LINEAR AND NONLINEAR REGRESSION OF EXPONENTIAL DISTRIBUTION

El desoky E. Afify, Dept. of mathematics, faculty of Eng. Shibeen El kom, Menoufyia Univ. EGYPT.

Key word: Least trimmed squares, nonlinear least squares method, Exponential distribution, quantile method, median, moments.

Abstract:

In this paper, we estimate the parameters of the exponential distribution by least trimmed squares (TLS), quantile estimation (QE) method, nonlinear least squares method (NLS) and median - moment method we compare between these methods of estimation by using sum absolute error of estimates (SAEE) and mean square errors (MSE).

Introduction:

Exponential distribution plays an important role in life testing and reliability problems. If failure rate appears to be more or less constant, it would be the most appropriate choice. Exponential distribution occurs in several context such as waiting time problems. The probability density function of two parameter exponential distribution is given by

$$f(t,\alpha,\beta) = \frac{1}{\alpha}\exp\{-(t-\beta)/\alpha\} \qquad \theta < t < \infty, \alpha > 0$$
$$= 0 \qquad \textit{other wise} \qquad (1)$$

where the parameters α and β are interpreted as measure of guarantee and failure rate respectively. Cohen and Helm (1973) used (BLUE), (MLE), ME, MVUE and MME to estimate the parameters of the exponential distribution. Jani (1991) obtained the scale parameter of the exponential distribution by using shrinkage estimators. Siu and Tso (1996) estimated the parameters of the exponential distribution by the use of shrinkage and maximum likelihood estimators. Peter (1974) used robust M-estimation method for the scale parameter exponential distribution. Bayes method of estimation was used by Rajesh et al (1995) to estimate the parameters of the exponential distribution.This paper introduces least trimmed squares, quantile estimators and nonlinear least squares to estimate the parameters of the exponential distribution. From (1) the cumulative distribution function is given by:

$$F(t) = 1 - e^{-(t-\beta)/\alpha}$$

or $\qquad e^{-(t-\beta)/\alpha} = 1 - F(t)\,.$ (2)

Taking the logrithem of both sides of (2), we have

$$t = \beta - \alpha \log[1 - F(t)]\ . \qquad (3)$$

Equation (3) can be repersented by y = a +bx, where y= t , a= β , b= -α

and x =log[1- F(t)].

Least Trimmed Squares (LTS) Method:

The least trimmed squares was introduced by Barnett and Lewis (1984) which was defined by :

$$Min\ \sum_{i=1}^{h} (r^2)_i$$

where h=n/2+p/2 = (n+p)/2 , and n is the number of observations, p is the number of parameters and $(r^2)_1 \le (r^2)_2 \le \ldots.. \le (r^2)_n$ are the ordered squared residuals (the residuals are first squared and then ordered).

Quantile Estimators:

From Enrique and Hadi (1997), Quantile estimation (QE) can be summarized as follows: Let $\alpha = \{\alpha_1, \alpha_2, \ldots., \alpha_r\}$ be the parameters to be estimated and $t_{1:n}, t_{2:n}, \ldots, t_{n:n}$ be the order statistics obtained from a random sample from F(t;α), then one can write

$$F(t_{I:n}, \alpha) = P_{i:n}$$

or

$t_{i:n} = F^{-1}(P_{i:n}, \alpha)$ where $P_{I::n} = (i-\alpha)/(n+\beta)$ and $Pi_{:n} = i/(n+1)$ is the emprical estimator.

