SOME METHODS OF ESTIMATION FOR LINDLEY-EXPONENTIAL DISTRIBUTION

***ARBËR QOSHJA.**

University of Tirana, Albania Faculty of Natural Science,Department of Applied Mathematics

e-mail: qoshjaa@gmail.com

Abstract

Recently, Bhati, Malik and Vaman (Metron, 73(3), 335-357, 2015) introduced two-parameter Lindley-Exponential distribution. It is observed that this distribution can be used quite effectively in analyzing lifetime data. Different estimation procedures have been used to estimate the unknown parameters and their performances are compared using Monte Carlo simulations.

Keyword: Lindley-Exponential Distribution; Maximum likelihood estimaton Least squares estimators; Weighted least squares estimators; Simulations.

Përmbledhje

Kohët e fundit, Bhati, Malik and Vaman (Metron, 73(3), 335-357, 2015) prezantuan shpërndarjen Lindley-Eksponenciale me dy parametra. Eshtë vënë re se kjo shpërndarje ka efekt në analizën e të dhënave të jetëgjatësisë. Janë përdorur disa metoda vlerësimi për parametrat e kësaj shpërndarje si dhe kemi krahasuar përfundimet e tyre duke përdorur simulimet Monte Karlo.

Fjalëkyçe: Shpërndarja Lindley-Eksponenciale, metoda e përgjasisë maksimale, metoda e katrorëve më të vegjel, metoda e katrorëve më të vegjël sipas peshave, Simulimet.

Introduction

The exponential distribution was the first widely used lifetime distribution model in areas ranging from studies on the lifetimes of manufactured items to research involving survival or remission times in chronic diseases. The main reason for its wide applicability as lifetime model is partly because of the availability of simple statistical methods for it and partly because it appeared suitable for representing the lifetimes of many things such as various types of manufactured items.
Lindley distribution is a mixture of exponential θ and $\gamma(2,\theta)$ distributions.

Recently, one parameter Lindley distribution has attracted the researchers for its use in modelling lifetime data, and it has been observed in several papers that this distribution has performed excellently. More details on the Lindley distribution can be found in Ghitany et al.
Some of the advances in the literature of Lindley distribution are given by Ghitany et al. (2011) who has introduced a two-parameter weighted Lindley distribution and has pointed that Lindley distribution is particularly useful in modeling biological data from mortality studies. Bakouch et al. (2012) have come up with extended Lindley (EL) distribution, Adamidis and Loukas (1998) have introduced exponential geometric (EG) distribution.

Shanker *et al.* (2013) have introduced a two-parameter Lindley distribution. Zakerzadeh *et al.* (2012) have proposed a new two parameter lifetime distribution: model and properties. M.K. Hassan (2008) has introduced convolution of Lindley distribution. Ghitany *et al.* (2013) worked on the estimation of the reliability of a stress-strength system from power Lindley distribution.

Definition 1. A random variable X is said to have Lindley distribution with parameter θ if its probability density function is defined as:

$$f_X(x,\theta)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x},\ x>0,\theta>0, \qquad (1)$$

with cumulative distribution function

$$F_X(x,\theta)=1-\frac{e^{-\theta x}(1+\theta+\theta x)}{\theta+1},\ x>0,\theta>0. \qquad (2)$$

In 2015, Bahti et al. introduced a new modified exponential distribution with two-parameters called Lindley-Exponential distribution with pdf and cdf, given by, respectively

$$f_X(x,\theta,\lambda)=\frac{\theta^2\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\theta-1}(1-\ln(1-e^{-\lambda x}))}{\theta+1}, \qquad (3)$$

and

$$F_X(x,\theta,\lambda)=\frac{(1-e^{-\lambda x})^{\theta}(1+\theta-\theta\ln(1-e^{-\lambda x}))}{\theta+1}. \qquad (4)$$

Maximum likelihood estimates

Let $x_1, x_2, ..., x_n$ be a (iid) observed random sample of size n from a new modified exponential distribution . Then, the likelihood function based on observed sample is defined as

$$L=\frac{\theta^{2n}\lambda^n e^{-\lambda\sum_{i=1}^{n}x_i}\prod_{i=1}^{n}(1-e^{-\lambda x_i})^{\theta-1}(1-\ln(1-e^{-\lambda x_i}))}{(\theta+1)^n} \qquad (5)$$

