_{LS}$ of α, β and θ are obtained by minimizing

$$Z(\alpha,\beta,\theta)=\sum_{i=1}^{n}\left(F(x_{(i)},\alpha,\beta,\theta)-\frac{i}{n+1}\right)^2$$

There for, $\hat{\alpha}_{LS},\hat{\beta}_{LS}$ and $\hat{\theta}_{LS}$ of α, β and θ can be obtained as the solution of the following system of equations:

$$\frac{\partial Z(\alpha,\beta,\theta)}{\partial\alpha}=\sum_{i=1}^{n}F_{\alpha}(x_{(i)},\alpha,\beta,\theta)\left(F(x_{(i)},\alpha,\beta,\theta)-\frac{i}{n+1}\right)=0$$

$$\frac{\partial Z(\alpha,\beta,\theta)}{\partial\beta}=\sum_{i=1}^{n}F_{\beta}(x_{(i)},\alpha,\beta,\theta)\left(F(x_{(i)},\alpha,\beta,\theta)-\frac{i}{n+1}\right)=0$$

$$\frac{\partial Z(\alpha,\beta,\theta)}{\partial\theta}=\sum_{i=1}^{n}F_{\theta}(x_{(i)},\alpha,\beta,\theta)\left(F(x_{(i)},\alpha,\beta,\theta)-\frac{i}{n+1}\right)=0$$

4 BAYES ESTIMATION

In this section, we developed the Bayes procedure for the estimation of the unknown model parameters based on observed sample x from beta Lindley distribution. In addition to having a likelihood function, the Bayesian needs a prior distributionfor parameter, which quantifiesthe uncertainty about parameter prior to having data. In many situations, existing knowledge may be difficult to summarise in the form of an informative prior. In such case, it is better to consider the non-informative prior for Bayesian analysis,(for more detail on the use of non-informative prior, see [1]).We take the non-informative priors for θ, α and β of the following forms

$$\pi(\theta) \propto \frac{1}{\theta}; \theta > 0 \tag{4.1}$$

$$\pi(\alpha) \propto M_1^{-1}; 0 < \alpha < M_1 \tag{4.2}$$

$$\pi(\beta) \propto M_2^{-1}; 0 < \beta < M_2 \tag{4.3}$$

It is to be noticed that the choice of M_1 and M_2 are unimportant and we can simply take

$$\pi_2(\alpha) \propto 1 \tag{4.4}$$

$$\pi(\beta) \propto 1 \tag{4.5}$$

Thus the joint posterior distribution of θ, α and β is given by

$$\pi(\alpha, \beta, \theta \mid x) = K \frac{\theta^{n-1} \exp\left(-\theta\beta \sum_{i=1}^{n} x_i\right)}{B^n(\alpha, \beta)(1+\theta)^{n\beta}} \prod_{i=1}^{n} \left[(1+\theta+\theta x_i)^{\beta-1}\left(1 - \frac{1+\theta+\theta x_i}{1+\theta} e^{-\theta x_i}\right)^{\alpha-1}\right] \tag{4.6}$$

Where, K is the normalizing costant.Under square error loss,the Bayes estimates of θ, α and β are the means of their marginal posteriors and defined as ,

$$\hat{\theta}_B = \int\int\int \theta\pi(\alpha, \beta, \theta \mid x)\, d\beta\, d\alpha\, d\theta \tag{4.7}$$

$$\hat{\alpha}_B = \int\int\int \alpha\pi(\alpha, \beta, \theta \mid x)\, d\beta\, d\theta\, d\alpha \tag{4.8}$$

and

$$\hat{\beta}_B = \int\int\int \beta\pi(\alpha, \beta, \theta \mid x)\, d\theta\, d\alpha\, d\beta \tag{4.9}$$

Respectively. It is not easy to calculate Bayes estimates through equantions(4.7), (4.8) and (4.9) and so the numerical approximation techniques are needed. Therefor we proposed the use of Monte Carlo Markow Chain (MCMC) techniques namely Gibbs sampler and Metropolis Hastings-it (MH) algorithm, see [6, 10, 2]. Since the conditional posteriors of the parameters can not be obtained in any standard forms,we,therefore used a hybrid MCMC strategy for drawing samples from the joint posterior of the parameters. To implement the Gibbs algorithm, the full conditional posteriors of α, β and θ are given by:

$$\pi(\alpha \mid \beta, \theta, x) \propto \frac{\Gamma^n(\alpha+\beta)}{\Gamma^n(\alpha)} \prod_{i=1}^{n} \left(1 - \frac{1+\theta+\theta x_i}{1+\theta} e^{-\theta x_i}\right)^{\alpha} \tag{4.10}$$

$$\pi(\beta \mid \alpha, \theta, x) \propto \frac{\Gamma^n(\alpha+\beta)}{\Gamma^n(\alpha)} \frac{\exp\left(-\theta\beta \sum_{i=1}^{n} x_i\right)}{(1+\theta)^{n\beta}} \prod_{i=1}^{n} (1+\theta+\theta x_i)^{\beta} \tag{4.11}$$

$$\pi_3(\theta \mid \alpha, \beta, x) \propto \frac{\theta^{n-1}\exp\left(-\theta\beta\sum_{i=1}^{n}x_i\right)}{(1+\theta)^{n\beta}}\prod_{i=1}^{n}\left[(1+\theta+\theta x_i)^{\beta-1}\left(1-\frac{1+\theta+\theta x_i}{1+\theta}e^{-\theta x_i}\right)^{\alpha-1}\right] \qquad (4.12)$$