In the two parameter exponential distribution (1) the cdf is

$$F(t;\alpha,\beta) = 1 - e^{-(t-\beta)/\alpha}$$

or $\qquad P_{i:n} = 1 - e^{-(t-\beta)/\alpha}$

$$e^{-(t-\beta)/\alpha} = 1 - P_{i\ :n}$$

Taking the logarithm of both sides, we have

$$(t_{i::n} - \beta) = -\alpha \log(1 - P_{i::n})$$

There are two parameters, so two equations are needed, as follows:

$$\begin{aligned} t_{i:n} &= \beta - \alpha \log(1 - p_{i:n}) \\ t_{j:n} &= \beta - \alpha \log(1 - p_{j:n}) \end{aligned} \qquad (4)$$

where i < j, eliminating β from the system (4), the estimates of α and β are given by

$$\hat{\alpha}_{i,j} = \frac{t_{i:n} - t_{j:n}}{\log(1-p_{j:n}) - \log(1-p_{i:n})} \quad (5)$$

$$\hat{\beta}_{i,j} = t_{i:n} + \hat{\alpha}_{i,j}\log(1-p_{i:n}) \quad (6)$$

If there are r parameters to be estimated, then there are ${}^{n}C_{r}$ elemental estimates.

In our case, there are ${}^{n}C_{2}$ elemental estimates, for choosing $P_{i:n}$, i=1,2,…,n overall estimates for α and β are obtained as

$$\hat{\alpha}_{QE} = median(\hat{\alpha}_{i,j}) \quad (7)$$

$$\hat{\beta}_{QE} = median(\hat{\beta}_{i,j}) \quad (8)$$

where QE repersents the quantile estimate.

As a special case, let i=n/2 and j=3n/4 in the first and second equations of (4) we have

$$t_{\frac{n}{2}:n} = \beta - \alpha\, Log(1/2) \quad (4a)$$

$$t_{\frac{3n}{4}:n} = \beta - \alpha\, Log(1/4) \quad (4b)$$

Solving (4a) and (4b), we obtain the estimates of α and β as

$$\hat{\alpha} = (t_{\frac{n}{2}:n} - t_{\frac{3n}{4}:n}) / Log(1/2)$$

$$\hat{\beta} = 2t_{\frac{n}{2}:n} - t_{\frac{3n}{4}:n}$$

Nonlinear Least Squares (NLS):

From Sindney and Ference (1989) nonlinear least squares estimators can be obtained as follows. The normal equations for the problem of finding the parameters α and β to minimize

$$S(t_i;\alpha,\beta) = \sum_{i=1}^{n}\left(R_i(t) - \exp\{-(t_i-\beta)/\alpha\}\right)^2 \quad (9)$$

The resulting nonlinear normal equations for (9) obtained by taking first partial derivative with respect to the parameters α and β respectively, are

$$\sum_{i=1}^{n}\left[R_i(t) - \exp\{-(t_i-\beta)/\alpha\}\right]\left[\frac{1}{\alpha}\ \exp\{-(t_i-\beta)/\alpha\}\right] = 0$$

$$\sum_{i=1}^{n}[R_i(t) - \exp\{-(t_i-\beta)/\alpha\}][\frac{t_i-\beta}{\alpha^2}\exp\{-(t_i-\beta)/\alpha\}] = 0$$

Newton's method for system uses then nXn Jacobian matrix in the vector situation and substitute by the derivative with the inversion of the Jacobian matrix, this method for finding the solution of the parameters α and β to the nonlinear system of equations (9) has the form

$$B^k = B^{k-1} - \left[J(B^{k-1})\right]^{-1} F(B^{k-1})$$

For k ≥ 1, given the initial approximation $B^{(0)}$ to the solution B, the initial values of B can be assumed to be zero, where J(t) is the first order derivative of R(t) in the matrix form is given by:

$$J(t)=\begin{bmatrix} \frac{\partial R_1(t)}{\partial \alpha} & \frac{\partial R_1(t)}{\partial \beta} \\ \frac{\partial R_2(t)}{\partial \alpha} & \frac{\partial R_2(t)}{\partial \beta} \end{bmatrix}$$

where $R(t) = \exp\{-(t-\beta)/\alpha\}$, $\frac{\partial R(t)}{\partial \beta} = \frac{1}{\alpha}\exp\{-(t-\beta)/\alpha\}$,

$$\frac{\partial R(t)}{\partial \alpha} = \frac{t-\beta}{\alpha^2}\exp\{-(t-\beta)/\alpha\}$$

and $\frac{\partial R(t)}{\partial a_k}$ is the partial derivative of the reliability function R(t) w.r.t. k^{th} parameter at i^{th} data point.