The log-likelihood function corresponding to (5) is given by

$$\ln L = 2n\ln(\theta) + n\ln(\lambda) - \lambda\sum_{i=1}^{n} x_i - n\ln(\theta+1)$$
$$+(\theta-1)\sum_{i=1}^{n}\ln(1-e^{-\lambda x_i}) + \sum_{i=1}^{n}\ln(1-\ln(1-e^{-\lambda x_i}))$$

The maximum likelihood estimates $\hat{\theta}$ and $\hat{\lambda}$ of θ and λ can be obtained as the solution of the following nonlinear equations $\frac{\partial \ln L}{\partial \theta} = \frac{2n}{\theta} - \frac{n}{\theta+1} + \sum_{i=1}^{n}\ln(1-e^{-\lambda x_i}) = 0$

$$\frac{\partial \ln L}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i + (\theta-1)\sum_{i=1}^{n}\frac{x_i e^{-\lambda x_i}}{1-e^{-\lambda x_i}} - \sum_{i=1}^{n}\frac{x_i e^{-\lambda x_i}}{(1-e^{-\lambda x_i})(1-\ln(1-e^{-\lambda x_i}))} = 0$$

It is usually more convenient to use nonlinear optimization algorithms such as the quasi-Newton algorithm to numerically maximize the sample likelihood function.

Maximum product spacing estimates

The maximum product spacing (MPS) method has been proposed by Cheng and Amin which is based on an idea that the differences 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

$$D_i = \int_{x_{(i-1)}}^{x_{(i)}} f(x,\theta,\lambda)dx, \quad i = 1,2,3...,n+1. \qquad (6)$$

where, $F(x_{(0)},\theta,\lambda) = 0$ and $F(x_{(n+1)},\theta,\lambda) = 0$. The MPS estimator $\hat{\theta}$ and $\hat{\lambda}$ of θ and λ is obtained by maximizing the geometric mean (GM) of the differences. Substituting (3) in (6) and taking logarithm of the above expression, we will have

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

$$=\frac{1}{n+1}\sum_{i=1}^{n+1}\log\left[\frac{(1-e^{-\lambda x_{(i)}})^{\theta}(1+\theta-\theta\ln(1-e^{-\lambda x_{(i)}}))}{\theta+1}\right]$$

$$-\frac{1}{n+1}\sum_{i=1}^{n+1}\log\left[\frac{(1-e^{-\lambda x_{(i-1)}})^{\theta}(1+\theta-\theta\ln(1-e^{-\lambda x_{(i-1)}}))}{\theta+1}\right]$$

The MPS estimator $\hat{\theta}$ and $\hat{\lambda}$ of θ and λ can be obtained as the simultaneous solution of the following non-linear equation:

$$\frac{\partial \log GM}{\partial\theta}=\frac{1}{n+1}\sum_{i=1}^{n+1}\left[\frac{F'_{\theta}\left(x_{(i)},\theta,\lambda\right)-F'_{\theta}\left(x_{(i-1)},\theta,\lambda\right)}{F\left(x_{(i)},\theta,\lambda\right)-F\left(x_{(i-1)},\theta,\lambda\right)}\right]$$

$$\frac{\partial \log GM}{\partial\lambda}=\frac{1}{n+1}\sum_{i=1}^{n+1}\left[\frac{F'_{\lambda}\left(x_{(i)},\theta,\lambda\right)-F'_{\lambda}\left(x_{(i-1)},\theta,\lambda\right)}{F\left(x_{(i)},\theta,\lambda\right)-F\left(x_{(i-1)},\theta,\lambda\right)}\right]$$

where, $F'_{\theta}\left(x_i,\theta,\lambda\right)=-\dfrac{\theta^2(1-e^{-\lambda x})^{\theta}(1-\ln(1-e^{-\lambda x}))}{(\theta+1)^2}$ and

$$F'_{\lambda}\left(x_i,\theta,\lambda\right)=\frac{\theta^2 xe^{-\lambda x}(1-e^{-\lambda x})^{\theta-1}(1-\ln(1-e^{-\lambda x}))}{\theta+1}$$