The simulation algorithm, we followed is given by

Step 1. Set starting points, say $\alpha^{(0)}$, $\beta^{(0)}$ and $\theta^{(0)}$, then at $i-$th stage

Step 2. Using MH algorithm, Generate $\alpha^{(i)} \sim \pi_1(\alpha \mid \beta^{(i-1)}, \theta^{(i-1)}, x)$

Step 3. Using MH algorithm, Generate $\beta^{(i)} \sim \pi_2(\beta \mid \alpha^{(i)}, \theta^{(i-1)}, x)$

Step 4. Using MH algorithm, Generate $\theta^{(i)} \sim \pi_3(\theta \mid \alpha^{(i)}, \beta^{(i)}, x)$

Step 5. Repeat steps $2-4$, $M(=20000)$ times to get the samples of size M from the corresponding posteriors of interest.

Step 6. Obtain the Bayes estimates of α, β and θ using the following formulae

$\hat{\alpha} = \frac{1}{M-M_0}\sum_{j=M_0+1}^{M}\alpha^{(j)}$, $\hat{\beta} = \frac{1}{M-M_0}\sum_{j=M_0+1}^{M}\beta^{(j)}$ and $\hat{\theta} = \frac{1}{M-M_0}\sum_{j=M_0+1}^{M}\theta^{(j)}$ respectively, where $M_0(\approx 5000)$ is the burn-in period of the generated Markow chains.

Step 7. Obtain the $100\times(1-\gamma)\%$ HPD credible intervals for α, β and θ by applying the methodology of [3].The HPD credible inervals for α, β and θ are $(\alpha_{(j^*)}, \alpha_{(j^*+[(1-\gamma)M])})$, $(\beta_{(j^*)}, \beta_{(j^*+[(1-\gamma)M])})$ and $(\theta_{(j^*)}, \theta_{(j^*+[(1-\gamma)M])})$ respectively. Where, j^* is chosen such that

$$\alpha_{(j^*+[(1-\gamma)M])} - \alpha_{(j^*)} = \min_{1\le j\le M-[(1-\gamma)M]}\left(\alpha_{(j+[(1-\gamma)M])} - \alpha_{(j)}\right)$$

$$\beta_{(j^*+[(1-\gamma)M])} - \beta_{(j^*)} = \min_{1\le j\le M-[(1-\gamma)M]}\left(\beta_{(j+[(1-\gamma)M])} - \beta_{(j)}\right)$$

$$\theta_{(j^*+[(1-\gamma)M])} - \theta_{(j^*)} = \min_{1\le j\le M-[(1-\gamma)M]}\left(\theta_{(j+[(1-\gamma)M])} - \theta_{(j)}\right)$$

Here, $[x]$ denotes the largest integer less than or equal to x.

Note that there have been several attempts made to suggesting the proposal density for the target posterior in the implementation of MH algorithm. By reparameterizing the posterior on the entire real line, [18] and [17] have suggested to use the normal approximation of the posterior as a proposal candidate in MH algorithm. Alternatively, it is also realistic to have the thought of using the truncated normal distribution without reparameterizing the original parameters. Therefore, we proposed the use of the truncated normal distribution as the proposal kernel to the target posterior.

5 SIMULATION STUDY

This subsection deals whith the comparisons study of the proposed estimators in term of their mean squareerror on the basis of simulates sample from pdf of a three parameters Lindely destribution whith varying sample sizes.For this purpose, we take $\alpha = 5, \beta = 3$ and $\theta = 1$ arbitrarily and $n = 10, 20, \ldots, 100$. All the algorithms are coded in R.

We calculate MLEs,LSEs,MPSs and Bayes estimators of α,β and θ based on each generated sample.This proses is repeated 1000 of times, and average estimates and corresponding mean square errors are computed and also reported in Table 1,2,3.

From Table 1,2 and 3 it can be clearly observed that as sample size increases the mean square error decreases, it proves the consistency of the estimators.The Bayesian and the least square estimator α, β, θ is superior than that of obtained by MLEs procedure and PSE.