Median – First Order Statistics Method (MOS):

The median of exponential distribution is given by:

$$t_{med} = \beta + \alpha \log 2 \tag{10}$$

If it is assumed that the observed first order statistic $t_{(1)}$ is equal to its expected value

That is $E(t_{(1)}) = t_{(1)}$ or

$$x_{(1)} = \beta + \alpha / n \tag{11}$$

Solving (10) and (11), will give the following estimates:

$$\hat{\alpha} = \frac{t_{med} - t_{(1)}}{\log 2 - 1/n}$$

$$\hat{\beta} = \frac{t_{med} - n t_{(1)} \log 2}{1 - n \log 2}$$

where t_{med} *and* $t_{(1)}$ are the median and the first order statistics of the sample respectively.

Goodness of Fit:

To compute the accuracy of parameter estimates and reliability function , we use the mean squared error (MSE) defined by

$$MSE = \frac{1}{n}\sum_{i=1}^{n}[\ R(t_i, \alpha, \beta) - \hat{R}(t_i, \alpha, \beta)]^2$$

where, R(t) is actual reliability function, $\hat{R}(t)$ is the estimated reliability function,and the sum absolute error of estimates (SAEE) is given by:

$$SAEE = \left|\frac{\alpha - \hat{\alpha}}{\alpha}\right| + \left|\frac{\beta - \hat{\beta}}{\beta}\right|$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the estimates of α and β respectively. R_i is the reliability function at the parameters, $\hat{R}_i$ is reliability function at the estimated parameters. The minimum value of MSE and SAEE indicate a goodness of the distribution.

Simulation Results:

To examine the goodness of fit of these methods, the mean square errors (MSE) and sum absolute errors of estimates (SAEE) for each method ware calculated using 1000 replications for each sample size. In this simulation study, we generate a data set for values of α and β and for samples 10, 20 and 30, the true values of pairs (α, β): (1,1), (1,2), (2,2). The generation of random sample by consider that if U is Uniform (0,1), then X= α + β Log (1-U) is exponential (α , β).

Conculusion:

The results are reported in the tables (1), (2) and (3), from the computations, we note that, the estimate of parameters from the LTS are too close to the true values, and the values of SAEE and MSE are very small. The parameter estimates from QE are close to the true values but not as the LTS estimates SAEE and MSE are greater than the corresponding values from LTS. The values SAEE and MSE from NLS are greater than the values from LTS and QE methods. Finally, the values of SAEE and MSE are greater than the corresponding values from LTS, QE and NLS. For the samples of large size, the values of SAEE and MSE obtained from MOS method are too large, LTS is the best method of all but MOS method takes little time on computer.

Table (1) The estimates for n =10

Method	True values α; β	Estimated values $\hat{\alpha}$	$\hat{\beta}$	SAEE	MSE
LTS	1,1	1	0.99999	0.0002	$71\text{x}10^{-10}$
	1,2	1	1.99999	0.00015	$71\text{x}10^{-10}$
	2,2	1.99999	1.99999	0.0001	$18\text{x}10^{-10}$
QE	1,1	0.9988	0.9987	0.0025	$11522\text{x}10^{-10}$
	1,2	0.9998	1.9998	0.0003	$284\text{x}10^{-10}$
	2,2	2.0012	1.9998	0.0007	$173\text{x}10^{-10}$
NLS	1,1	0.9987	0.9989	0.0024	$9438\text{x}10^{-10}$
	1,2	0.9996	1.9996	0.0006	$1135\text{x}10^{-10}$
	2,2	1.9976	1.9984	0.002	$5851\text{x}10^{-10}$
MOS	1,1	0.9028	0.9297	0.1675	$458411\text{x}10^{-10}$
	1,2	0.9028	1.9297	0.13235	$4584144\text{x}10^{-10}$
	2,2	0.9315	1.9479	0.5603	$4661486\text{x}10^{-10}$