Method of Cramer-von-Mises

The Cramer-von - Mises uses the integral of the squared difference between the empirical and the estimated distribution function. The Crame-von Mises estimates $\hat{\theta}$ and $\hat{\lambda}$ of θ and λ *is* obtained by minimizing, with respect to θ and λ the function:

$$C(\theta,\lambda)=\frac{1}{12n}+\sum_{i=1}^{n}\left(F(x_i\,|\,\theta,\lambda)-\frac{2i-1}{2n}\right)^2$$

$$=\frac{1}{12n}+\sum_{i=1}^{n}\left(\frac{(1-e^{-\lambda x_{(i)}})^{\theta}(1+\theta-\theta\ln(1-e^{-\lambda x_{(i)}}))}{\theta+1}-\frac{2i-1}{2n}\right)^2$$

Least square estimators

The least square estimators and weighted least square estimators were proposed by Swain, Venkataraman and Wilson to estimate the parameters of Beta distributions.

Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the ordered sample of size n drawn the new modified exponential population (3). Then, the expectation of the empirical cumulative distribution function is defined as

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

The least square estimates (LSEs) $\hat{\theta}$ and $\hat{\lambda}$ of θ and λ are obtained by minimizing:

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

Therefore, $\hat{\theta}$ and $\hat{\lambda}$ of θ and λ can be obtained as the solution of the following equation:

$$\frac{\partial Z(\theta,\lambda)}{\partial\theta} = \sum_{i=1}^{n} F'_{\theta}(x_{(i)}\,|\,\theta,\lambda)\left(F(x_{(i)}\,|\,\theta,\lambda) - \frac{i}{n+1}\right)$$

$$= \sum_{i=1}^{n} \frac{\theta^2 (1-e^{-\lambda x})^{\theta}(1-\ln(1-e^{-\lambda x}))}{(\theta+1)^2}\left(F(x_{(i)}\,|\,\theta,\lambda) - \frac{i}{n+1}\right) = 0$$

$$\frac{\partial Z(\theta,\lambda)}{\partial\lambda} = \sum_{i=1}^{n} F'_{\lambda}(x_{(i)}\,|\,\theta,\lambda)\left(F(x_{(i)}\,|\,\theta,\lambda) - \frac{i}{n+1}\right) =$$

$$= \sum_{i=1}^{n} \frac{\theta^2 x e^{-\lambda x}(1-e^{-\lambda x})^{\theta-1}(1-\ln(1-e^{-\lambda x}))}{\theta+1}\left(F(x_{(i)}\,|\,\theta,\lambda) - \frac{i}{n+1}\right) 0$$

Weighted least square estimators

The weighted least square estimators of the unknown parameters can be obtained by minimizing

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

with respect to θ and λ. The weights w_j are equal to $\frac{1}{V(X_{(i)})} = \frac{(n+1)^2(n+2)}{j(n-j+1)}$.

Simulation algorithms and study

To generate a random sample of size *n* from new modified exponential distribution, we follow the following steps:

1. Set n, $\Theta = (\theta, \lambda)$ and initial value x^0.
2. Generate $U \sim Unifrom\,(0,1)$.
3. Update x^0 by using the Newton's formula

$x^* = x^0 - R(x^0, \Theta)$

where, $R(x^0, \Theta) = \frac{F_X(x^0, \Theta) - U}{f_X(x^0, \Theta)}$, $F_X(.)$ and $f_X(.)$ are cdf and pdf of new modified exponential distribution, respectively.

4. If $|x^0 - x^*| \leq \epsilon$, (very small, $\epsilon > 0$ tolerance limit), then store $x = x^*$ as a sample from new modified exponential distribution.
5. If $|x^0 - x^*| > \epsilon$, then, set $x^0 = x^*$ and go to step 3.
6. Repeat steps 3-5, n times for $x_1, x_2, \dots, x_n$ respectively.

This subsection deals with the comparisons study of the proposed estimators in terms of their mean square error on the basis of simulates sample from pdf of new modified exponential distribution with varying sample sizes. For this purpose, we take θ = 3, λ = 0.7 arbitrarily and n = 10, 20, . . . , 50. All the algorithms are coded in R, a statistical computing environment and we used algorithm given above for simulations purpose. We calculate MLEs, LSEs, WLSEs, MPSs and CMEs estimators of θ and λ 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 and 2. From Table 1 and Table 2 it can be observed that as sample size increases the mean square error decreases, it proves the consistency of the estimators. The least square estimator of parameter θ is superior than the others methods of estimation, on other hand, the maximum likelihood estimator of λ performs well in the comparison to others methods.