Table 1: Estimates and mean square errors in second row of each cell for α

MLE	LSE	PSE	Bayes
5.4442	5.9901	5.7470	5.9901
1.3727	1.4215	0.7798	1.4215
5.2293	5.9866	5.8240	5.9866
0.4945	0.5933	0.3969	0.5933
5.1573	5.9886	5.8583	5.8876
0.2498	0.2682	0.2102	0.1764
5.1215	4.9952	4.8805	4.9912
0.1654	0.2232	0.1385	0.1232
5.0965	4.9934	4.8932	4.9923
0.1324	0.1710	0.1234	0.1610
5.0804	4.9954	4.9031	4.9956
0.0998	0.1379	0.0931	0.1189
5.0711	4.9973	4.9140	4.9975
0.0852	0.1175	0.0805	0.1011
5.0553	4.9937	4.9138	0.4997
0.0691	0.1009	0.0681	0.9120
5.0511	4.9944	4.9219	4.9981
0.0619	0.0889	0.0610	0.064
5.0481	4.9943	4.9290	4.9982
0.0565	0.0785	0.0557	0.0543

Table 2: Estimates and mean square errors in second row of each cell for β

MLE	LSE	PSE	Bayes
3.0175	3.0404	3.0532	3.0175
0.1300	0.1160	0.0873	0.1300
3.0022	3.0185	3.0302	3.0022
0.0405	0.0489	0.0422	0.0507
2.9949	2.0060	3.0181	2.9949
0.0407	0.0299	0.0269	0.0307
2.9952	2.0034	3.0146	2.9952
.03001	0.0223	0.0202	0.0224
2.9954	3.0017	3.0123	2.9954
0.0184	0.0181	0.0169	0.0184
2.9956	3.0014	3.0106	2.9956
0.0148	0.0149	0.0137	0.0148
2.9966	3.0020	3.0101	2.9966
0.0125	0.0129	0.0118	0.0125
2.9978	3.0027	3.0100	2.9978
0.0106	0.0110	0.0100	0.0106
2.9992	3.0038	3.0104	2.9992
0.0095	0.0099	0.0090	0.0095
2.9982	3.0025	3.0087	2.9982
0.0087	0.0091	0.0083	0.0087

Table 3: Estimates and mean square errors (in IInd row of each cell) for θ

MLE	LSE	PSE	Bayes
1.0234	1.0202	1.0532	1.0175
0.1312	0.1425	0.0873	0.1000
1.0142	1.0179	1.0302	1.0022
0.0415	0.0478	0.0422	0.0477
0.9989	1.0054	1.0181	0.9949
0.0417	0.0219	0.0269	0.0207
0.9934	1.0084	1.0146	0.9952
0.0312	0.0254	0.0202	0.0211
0.9875	1.0021	1.0132	0.9954
0.0184	0.0194	0.0169	0.0165
0.9956	1.0014	1.0111	0.9956
0.0148	0.0149	0.0137	0.0148
0.9966	1.0020	1.0101	0.9966
0.0125	0.0121	0.0118	0.0120
0.9978	1.0027	1.0100	0.9978
0.0106	0.0104	0.0100	0.0106
0.9992	1.0038	1.0104	0.9992
0.0095	0.0089	0.0090	0.0054
0.9982	1.0025	1.0087	0.9982
0.0087	0.0061	0.0083	0.0047

References

1. Berger, J. O. (1985). Statistical decision theory and Bayesian analysis. Springer.
2. Brooks, S. (1998). Markov chain Monte Carlo method and its application. Journal of the royal statistical society: series D (the Statistician), 47(1), 69-100.
3. Chen, M. H., and Shao, Q. M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. Journal of Computational and Graphical Statistics, 8(1), 69-92. with applications to lifetime data *Stat Papers* 54:133–161.
4. Cordeiro G.M., Nadarajah S.,(2011). Closed form expressions for moments of a class of Beta generalized distributions, Brazilian Journal of Probability and Statistics 25 , 14-33.
5. Gradshteyn IS, Ryzhik IM (2000) Table of integrals, series, and products. Academic Press, San Diego
6. Gelfand, A. E., and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. Journal of the American statistical association, 85(410), 398-409.
7. Johnson, N. L., Kotz, S and N. Balakrishnan (1995). Continuous Univariate Distribution, V(2), 2nd edition, New York, Wiley.
8. Gusmao FRS, Ortega EMM, Cordeiro GM (2011) The generalized inverse Weibull distribution. Statistical Papers 52:591-619.
9. Ghitany M.E., Atieh B., Nadarajah S.,Lindley distribution and its application,*Mathematics and Computers in Simulation*78, 493-506.
10. Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. Biometrika, 57(1), 97-109.
11. Jones MC (2004) Families of distributions arising from distributions of order statistics. Test 13:1-43.

12. Merovci F., and Sharma V. K., The Beta-Lindley Distribution: Properties and Applications, Journal of Applied Mathematics, vol. 2014, Article ID 198951, 10 pages, 2014. doi:10.1155/2014/198951

13. Nadarajah S., Bakouch H.S., and Tahmasbi R., (2011). A generalized Lindley distribution, Sankhya B - Applied and Interdisciplinary Statistics,73:331-359