Table (2) The estimates for n=20

Method	True values α; β	Estimated values $\hat{\alpha}$	$\hat{\beta}$	SAEE	MSE
LTS	1,1	1	0.9999	0.0001	$39\text{X}10^{-10}$
	1,2	1	1.9999	0.00005	$39\text{X}10^{-10}$
	2,2	2	1.9999	0.00005	$1\text{X}10^{-10}$
QE	1,1	0.9983	0.9974	0.0043	$38495\text{x}10^{-10}$
	1,2	0.9974	1.9887	0.00825	$565385\text{x}10^{-10}$
	2,2	1.9986	1.9977	0.00185	$7339\text{x}10^{-10}$
NLS	1,1	0.9988	1.9987	0.00185	$1119\text{x}10^{-9}$
	1,2	0.9988	1.9887	0.00185	$1119\text{x}10^{-9}$
	2,2	1.9996	1.9885	0.00595	$130864\text{x}10^{-10}$
MOS	1,1	0.9627	0.9237	.00594997	$130864\text{x}10^{-10}$
	1,2	0.9627	1.9237	.00594997	$130864\text{x}10^{-10}$
	2,2	1.9245	1.8486	.00594997	$130864\text{x}10^{-10}$

Table (3) The estimates for n=30

Method	True values α; β	Estimated values		SAEE	MSE
		$\hat{\alpha}$	$\hat{\beta}$		
LTS	1,1	0.9999	1	0.0001	$7\text{x}10^{-10}$
	1,2	0.9999	1.9999	0.00015	$67\text{x}10^{-10}$
	2,2	2.0001	1.9999	0.0001	$5\text{x}10^{-10}$
QE	1,1	0.9986	0.9985	0.0029	$14529\text{x}10^{-10}$
	1,2	0.9993	1.9982	0.0016	$15156\text{x}10^{-10}$
	2,2	1.9981	1.9881	0.0069	$143369\text{x}10^{-10}$
NLS	1,1	0.9999	0.9881	0.0029	$18357\text{x}10^{-10}$
	1,2	0.9983	1.9983	0.00255	$19382\text{x}10^{-10}$
	2,2	1.9972	1.9975	0.00265	$11201\text{x}10^{-10}$
MOS	1,1	0.993	0.9207	0.0863	$288\text{x}10^{-9}$
	1,2	0.993	1.9207	0.0466501	0.002287875
	2,2	1.986	1.8424	0.0858	$2261\text{x}10^{-9}$

Acknowledgements:

The auther would like to thank the editor of the Interstat journal Profeeesor Richard G. Krutchkoff for guidance and also thank Professor Rao Chaganty for his valuable comments, and thank Professor Tariq El-Ghonaimy, Tanta University, Egypt for assisting in the prepration of the manuscript.

REFERENCES

1) Barnett, V., and Lewis,T. (1984) Outliers in Statistical Data .(2nd ed), New York, John Wiley.
2) Cohen, A.C., and Helm F.R. (1973) Estimation in Exponential Distribution. Technometrics,Vol.15, 415 - 418.
3) Enrique, Castillo and Ali, S. Hadi (1997) Fitting the Generalized Pareto Disribution to Data . Journal of the American Statistical Association, Vol. 92, 1609 – 1620.
4) Jani, P.N.(1991) A Class of Shinkage Estimators for the Scale Parameter of the Exponential Distribution. IEEE, Trans. Reliability, Vol.40, 68-70.
5) Peter, F. Thall (1974) Huber – Sense Robust M- Estimation of a Scale Parameter, with Application to the Exponential Distribution. JASA, Vol.74, 147 – 152.

6) Rajesh Singh, Upadhyay S.K., and Singh,U. (1995) Bayes Estimators for Two Parameter Exponential Distribution. Commun. Statist. Theory Meth., Vol. 24, 227-240.
7) Sidney, Y. and Ferenc, S. (1989) An Introduction to Numerical Computations. Second Edition. Macmillan Publishing Company, New York. Coolier Macmillan Publishing, London.
8) Siu – Keung Tse and Geoffrey Tso (1996) Shrinkage Estimation of Reliability for Exponentially Distribution Lifetimes. Commun. Statist.- Simula., Vol. 25, 415-430.