Table 1. Estimates and mean square errors (in IInd row of each cell) of the proposed estimators with varying sample size

θ					
n	MLE	LSE	WLSE	MPS	CME
10	2.8442	3.1442	3.6470	3.3442	2.6901
	1.1717	1.0717	1.3215	1.0798	1.2717
20	2.9193	3.0893	3.2240	3.2293	2.7866

	0.3925	0.32145	0.5969	0.3925	0.4933
30	2.94573	3.0173	3.2283	3.1273	2.8886
	0.2298	0.2098	0.3902	0.2598	0.3182
40	2.97235	3.0115	3.2005	3.1115	2.9052
	0.1590	0.1420	0.2385	0.1590	0.3119
50	3.0965	3.0110	3.0932	3.1065	2.9134
	0.1252	0.0552	0.2141	0.1252	0.1982

Table 2. Estimates and mean square errors (in IInd row of each cell) of the proposed estimators with varying sample size

λ					
n	MLE	LSE	WLSE	MPS	CME
10	0.2975	0.5442	0.6470	0.3442	0.6801
	0.5479	1.2717	1.0798	1.4717	1.3215
20	0.87401	0.5642	0.5310	0.3991	0.5631
	0.3717	0.9717	0.5215	0.6098	0.4717
30	0.77401	0.5942	0.5814	0.4897	0.6331
	0.2803	0.6217	0.3098	0.2917	0.4215
40	0.73121	0.6342	0.6411	0.5394	0.6402
	0.1842	0.4217	0.2498	0.2117	0.2915
50	0.70121	0.6402	0.6715	0.6092	0.6521
	0.0812	0.2217	0.1490	0.2007	0.1905

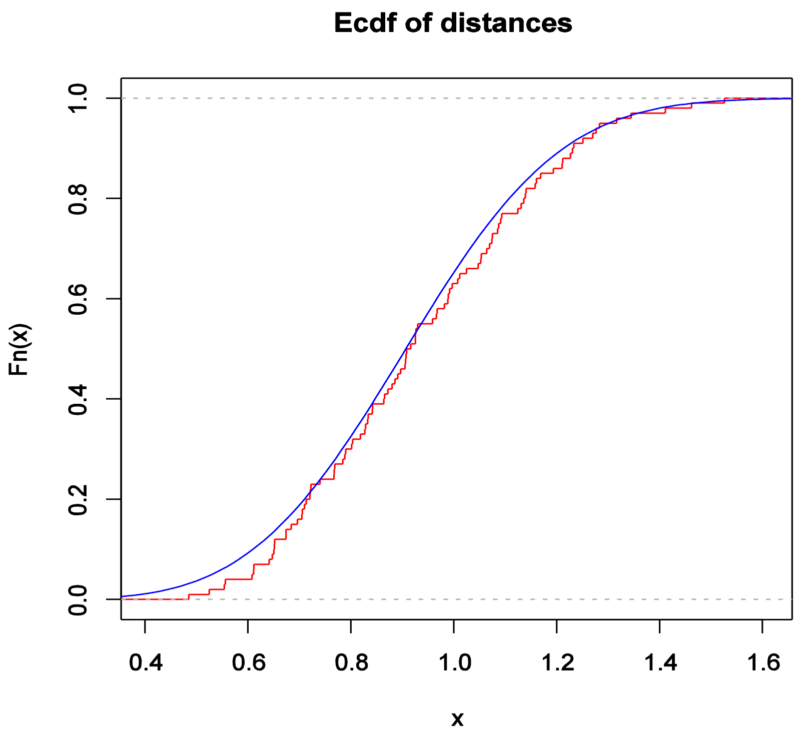

Figure 1. Empirical cdf plots for the fits of the LE distribution for simulated data.

R code for simulated data, average bias and average mean square error for maximum likelihood estimation

```
rm(list=ls(all=TRUE))
library(gsl) # Package for Lambert Function
n=1000
simulated_data=matrix(2*n, nrow=n, ncol=2, byrow=TRUE)
for(j in 1:n){
ld0=0.7;th0=3;m=50
u= runif(m,0,1)
a= -exp(-(1+th0))*(1+th0)*u
Lw=lambert_Wm1(a)# Lambert Function
x= -(1/ld0)*log(1-(-Lw/(u+u*th0))^(-1/(th0)))
mean(x)
```

```
var(x)
negloglike = function(p)
{-m*log(p[1])-2*m*log(p[2])+m*log(1+p[2])-(p[2]-1)*sum(log(1-exp(-
x*p[1])))+(p[1]*sum(x))-sum(log(1-log(1-exp(-x*p[1])))) }
out = nlm(negloglike,c(ld0,th0),hessian=TRUE)
simulated_data [j,1]=out$estimate[1]
simulated_data [j,2]=out$estimate[2] }
mean(simulated_data [,1])
mean(simulated_data [,2])
bias.th0 = mean(simulated_data [,2]-th0)
bias.ld0 = mean(simulated_data [,1]-ld0)
mse.th0 = (1/n)*sum((simulated_data [,2]-th0)^2)
mse.ld0 = (1/n)*sum((simulated_data [,1]-ld0)^2)
bias.th0
bias.ld0
mse.th0
mse.ld0
out
```

References

Adamidis K., and Loukas S.,(1998): A lifetime distribution with decreasing failure rate, Statistics and Probability Letters, Vol(39), 35-42

Arnold B.C., Balakrishnan N. and Nagaraja H.N.(2013): A First Course in Order Statistics, Wiley, New York, 1992

Bakouch H. S., Al-Zahrani B. M., Al-Shomrani A. A., Marchi V. A., and Louzada F.,(2012) An extended Lindley distribution, Journal of the Korean Statistical Society, Vol(41), 75-85

Bhati, D., Malik, M. A., Vaman, H. J. (2015). Lindley-Exponential distribution: properties and applications. Metron, 73(3), 335-357

Cheng, R. C. H., and N. A. K. Amin.(1983) . Estimating parameters in continuous univariate distributions with a shifted origin. Journal of the Royal Statistical Society. Series B (Methodological) (1983): 394-403

Ghitany M. E., Alqallaf F., Al-Mutairi D. K., and Husain H. A., (2011) A two-parameter weighted Lindley distribution and its applications to survival data, Mathematics and Computers in Simulation, Vol. (81), no. 6,1190-1201

Ghitany M. E., Al-Mutairi D. K., and Aboukhamseen S. M. , (2013): Estimation of the reliability of a stress-strength system from power Lindley distributions, Communications in Statistics - Simulation and Computation, Vol (78), 493-506.

Ghitany M. E., Atieh B., and Nadarajah, S., (2008): Lindley distribution and its application, Mathematics and Computers in Simulation, Vol (78), 493-506

Hassan M.K., (2014): On the Convolution of Lindley Distribution, Columbia International Publishing Contemporary Mathematics and Statistics, Vol. (2) No. 1,47-54

Jo´rda P.,(2010): Computer generation of random variables with Lindley or Poisson-Lindley distribution via the Lambert W function, Mathematics and Computers in Simulation, Vol(81), 851-859

Lindley D. V.,(1958): Fiducial distributions and Bayes theorem, Journal of the Royal Statistical Society, Series B (Methodological),102-107

Laio, F. (2004): Cramer–von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters. Water Resources Research, 40(9)

Mahmoudi E., and Zakerzadeh H., (2010): Generalized Poisson Lindley distribution, Communications in Statistics: Theory and Methods, Vol (39), 1785-1798

Ranneby, B. (1984): The maximum spacing method. An estimation method related to the maximum likelihood method. Scand. J. Statist. 11, 93-112

Shanker R., Sharma S., and Shanker R., (2013): A Two-Parameter Lindley Distribution for Modeling Waiting and Survival Times Data, Applied Mathematics, Vol (4), 363-368

Zakerzadeh, H. and Mahmoudi, E.,(2012): A new two parameter lifetime distribution: model and properties. arXiv:1204.4248 v1 [stat.CO]

Swain, J. J., Venkatraman, S., & Wilson, J. R. (1988): Least-squares estimation of distribution functions in Johnson's translation system. Journal of Statistical Computation and Simulation, 29(4), 271-297